

Reading and writing skills: Cognitive, emotional, creative, and digital approaches

Edited by

María Isabel de Vicente-Yagüe Jara, Elena Jiménez-Pérez,
Pedro García Guirao and Olivia López Martínez

Published in

Frontiers in Psychology
Frontiers in Education



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-4095-4
DOI 10.3389/978-2-8325-4095-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Reading and writing skills: Cognitive, emotional, creative, and digital approaches

Topic editors

Maria Isabel de Vicente-Yagüe Jara — University of Murcia, Spain

Elena Jiménez-Pérez — University of Malaga, Spain

Pedro García Guirao — WSB Universities, Poland

Olivia López Martínez — University of Murcia, Spain

Citation

de Vicente-Yagüe Jara, M. I., Jiménez-Pérez, E., García Guirao, P., López Martínez, O., eds. (2023). *Reading and writing skills: Cognitive, emotional, creative, and digital approaches*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4095-4

Table of contents

- 07 **Editorial: Reading and writing skills: cognitive, emotional, creative, and digital approaches**
María Isabel de Vicente-Yagüe Jara, Pedro García Guirao, Elena del Pilar Jiménez-Pérez and Olivia López Martínez
- 10 **Mechanisms for Interpretative Cooperation: Fan Theories in Virtual Communities**
José Manuel de Amo and Anastasio García-Roca
- 18 **Validation of the Writing Strategies Questionnaire in the Context of Primary Education: A Multidimensional Measurement Model**
Olga Arias-Gundín, Sara Real, Gert Rijlaarsdam and Paula López
- 27 **Beyond the Educational Context: Relevance of Intrinsic Reading Motivation During COVID-19 Confinement in Spain**
Raquel De Sixte, Inmaculada Fajardo, Amelia Mañá, Álvaro Jáñez, Marta Ramos, María García-Serrano, Federica Natalizi, Barbara Arfé and Javier Rosales
- 41 **Comparing the Effects of Stroke-Appearing and Stroke-Disappearing on Learning the Order of Strokes in Chinese Characters**
Jon-Chao Hong, Kai-Hsin Tai, Ming-Yueh Hwang and Pei-Hsin Lin
- 50 **Cognitive Strategies and Textual Genres in the Teaching and Evaluation of Advanced Reading Comprehension (ARC)**
Jesús-Nicasio García-Sánchez and Judit García-Martín
- 67 **By Toutatis! Trainee Teachers' Motivation When Using Comics to Learn History**
Juan Ramón Moreno-Vera, Santiago Ponsoda-López de Atalaya and Rubén Blanes-Mora
- 78 **Implicit Teacher Theories Regarding the Argumentative Commentary of Multimodal Texts in the Teaching of Spanish as a Native and Foreign Language**
María Teresa Caro Valverde, José Manuel de Amo Sánchez-Fortún and Juana Celia Domínguez-Oller
- 96 **The Importance of Phonological Awareness in Learning Disabilities' Prevention: Perspectives of Pre-School and Primary Teachers**
Lurdes Veríssimo, Marisa Costa, Francisca Miranda, Catarina Pontes and Isabel Castro
- 103 **Vocabulary Repetition Following Multisensory Instruction Is Ineffective on L2 Sentence Comprehension: Evidence From the N400**
Reza Pishghadam, Haniyeh Jajarmi, Shaghayegh Shayesteh, Azin Khodaverdi and Hossein Nassaji

- 114 **School Belonging and Reading Literacy: A Multilevel Moderated Mediation Model**
Yuting Tan, Zhengcheng Fan, Xiaoman Wei and Tao Yang
- 125 **The Role of Digital Technologies to Promote Collaborative Creativity in Language Education**
Moisés Selfa-Sastre, Manoli Pifarré, Andreea Cujba, Laia Cutillas and Enric Falguera
- 138 **An Intervention in Reading Disabilities Using a Digital Tool During the COVID-19 Pandemic**
Irene Cadime, Iolanda Ribeiro, Joana Cruz, Maria do Céu Cosme, Diana Meira, Fernanda Leopoldina Viana and Sandra Santos
- 147 **Investigating Effects of Small-Group Student Talk on the Quality of Argument in Chinese Tertiary English as a Foreign Language Learners' Argumentative Writing**
Hui Helen Li and Lawrence Jun Zhang
- 161 **From Text on Paper to Digital Poetry: Creativity and Digital Literary Reading Practices in Initial Teacher Education**
Moisés Selfa Sastre and Enric Falguera Garcia
- 170 **The Image as Language: The Creation and the Use of the Visual Message by Young University Students in Their Communicative Social Activity**
Alfredo Ramón-Verdú and José Víctor Villalba-Gómez
- 181 **Research on correlation between English writing self-efficacy and psychological anxiety of college students**
Bai Li
- 191 **Contribution of oral narrative textual competence and spelling skills to written narrative textual competence in bilingual language-minority children and monolingual peers**
Giulia Vettori, Lucia Bigozzi, Oriana Incognito and Giuliana Pinto
- 203 **Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021**
Nieves Gutiérrez-Ángel, Jesús-Nicasio Sánchez-García, Isabel Mercader-Rubio, Judit García-Martín and Sonia Brito-Costa
- 238 **The role of learner character strengths and classroom emotions in L2 resilience**
Fakieh Alrabai and Abdullah Alamer
- 252 **Influence of performance in Spanish language and literature on physical education and music grades**
Esperanza Clares-Clares and Alberto Gómez-Mármol
- 257 **Research on Papua, a digital tool with artificial intelligence in favor of learning and linguistic attitudes towards the learning of the English language in students of Spanish language as L1**
Beatriz Peña-Acuña and Rafael Crismán-Pérez

- 272 **Predictors of Portuguese teachers' use of Information and Communication Technologies in literacy classes**
Andreia Nunes, Teresa Limpo and São Luís Castro
- 280 **Investigating high schoolers' L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement: Relationships and mediator**
Jie Zhou, Shusheng Wang and Junju Wang
- 289 **The integration of affective characteristics of the family environment for a more comprehensive explanatory model of reading abilities**
Andréanne Gagné, Nathalie Bigras, Annie Charron and Julie Lemire
- 300 **Analysis of pre-service teachers' argumentation-based academic writing process**
Bekir Direkci, Serdar Akbulut, Bilal Şimşek, Mevlüt Gülmez and Emel Nalçacıgil Çopur
- 316 **Psychometric properties and invariance of the self-efficacy for writing scale in Peruvian high school students**
Matilde León-Gutiérrez, Wilter C. Morales-García, Mardel Morales-García, Jania Jaimes-Soncco, Jacksaint Saintila and Percy G. Ruiz Mamani
- 327 **Academic literacy among the university students in Mexico and Spain: A holistic perspective**
Isolda Margarita Castillo-Martínez, Cynthia Patricia Cerros Regalado, Leonardo David Glasserman-Morales and María Soledad Ramírez-Montoya
- 341 **Relationships among students' reading habits, study skills, and academic achievement in English at the secondary level**
Nisar Abid, Sarfraz Aslam, Abdulelah A. Alghamdi and Tribhuwan Kumar
- 351 **Emotions and reading: When reading is the best way to improve skills in adolescents**
Elena del Pilar Jiménez-Pérez, María Isabel de Vicente-Yagüe Jara, Manuel León Urrutia and Pedro García Guirao
- 357 **The relationship between grammatical knowledge and reading comprehension: A meta-analysis**
Haoyuan Zheng, Xuecong Miao, Yang Dong and Daniel-Chongbo Yuan
- 370 **The promotion of critical reading through the digital environment: A study on the virtual epitexts used to promote children's picturebooks**
Rosa Tabernero-Sala and María Jesús Colón-Castillo
- 387 **A longitudinal study on sensitivity to symmetry in writing and associations with early literacy abilities**
Li Yin and Catherine McBride

- 395 **What spelling errors can tell us about the development of processes involved in children's spelling**
Georgia Z. Niolaki, Alexandra Negoita, Janet I. Vousden, Aris R. Terzopoulos, Laura Taylor and Jackie Masterson
- 405 **Development and validation of a genre-based second language (L2) writing self-efficacy scale**
Jianhua Zhang, Lawrence Jun Zhang and Ye Zhu
- 418 **Longitudinal contributions of morphological awareness, listening comprehension, and gains in word reading fluency to later word- and text-reading fluency**
Jamie L. Metsala
- 428 **A study on the emergence of sound-sign correspondence in Italian-speaking 5-year-old pre-schoolers**
Lucia Bigozzi, Oriana Incognito, Alice Mercugliano, Diletta De Bernart, Lucia Botarelli and Giulia Vettori
- 437 **Eye movements are stable predictors of word reading ability in young readers**
Andrea Strandberg, Mattias Nilsson, Per Östberg and Gustaf Öqvist Seimyr
- 452 **Reliability, validity, and measurement invariance of a Chinese handwriting legibility scale among primary students in central China**
Hong Lu, Xin Chen, Frederick K. S. Leung and Haode Zuo
- 465 **Internet use predicts Chinese character spelling performance of junior high school students: multiple mediating roles of pinyin input proficiency and net-speak experience**
Rong Luo, Yifan Peng and Jingjun Chen
- 473 **Effects of cognitive task complexity and online planning on second language learners' argumentative writing**
Ting Sophia Xu and Lawrence Jun Zhang



OPEN ACCESS

EDITED AND REVIEWED BY
Douglas F. Kauffman,
Medical University of the Americas,
United States

*CORRESPONDENCE
María Isabel de Vicente-Yagüe Jara
✉ isabelvyague@um.es

RECEIVED 17 August 2023
ACCEPTED 20 October 2023
PUBLISHED 27 November 2023

CITATION
Vicente-Yagüe Jara MI de, García Guirao P,
Jiménez-Pérez EdP and López Martínez O
(2023) Editorial: Reading and writing skills:
cognitive, emotional, creative, and digital
approaches. *Front. Psychol.* 14:1279276.
doi: 10.3389/fpsyg.2023.1279276

COPYRIGHT
© 2023 Vicente-Yagüe Jara, García Guirao,
Jiménez-Pérez and López Martínez. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Editorial: Reading and writing skills: cognitive, emotional, creative, and digital approaches

María Isabel de Vicente-Yagüe Jara^{1*}, Pedro García Guirao²,
Elena del Pilar Jiménez-Pérez³ and Olivia López Martínez⁴

¹Department of Didactics of Language and Literature, University of Murcia, Murcia, Spain, ²Department of Linguistics, WSB University in Wrocław, Wrocław, Poland, ³Department of Didactics of Language, Art and Sports, University of Málaga, Málaga, Spain, ⁴Department of Didactics of Evolutionary and Educational Psychology, University of Murcia, Murcia, Spain

KEYWORDS

reading, writing, cognitive, emotional, creative, digital

Editorial on the Research Topic

Reading and writing skills: cognitive, emotional, creative, and digital approaches

Within the Research Topic “*Reading and Writing Skills: Cognitive, Emotional, Creative, and Digital Approaches*”, 40 articles were published in the “Educational Psychology” section of the following two journals in a continuous publication: *Frontiers in Psychology* and *Frontiers in Education*. These are studies conducted in an educational context by researchers from Italy, Portugal, Spain, the Netherlands, Poland, Sweden, the United Kingdom, the United States, Canada, Peru, Mexico, Colombia, Turkey, Iran, China, Saudi Arabia, Pakistan, Malaysia, and New Zealand.

The development of literacy, defined as the psycholinguistic and social components of communication, is crucial in the educational context from primary school to university. Although reading and writing are two well-studied pillars of education, this Research Topic is currently interested in rethinking these educational processes, highlighting the cognitive, emotional, creative, and digital needs that have emerged as a result of the pandemic, using a multidisciplinary and holistic approach.

With confinement limits and social distancing measures, we are now trying to operate flexibly in a new health situation resulting from the worldwide outbreak of COVID-19. The growing “new normal” has reshaped instructors’ roles, techniques, and work schedules. Similarly, increasing digital literacy and developing strategies for mastering new skills are critical components of using, interpreting, communicating, and sharing information in this environment. In addition, parental involvement in the student learning process impacts not only academic success, but also a variety of emotional and social issues. In addition to the changing role of teachers and the affective social impact of the family, literacy skills can be useful in nurturing creativity.

Within this thematic perspective, the published articles have addressed the study of literacy from a variety of perspectives that provide an in-depth look at the current diverse educational landscape, as described below.

First, both reading comprehension and writing require the interaction of several cognitive processes. If we start with the learning of these skills at an early age, we find the studies “*A longitudinal study on sensitivity to symmetry in writing and associations with*

early literacy abilities” (Yin and McBride); and “A study on the emergence of sound-sign correspondence in Italian-speaking 5-year-old pre-schoolers” (Bigozzi et al.).

Following the cognitive strategies involved in reading skills, these are addressed in relation to morphological awareness, grammatical knowledge, text genres, eye movements, school belonging, academic achievement between linguistic and non-linguistic subjects. This is the case of the articles entitled “Longitudinal contributions of morphological awareness, listening comprehension, and gains in word reading fluency to later word- and text-reading fluency” (Metsala); “The relationship between grammatical knowledge and reading comprehension: A meta-analysis” (Zheng et al.); “Cognitive Strategies and Textual Genres in the Teaching and Evaluation of Advanced Reading Comprehension (ARC)” (García-Sánchez and García-Martín); “Eye movements are stable predictors of word reading ability in young readers” (Strandberg et al.); “School Belonging and Reading Literacy: A Multilevel Moderated Mediation Model” (Tan et al.); “Relationships among students’ reading habits, study skills, and academic achievement in English at the secondary level” (Abid et al.); and “Influence of performance in Spanish language and literature on physical education and music grades” (Clares-Clares and Gómez-Mármol).

In line with the articles on reading comprehension, the study “Vocabulary Repetition Following Multisensory Instruction Is Ineffective on L2 Sentence Comprehension: Evidence from the N400” (Pishghadam et al.), which used electroencephalography to record students’ electrophysiological neural activity, is also noteworthy.

Within this general background, learning difficulties in reading are studied from both prevention and intervention perspectives in “The Importance of Phonological Awareness in Learning Disabilities’ Prevention: Perspectives of Pre-School and Primary Teachers” (Veríssimo et al.); and “An Intervention in Reading Disabilities Using a Digital Tool During the COVID-19 Pandemic” (Cadime et al.).

Continuing with the teaching and learning of writing, the relevance of writing strategies in the cognitive processes of expression is studied in the articles “Validation of the Writing Strategies Questionnaire in the Context of Primary Education: A Multidimensional Measurement Model” (Arias-Gundín et al.); and “What spelling errors can tell us about the development of processes involved in children’s spelling” (Niolaki et al.).

In addition, the specific case of learning Chinese characters is addressed in the study “Comparing the Effects of Stroke-Appearing and Stroke-Disappearing on Learning the Order of Strokes in Chinese Characters” (Hong et al.); and “Reliability, validity, and measurement invariance of a Chinese handwriting legibility scale among primary students in central China” (Lu et al.).

Among the cognitive processes involved in writing, the mental operations of reasoning, interpreting, and arguing are of paramount importance in academic writing, as discussed in articles such as “Analysis of pre-service teachers’ argumentation-based academic writing process” (Direkci et al.); “Investigating Effects of Small-Group Student Talk on the Quality of Argument in Chinese Tertiary English as a Foreign Language Learners’ Argumentative Writing” (Li and Zhang); “Implicit Teacher Theories Regarding the Argumentative Commentary of Multimodal Texts in the Teaching

of Spanish as a Native and Foreign Language” (Caro Valverde et al.); and “Effects of cognitive task complexity and online planning on second language learners’ argumentative writing” (Xu and Zhang).

In particular, the belief in one’s own ability to communicate successfully through writing has been studied from the perspective of self-efficacy in several published articles, as is the cases of “Psychometric properties and invariance of the self-efficacy for writing scale in Peruvian high school students” (León-Gutiérrez et al.); and “Development and validation of a genre-based second language (L2) writing self-efficacy scale” (Zhang et al.). Moreover, self-efficacy can regulate anxiety and other emotional variables, as seen in the articles “Research on correlation between English writing self-efficacy and psychological anxiety of college students” (Li); and “Investigating high schoolers’ L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement: Relationships and mediator” (Zhou et al.).

Second, regarding the emotional education approach in language and literature teaching and learning processes, recent and significant publications on this Research Topic stand out. In recent years, a great deal of research has been published on reading literacy. However, there are hardly any studies that link reading literacy with emotional intelligence. In this line we find the article “Emotions and reading: When reading is the best way to improve skills in adolescents” (Jiménez-Pérez et al.), in which evidence was obtained of a direct relationship between both variables, especially in the experimental group in which the reading habit was stimulated.

From a learner resilience perspective, the article “The role of learner character strengths and classroom emotions in L2 resilience” (Alrabai and Alamer) examines a theory-based model to explain how trait emotional intelligence and language learner effort, as two-character strengths, predict learner enjoyment (positive emotions) and anxiety and boredom (negative emotions), and how these variables together predict resilience in language learning.

Moreover, motivation in the reading process is one of the main concerns of teachers today, as can be seen in the articles “Beyond the Educational Context: Relevance of Intrinsic Reading Motivation During COVID-19 Confinement in Spain” (De Sixte et al.); and “By Toutatis! Trainee Teachers’ Motivation when Using Comics to Teach History” (Moreno-Vera et al.).

Finally, the topic of the affective social influence of the family on the learning of reading is addressed in the following article: “The integration of affective characteristics of the family environment for a more comprehensive explanatory model of reading abilities” (Gagné et al.).

Thirdly, we find the role that creativity is playing in recent years in the educational field, in general, and in language learning and writing practice, in particular. This is the case of certain studies published in this Research Topic, which connect student creativity with digital practices, such as “The Role of Digital Technologies to Promote Collaborative Creativity in Language Education” (Selfa-Sastre et al.); and “From Text on Paper to Digital Poetry: Creativity and Digital Literary Reading Practices in Initial Teacher Education” (Selfa-Sastre and Falguera García).

Creative expression is addressed in another study that examines narrative textual competence, both oral and written, entitled “Contribution of oral narrative textual competence and spelling

skills to written narrative textual competence in bilingual language-minority children and monolingual peers" (Vettori et al.).

On the other hand, the creative development of students is also studied based on the image as language, in an article that aims to delve into the visual message in its productive and interpretive aspects, entitled "*The Image as Language: The Creation and the Use of the Visual Message by Young University Students in Their Communicative Social Activity*" (Ramón-Verdú and Villalba-Gómez).

Finally, in relation to the predominance of the digital era in the educational context, which requires the development of new skills and learning models, the article entitled "*Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021*" (Gutiérrez-Ángel et al.) aims to analyze the empirical evidence provided by international research related to digital literacy among university students.

More specifically, the use of information and communication technologies for language learning has been particularly studied in the contexts of Portugal and China, through the articles "*Predictors of Portuguese teachers' use of Information and Communication Technologies in literacy classes*" (Nunes et al.); and "*Internet use predicts Chinese character spelling performance of junior high school students: multiple mediating roles of pinyin input proficiency and net-speak experience*" (Luo et al.). Similarly, the pedagogical relevance of using artificial intelligence for learning English as an L2 is addressed in the article "*Research on Papua, a digital tool with artificial intelligence in favor of learning and linguistic attitudes toward the learning of the English language in students of Spanish language as L1*" (Peña-Acuña and Crismán-Pérez).

Likewise, literature is not excluded from this digital universe, which is reflected in the promotion of children's picturebooks in the study "*The promotion of critical reading through the digital environment: A study on the virtual epitexts used to promote children's picture books*" (Tabernero-Sala and Colón-Castillo). Similarly, reading-literary cooperation through spaces of participatory culture is developed in the article "*Mechanisms for Interpretative Cooperation: Fan Theories in Virtual Communities*" (de Amo and García-Roca).

Finally, the article "*Academic literacy among university students in Mexico and Spain: A holistic perspective*" (Castillo-Martínez et al.) aims to identify the extent to which cognitive, emotional, attitudinal, digital and personality aspects influence the development of academic literacy among university students.

In conclusion, through the reading of the different articles that make up this Research Topic, the value of a holistic view that integrates the study of literacy from a pluralistic research perspective has been demonstrated in order to meaningfully delve into the educational reality and be able to face the current problems, challenges and opportunities presented by the teaching and learning of reading and writing.

Author contributions

MV-Y: Conceptualization, Writing—original draft, Writing—review & editing. EJ-P: Writing—review & editing. OL: Writing—review & editing. PG: Writing—review & editing, Writing—original draft.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Mechanisms for Interpretative Cooperation: Fan Theories in Virtual Communities

José Manuel de Amo ^{*†} and Anastasio García-Roca [†]

Department of Education, University of Almería, Almería, Spain

OPEN ACCESS

Edited by:

Elena Jiménez-Pérez,
University of Malaga, Spain

Reviewed by:

Roberto Cuadros-Muñoz,
Sevilla University, Spain
Natalia Martínez León,
University of Granada, Spain

*Correspondence:

José Manuel de Amo
jmdeamo@ual.es

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 24 April 2021

Accepted: 18 May 2021

Published: 10 June 2021

Citation:

Amo JM and García-Roca A (2021)
Mechanisms for Interpretative
Cooperation: Fan Theories in Virtual
Communities.
Front. Psychol. 12:699976.
doi: 10.3389/fpsyg.2021.699976

This work focuses on analyzing fan theories as interpretive processes shared by users of virtual reading communities. Within these spaces of participatory culture, a complex strategy of interactions among its members is encouraged to formulate conjectures about the text's intentions and negotiate its degree of relevance. We have based our research on a methodology linked to the ethnography of reading. From a representative sample of the narrative universes of *A Song of Ice and Fire* and *Harry Potter*, the different modes of agreed collaborative textual interpretation are explored. The data show that within these communities of practice, three reading models are developed: predictive theories in which future narrative contents are inferred; explanatory theories in which narrative arcs are endowed and charged with meaning through the analysis of canon and, finally, alternative theories with a highly creative component in which the interpretative limits of the text are explored. Within these virtual communities, hermeneutical proposals are characterized by the activation of complex, heavily referenced literary argumentation to maintain semiosis active and expand the horizon of expectations of its members.

Keywords: digital reading, transmedia, fan theories, interpretation, virtual community, informal learning, literacy

INTRODUCTION

Numerous studies show how technological revolutions are changing the modes and habits of reading (Landow, 2009; Naseri and Noruzi, 2016; Cruces, 2017; Chartier, 2018; Amo, 2019). In addition to the reception of a hypertext, focused on hyperconnected fragments, the Internet has made the act of reading cease to be conceived exclusively as a solitary or individual cognitive activity and acquire a more social dimension: reading becomes conversation (Cordón, 2016; Lluch, 2017, 2018; Rovira-Collado, 2017). Thus, affinity spaces are created so members can share hobbies, objectives, and interests around books, authors, or themes (Gee and Hayes, 2011). They also collaborate to develop complex negotiations of meaning and sustain the collective semiosis of the text. Within these flexible and decentralized interactive structures, reading becomes linked to writing (the read/write culture). Readers not only activate their knowledge about code conventions (literary, audiovisual, or hypermedia), generate reading hypotheses, establish connections, and either confirm or frustrate their expectations (Culler, 2004: p. 80), but also play a more dynamic and creative role that transcends textual interpretation (from the canon), becoming content generators or protagonists in the circulation of meaning.

In these scenarios, there is a particular type of participatory commitment to the text and with other users, since belonging and recognition within the group, as well as its fascination with the source work, enhance the motivation to develop the social value of reading and the various literacy manifestations (Paladines-Paredes and Margallo, 2020). Fandom, from this perspective, is a committed, critical and intertextual practice (Gray et al., 2007). Fandom users behave as an authentic collective intelligence (Lévy, 2004; Jenkins et al., 2015) in which the interpretation of the text (fanon) together with promoting and guiding the creation of hypertextual manifestations is jointly negotiated but does not always reach consensus. In short, these are collective strategies leading to the formation of a reading community that interprets texts beyond the guidelines set by hegemonic culture (Rovira-Collado, 2017).

In this context, transmedia narratives (hereafter, TN) are particularly relevant (Jenkins, 2009). They are defined as complex semiotic systems that extend to different media, platforms, or formats and demand from their audience the activation of more sophisticated and complex reception mechanisms than in traditional reading. In an attempt to synthesize this dense concept, Scolari (2013) places its specificity in the sum of canon and fandom. The first refers to the contents that follow a top-down or descending logic and the diegetic universe that they develop (Guerrero-Pico, 2014). It is essential to clarify that they include official elements and, as described by Genette (1989, p. 316), all epitexts such as interviews with authors, marketing strategies, and official social networks (Lluch et al., 2015). Fandom, meanwhile, refers to the amateur community created around a cultural phenomenon, as well as its activities.

Important transformative or hypertextual creations are created within fandoms (*fanfics*, *fanfilms*, *remixes*, *fansubs*), characterized by being constantly built on and are, therefore, incomplete. Its production and reception mechanisms question the competencies traditionally attributed to the author and the reader (Fathallah, 2017). They are also part of the popular cultural and daily practices of young people (Lankshear and Knobel, 2010) and are developed in parallel with formal learning environments, linguistic and literary skills (Black, 2008; García-Roca, 2020), as well as the collaborative skills that new information and communication technologies require.

TNs develop complex fictional worlds from their possible narrative extensions into different media. In the words of Iser (1987), they intentionally play with the unsaid and encourage intersubjective discussions on individual or collective realizations of meaning and possible interpretations. Non-textual connections are made, and persuasive reasoning is made by searching for evidence or clues in corroborative works.

The creation of glossaries and other narratological elements in wikis are collaborative activities linked to collective intelligence (Booth, 2009). From this perspective, in the shared construction of community meanings and interpretation (fanon), there is often disagreement, proof of the subjective, polyphonic, and heteroglossic nature of TN (Thomas, 2018). These hermeneutic experiences have critical implications for multiliteracy and, in particular, reading training (Duncan, 2008).

Information processing is also a relevant aspect of fandom. For example, in cross-media narratives, readers of the literary saga (and other media) are often separated from those who watch films or television series. This segregation is due to the availability of different canon. The readings of members of virtual communities are also synchronized (Silva et al., 2015; García-Roca and Amo, 2019): accessing fragments of information in advance, which could ruin the surprise factor of narrative turns. Spoilers are fragments of narrative information shared among fans before the official narrative is made available (Hills, 2012). Some within the fandom focus on the creative challenge of drawing inferences and guesswork and enjoy advances, leaks, theories, and rumors (Johnson and Rosenbaum, 2017; Völcker, 2017). Mittell (2009) uses the concept of forensic fandom to describe the long-term commitment experienced by users who dissect the canon and immerse themselves in the deepest interpretative layers. He states that being a follower of these narratives implies:

to embrace a detective mentality, seeking out clues, charting patterns, and assembling evidence into narrative hypotheses and theories. This forensic engagement finds a natural home in online forums, where viewers gather to posit theories and debate interpretations, and fan wikis (Mittell, 2009, pp. 14–15).

Fan theories occupy a preferential place among fanon practices (Gray and Mittell, 2007). However, there is a lack of empirical studies that systematically and comprehensively analyze their structure, function, and definition. Prior research links fan theories, speculations, or hypotheses to forensic fandom and expand the narrative (Mittell, 2009). In addition, these theories are integrated into mechanisms for constructing shared meaning that allows readers to access more profound levels of interpretation than those achieved individually.

Fan theories are interpretative practices published within the virtual community in order that they may be discussed, modified, or ratified by the rest of its members. These are creative activities whose producers want to demonstrate to other users their reading competence (Aranda et al., 2013) or their degree of proximity to the model reader registered in the canon. This is a “lazy machine” (Eco, 1993) since the text requires the reader's participation to be potentially updated. It follows that the interpreter takes possible interpretations, and the various reading paths are foreseen and plotted by the text from the beginning since these movements form part of its own generative mechanism (Eco, 1993: p.79). It is, therefore, a textual strategy that the author must design or project. The work requires the receiver to activate the same elements of the reading competence described in the text to exhaust all possible meanings. This requisite receiver is not an empirical recipient but another mechanism of the text, built for interpretative cooperation: the Model Reader. Interpretation, therefore, necessarily implies a dialectic between the strategy of the text (model reader) and the actual reader's response (p. 86).

From this perspective, this work aims to explore the mechanisms of interpretative cooperation of TN in their natural context and, specifically, seeks to answer the following question: What role do fan theories play in the reception of narratives in which they are included? It explores the

meanings that the analyzed communities confer on reading to understand the implications of reading in the configuration and functioning of the community and what rules they use in the interpretation process. From this perspective, the following operational objectives have been established:

- 1 Explore the modes of interpretation that are generated in virtual reading communities.
- 2 Analyze and describe the development and impact of interpretative processes within the fandom.
- 3 Classify fan theories according to the type of interpretation and their argumentative and persuasive techniques.

MATERIALS AND METHODS

Qualitative research has been carried out in which the modes of textual interpretation are addressed by analyzing the discourse of different hermeneutical proposals shared within the network. It is an ethnographic study in which specific aspects of communities are studied in their natural context to know and understand them (Hine, 2004). Based on the idea of interpretation as a social practice, it is interesting to observe the processes of communicative interaction of its members, the types of literary practices, their rules of negotiation of meaning, and their reading canon.

We have taken a sample of fan theories between 2015 and 2020 of two of the most socially vibrant fandoms linked to the TN originating from literary texts: *A Song of Ice and Fire* and *Harry Potter*. Regarding the sample selection, we have opted for the most relevant theories of the two official fandom spaces (*Pottermore* and *Westeros*) and related virtual scenarios created by fans (e.g., *lossitereinos*, wikis, or fandom communities). Texts have been analyzed until the theoretical saturation of the data and analysis categories was achieved. In the search and data collection process, the digital archive *Wayback Machine* files from *Archive.org* and the *Google* search tool were utilized to analyze fan theories at the time of their publication vis-à-vis the canon. Note that while selected narratives began more than 20 years ago, fan theories, in contrast, are usually phenomena with limited temporal vibrancy.

The data analysis has been based on a holistic and emerging coding and categorization process with the help of the *Atlas.ti* data analysis program. Despite being based on predetermined categories extracted from the bibliographic review, they have been developed and re-coded throughout the research process.

- Each fan theory has been analyzed around the following variables.
- Argumentation: the primary narrative evidence that underpins fan theories has been categorized.
- Intent and projection: we have analyzed what creative strategies are used to complete the indetermined gaps in the narrative and expand the narrative universe, that is, prediction of the end of the story, more detailed description of a narrative

situation, development of motivations of secondary characters, clearing unknowns linked to the general plot, etc.

- Adaptation to the canon (as shared interpretation): in this category, the reactions of other users have been essential.
- In addition, four elements have been added that allow us to delve into the characteristics of each of the fan theories analyzed.
- Impact on the general readership and appropriation by the community of followers.
- Inclusion of insider information or spoilers.
- The obsolescence of theory is whether reading solves aspects that will soon be developed in the canon or proposed in the long term (outcome).
- Reaction from the authors and producers in the cases where the theories have been officially confirmed or refuted.

A researcher triangulation has been performed with the help of research experts with whom the mechanisms of coding, categorization, and development of hermeneutic models have been discussed and adjusted to minimize the subjective bias of the data analysis process.

RESULTS

Fan theories are interpretative proposals made by readers which are discussed, contrasted, and shared within affinity spaces. They anticipate or infer future content, explain specific events, or propose alternative visions. These interpretations are accompanied by solid arguments and precise references to canonical elements that give them greater likelihood and credibility: textual fragments are cited, frames are shared, or information from other media is shared as links.

Fan theories are constantly being evaluated and evolved. Some fan theories are finally confirmed, others are thwarted by the established canon. This notwithstanding, it is rare for the author's intention to prevail in these processes of interaction between the text and its readers and between the indications of the text and the readers' response. Other theories are reconfigured and end up being adapted to the canon. Examples of these include those regarding the homosexuality of Dumbledore or the prophecy of the Three Brothers, which theorizes that these characters are actually Harry Potter, Voldemort, and Dumbledore. J.K. Rowling, the original author, has confirmed these theories. Starting in 2011 on the *Westeros* forum with a simple paragraph about Canna (2011), another theory states that Bran is the King of the Night. This theory is currently in force, although much more consolidated, having been further developed with different proposed arguments from those initially put forward. Therefore, they are loaded with content through collective intelligence with details from the most up-to-date hypertext or base text are constantly added.

Likewise, theories that question and contradict others are detected. In this sense, there is evidence of different outcomes proposed related to the *Iron Throne*.

Sometimes the lines that separate fan theories from fanfictions are too fuzzy: theories become so complex that they create

possible and alternative lines of argument. The creativity and over-interpretation of fans interact to expand the history of hypertext (or canon) and thus the process of receiving these stories. Some theories, fully integrated into the fanon, acquire such relevance in fandom that they are not refuted, even though they may not fit with the coherence of the narrative in the subsequent canon.

The inductive analysis of the data has established three broad categories depending on the objective of the theory, the arguments used, the degree of obsolescence, their temporal location, their impact, and the textual appropriation by the members of the community. However, it is important to note that the categories are not siloed theories but instead often complement each other.

Predictive Theories

Predictive theories aim to advance and infer possible narrative actions that will be developed in the canon. From the reception point of view, these hermeneutical proposals reflect the unconscious and individual exercise of conjecture performed during reading.

These theories are not necessarily developed by fans, that is, by users with a significant intellectual, emotional and social involvement with the narrative or cultural product, but rather, are proposed and led by a less specialized and single-media readership, that of the literary, cinematographic, or television saga. Although they generate considerable debates, they do not go beyond the surface layer of the narrative. These are individual and disconnected theories that do not undergo a clear evolution.

Predictive theories are the most numerous and decentralized: being found in non-specialized spaces such as social media networks such as *Twitter* and *Facebook* through the use of *hashtags*. As narrative hypotheses, they are ephemeral and have little transcendence for the community: any user can guess the plot's outcome without mastering all available content. In this sense, these ideas usually lack a rigorous foundation or a clear justification: they are directly linked to the outcome (the Iron Throne or Voldemort's death) or explicit interpretative gaps in history.

Readers generate their hypotheses individually and then share them on a network for discussion. Generally, during this reading process, the underlying arguments for these theories are not examined. They often rely on the latest published narrative elements, but occasionally on epitextual marketing elements such as trailers or narrative advances.

They are temporary interpretative proposals since, over time, they end up either being confirmed or refuted. It should be noted that predictions are only generated in serialized productions where there is synchronization in its reception by the reader communities. They are contextualized theories at a time in the narrative. For example, the predictive theories developed in 2006 (with *Order of the Phoenix* and *Feast of Crows* as the latter publications of these narratives) are not the same as those developed in 2012 (*Dance of Dragons* and *Deathly Hallows*). For example, new followers who have read the Harry Potter heptalogy will not raise or read these shared

predictions on the Internet as they already have the outcome at their disposal.

Some notable examples of predictive theories are:

- The quest for *Valonqar*: it was speculated that it could be Jaime (and return to being the *kingslayer*), Tyrion, Jon, Danaerys, Arya, Stanis, or even Sam.
- Who would sit on the Iron Throne: the house Stark, Lanister, Targaryen, Baratheon, *The Night King*, all or none of them.
- The resurrection of Jon Snow was a consolidated and shared theory. For one part of the fandom, it was not a surprise but rather a confirmation.
- *Bran is the Night King*, which, despite being developed before the television series, gained particular relevance in the *fandom* due to the physical resemblance of the two characters. This theory has evolved in different directions, for example, in the alternative theories of Faillace (2019) and Khaled Comics (2020).
- Harry Potter's fandom speculated about the necessary death of Harry Potter (or one of his friends) to end Voldemort.
- For a long time, there were theories related to the different *ships* (relationships) of various characters in J. K. Rowling's saga.

Explanatory Theories

The second block of theories includes interpretative proposals that attempt to cover hermeneutical gaps or gaps in the text. Explanatory theories are canonical interpretations that endow and (over) subscribe meaning to the events of the available canon. While predictive theories aim to infer outcomes explicitly, explanatory theories detail the complex process before the outcome. These semiotic mechanisms favor these interpretative proposals and explore the subtext, nuance, or implicit content linked to characters, functions, and narrative actions. In this sense, the theories that analyze the true motivations of the King of the Night and Voldemort are noteworthy.

Explanatory theories are found in specialized affinity spaces such as specific sections of *Reddit*, *Westeros*, the *lossiterinos*, *Archive Of Our Own* (AO3), and more recently, on specialized YouTube channels. In this sense, influencers in YouTube such as *Javi Marcos* are popular for their meticulous analysis and fan theories such as the one entitled *Why is Jon Snow's real name Aegon Targaryen also in the books* (Marcos, 2018) or *Capa Invisible* with videos such as *Theories of the founders of Hogwarts* (Capa Invisible, 2020). In these scenarios, complex theories and their accompanying evidence are presented and discussed through fragments from all TNs. They are collective constructions created by forensic fandom and readers with a highly developed sense of literary reading competence that show great expertise in recognizing and gathering clues that help to understand what has not been explicitly narrated. The fandom carries out original research, collection, documentation, and analysis of works.

Hence, in addition to all the official components of transmedia narratives, epitextual elements are fundamental for the development of these hermeneutical proposals: interviews

with authors, statements on social networks, rumors, and leaks are some of the aspects that keep the fandom (and semiosis) active among publications, releases or broadcasts. Eco (1993, 1997) stated that empirical readers, in the form of a community, develop interpretative hypotheses that reconstruct the narrative, taking into account the author's intention or author-model.

They are not necessarily correct (they are hypotheses), but they are possible, canonical, and convincing. They provide answers to essential questions about the plot and justify and use narrative arguments to support a theory. They are accompanied by precise references to canonical elements: specific pages or fragments of the literary sagas, concrete frames in the audiovisual version, and epitextual elements. To this, we must add that collective intelligence feeds back and reinforces (or refutes) fan theories. In this regard, G. R. R. Martin notes:

The Internet affects all this to a degree it was never affected before [...] Like Jon Snow's parentage. There were early hints about it in the books, but only one reader in 100 put it together. And before the Internet that was fine—for 99 readers out of 100 when Jon Snow's parentage gets revealed, it would be, "Oh, that's a great twist!" But in the age of the Internet, even if only one person in 100 figures it out then that one person posts it online and the other 99 people read it and go, "Oh, that makes sense" (Hibberd, 2019, par. 2).

Among the most outstanding examples we can note:

- The theory of $R+L=J$ was raised in 1997 when only the first book was published. It has been one of the most widespread fandom theories and was confirmed in the canon almost 10 years later. This theory gives meaning to many details of the plot.
- Tyrion is the third head and, therefore, is Targaryen. Many events lead to this theory. We can find countless fragments and scenes used as a basis for this theory on the Internet.
- Dumbledore represents Death in *The Fable of the Three Brothers*, an interpretation of the fable that the author has confirmed on social networks (Rowling, 2015).

Alternative Theories

Alternative theories are the least common and the most complex, as they offer unique, personal, improbable, and creative visions of the fictional universe. These interpretative proposals defy the limits of interpretation.

Structurally and narratively, they are similar to explanatory theories: they cover argumentative gaps and propose well-founded explanations. However, they veer sharply away from the shared interpretation of the community (fanon). The fandom would automatically discard these theories if it were not for the solid arguments that accompany them. It is precisely one of its main characteristics. Given the impact and breakdown of the horizon of expectations they cause, they need to be duly justified with explicit evidence to soften the proposals' implausibility. Despite this, they are understood as amusing theories but are not deemed viable interpretations by other users.

They often display a high level of formulation, creativity, and complexity. In this way, they allow readers with low literary and literary skills, through social reading, to achieve a greater textual

understanding. What is relevant is that forensic fans exclusively create them, that is, those who dissect the canon and gather different fragments to construct the theory and persuade other users that such interpretations are viable, acceptable, and reliable.

Alternative theories comprise a high creative component and, therefore, usually have recognized authorship or source. Unlike other *fanworks* such as *fanfics* (transformative works which are character-centric), these readings keep the characters, the fictitious world in which they are placed, and the story narrated in the canon intact without including new narrative elements. While it is true that they update any potential meanings and interpretations of the text and are often overloaded with content, they are still suggested readings.

These theories are formed in specialized affinity spaces, but their playful character and interpretive shock value transcend and are subsequently published by other media and websites with more varied content. Most theories are timeless and current (even though new content has been published later): Could they be confirmed in the future? Yes, all alternative theories are possible, albeit improbable. From this perspective, alternative theories can pose new readings and can appear long after the narrative is completed. Some notable examples:

- Hogwarts is, in reality, a mental health institution in which Harry Potter is admitted, and the whole saga is a product of his imagination (see Roning_Ikari, 2016).
- Ron Wesley and Albus Dumbledore are actually the same character who has traveled in time with the aid of a Time-Turner to end Voldemort (later known as *Ronbledore*). Knight2King raised this theory in late 2003 (see Knight2king, 2004; Mallori, 2014). This particular reading has been maintained over time, and new arguments have been incorporated. However, 10 years later, the author, J. K. Rowling, intervened to refute this theory as it had garnered great relevance in the fandom (Rowling, 2015).
- Ned Stark is still alive. Different theories support this thesis:
 - Eddard Stark possesses shifter powers (like his son). Before being executed, he changed his body, specifically to that of the birds flying away.
 - Ned shared a cell with Jaqen (Syrian Forel), who could change his face. It was the latter who was executed. This would resignify Catelny's surprise on seeing that Ned's bones were too small.
- The whole game of thrones responds to an evil plan of the Three-Eyed Raven and Brandon Stark (Faillace, 2019). This theory has inspired different *fanfics* and alternative endings (see, for example, Khaled Comics, 2020).
- Hagrid is actually a *Death Eater* (see Whoofph, 2019).

Below, **Table 1** schematically develops the characteristics of each type of theory.

As vernacular literary practices, these three interpretative models are developed by highly motivated and committed users. Their participation in the processes of negotiation of meaning is mainly determined by the need to belong and be recognized in their affinity space and generate

TABLE 1 | Summary of analysis of fan theories.

	Objective	Arguments	Elaboration	Adequacy	Impact and obsolescence
Predictive	Infers narrative elements	No significant arguments	Individual	Canonical	Of little relevance to the fandom. Very abundant. Soon to be confirmed or refuted
Explanatory	Provides meaning to narrative events	Solid arguments with precise references to canonical elements	Fandom		The most transcendental theories. Expand the horizon of expectations of the fandom
Alternative	Rethinks reading and offers an alternate view		Individual	Non-canonical	Complex, unique, and peripheral. They are neither confirmed nor refuted by the canon

emotions linked to the canon or canonical text (Marina, 2007, 2011; Pink, 2010). This finding is highly relevant when transferring the proposal to the dynamics of formal education and, in particular, to “interpretive community” classes.

DISCUSSION

The Internet promotes the creation of virtual reading communities within which norms of production and reception of texts are negotiated and agreed on. The horizon of expectations common to its members is outlined: a reference system that guides them in reading and that translates into the activation of their previous experiences and knowledge about narrative sub-genres, formal, thematic, and discursive traits, as well as their ideas about what is literary (and transmedia) language as opposed to everyday or functional language.

As Culler (2004: p.80) stated, the interpretation of a work is developed to answer the questions formulated by this horizon of expectations. It is one in which the virtual reading community determines the types of valid or possible answers. In this case, the key lies in how textual details are used to link them to those answers. Fan theories, from this perspective, are forms of reading inscribed in the hermeneutic mechanisms of the community.

Fan theories are interpretive hypotheses proposed by empirical readers who collaborate and debate with other readers about the signs and clues that appear in the appellative structure of the text. Even when they occasionally successfully resolve certain plot unknowns ahead of the official narrative, the fandom does not reject these theories or consider them spoilers since they usually do not incorporate insider information. Given the large number and variety of interpretative possibilities offered, individual readers only access readings from other users and blur the lines of what is considered a spoiler, rumor (or hoax), fan theories, and *fanfics*. It enables readers in fan communities with low reading and literary competence to understand and access more profound levels of interpretation assisted by other followers in the community that helps to broaden their members’ expectations horizon. These results match those of Ellithorpe and Brookes (2018).

Therefore, rather than being spoilers, these theories are an integrated and inherent part of the process of receiving (transmedia) analyzed narratives: readers enjoy the text individually and subsequently develop their own readings; adjusting and contrasting their interpretations with the community (*fanon*) (Jenkins, 2009; Jenkins et al., 2015; Völcker, 2017). Users thus carry out negotiated and consensual re-readings in order to understand and maintain an active semiosis.

Readers are often reluctant to say goodbye to their favorite texts and, to avoid doing so, participate in new social and creative activities that allow them to continue enjoying the plot or developing their favorite characters. From this point of view, the limits of interpretation are sometimes intentionally playfully outstripped with what Eco (1993, p. 86) calls “an aesthetic of the free, aberrant, intentional and malicious use of texts.” In this way, manifestly aberrant interpretations are shared. Fan theories are located in the semiotic debates of delimitation of these interpretative limits established by fandom, that is, the fanon. They generate extensive hermeneutical discussions within fan communities, as evinced by the research by Chaney and Liebler (2007) and Thomas (2018).

Alternative theories result from exhaustive textual analysis by readers with a developed literary competence in which information is collected and reconstructed in exceptionally creative ways: an invitation to rereading and reinterpret the text (Scolari et al., 2018).

However, it is pertinent to note that fan theories are only operational and functional (except for alternative theories, which are timeless) when the narrative is serialized and not concluded: they are formed, enjoyed, and discussed in the breaks between distributions of plot content. Generally speaking, if an asynchronous reading is performed (that is to say, approaching the text in a chronological period separate from the processes of cultural dissemination) or if it occurs after the saga has ended, fan theories are not relevant to the readership.

The hypothesis is that fan theories influence the artificial generation of expectations (hype) by overvaluing narrative possibilities to TN viewers and readers. Confirmation of these hypotheses could explain the rejection of the fandom to the outcomes of the main transmedia narratives: *Game of Thrones*, *Harry Potter*, *Star Wars*, or *Lost*.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Amo, J. M. (2019). *Nuevos modos de lectura en la era digital*. Madrid: Síntesis.
- Aranda, D., Sánchez-Navarro, J., and Roig, A. (2013). *Fanáticos. La cultura fan*. Sevilla: OUC Press.
- Black, R. (2008). *Adolescents and Online Fan Fiction*. New York, NY: Peter Lang.
- Booth, P. (2009). Narrativity and the narrative database. Media-based wikis as interactive fan fiction. *Narrat. Inq.* 19:2. doi: 10.1075/ni.19.2.09boo
- Cainna (2011). *The Night's King=Brandon the Builder? (...). A Forum of Ice and Fire*. Available online at: https://asoiaf.westeros.org/index.php?/topic/51144-three-eyed-crow/page__view__findpost__p__2527735/ (accessed January 15, 2021).
- Capa Invisible (2020). *Teorías de los fundadores de Hogwarts [video]*. YouTube. Available online at: https://www.youtube.com/channel/UC_2Wa-00YsCMVYFZSCrbzAw (accessed May 12, 2021).
- Chaney, K., and Liebler, R. (2007). "Canon vs fanon: folksonomies of fan culture," in *Mit Media in Transition 5: Creativity Ownership and Collaboration in the Digital Age (1-16)*. Available online at: http://web.mit.edu/comm-forum/legacy/mit5/papers/Chaney-Liebler_MIT5.pdf (accessed January 30, 2021).
- Chartier, R. (2018). *Las revoluciones de la cultura escrita*. Barcelona: Gedisa.
- Cordón, J. A. (2016). La lectura en el entorno digital: nuevas materialidades y prácticas discursivas. *Rev. Chil. Lit.* 94, 15–38. doi: 10.4067/S0718-22952016000300002
- Cruces, F. (2017). *¿Cómo leemos en la sociedad digital? Lectores, booktubers y prosumidores*. Barcelona: Ariel.
- Culler, J. (2004). *Breve introducción a la teoría literaria*. Barcelona: Crítica.
- Duncan, S. (2008). "Literacy implications of online fan debates," in *Annual Meeting of the American Educational Research Association (24-28th of March)* (New York, NY).
- Eco, U. (1993). *Lector in fabula. La cooperación interpretativa en el texto narrativo*. Montevideo: Lumen.
- Eco, U. (1997). *Interpretación y sobreinterpretación*. Cambridge: Cambridge University Press.
- Ellithorpe, M., and Brookes, S. (2018). I didn't see that coming: spoilers, fan theories, and their influence on enjoyment and parasocial breakup distress during series finale. *Psychol. Pop. Media Cult.* 7, 250–263. doi: 10.1037/ppm0000134
- Faillace, D. (2019). *Según esta teoría Bran movió malignamente los hilos para ser Rey. Spoiler Time*. Available online at: <https://spoilerime.com/teoria-bran-stark-malvado/> (accessed March 1, 2020).
- Fathallah, J. (2017). *Fanfiction and the Author: How Fanfic Changes Popular Cultural Texts*. Amsterdam: Amsterdam University Press.
- García-Roca, A. (2020). Los fanfictions como escritura en colaboración: modelos de lectores beta. *Prof. Inf.* 28:4. doi: 10.3145/epi.2019.jul.04
- García-Roca, A., and Amo, J. M. (2019). Jóvenes escritores en la red: un estudio exploratorio sobre perfiles de Wattpad. *Ocnos Mag. Read. Stud.* 18:3. doi: 10.18239/ocnos_2019.18.3.1968
- Gee, J. P., and Hayes, E. (2011). *Language and Learning in the Digital Age*. London: Routledge.
- Genette, G. (1989). *Palimpsestos. Literatura en Segundo grado*. Barcelona: Taurus.

FUNDING

This research had been funded by the University of Almería's Own Research and Transfer Plan and is part of the R + D + I Project entitled Epistemic innovation of an argumentative comment model of multimodal texts in the teaching of Spanish as a mother tongue and foreign (ref. PGC2018-101457-B-I00) of the Ministry of Science, Innovation and University of the Government of Spain.

- Gray, J., and Mittell, J. (2007). *Speculation on Spoilers: Lost Fandom, Narrative Consumption, and Rethinking Textuality. Participations*. Available online at: https://www.participations.org/Volume%204/Issue%201/4_01_graymittell.htm (accessed February 5, 2021).
- Gray, J., Sandvoss, C., and Harrington, C. L. (2007). *Fandom Identities and Communities in a Mediated World*. New York, NY: University Press.
- Guerrero-Pico, M. (2014). *Historias más allá de lo filmado: Fan fiction y narrativa transmedia en series de televisión* (dissertation). Universitat Pompeu Fabra, Barcelona, Spain.
- Hibberd, J. (2019). *George R.R. Martin: Game of Thrones fan reactions won't change my books. Entertainment Weekly*. Available online at: <https://ew.com/tv/2019/07/15/george-rr-martin-game-thrones-fan-reactions> (accessed March 21, 2020).
- Hills, M. (2012). "Psychoanalysis and digital fandom: theorizing spoilers and fans' self-narratives," in *Producing Theory in a Digital World: The Intersection of Audiences and Production in Contemporary Theory*, ed R. Lind (Bern: Peter Lang), 105–122.
- Hine, C. (2004). *Etnografía virtual*. Sevilla: Editorial UOC.
- Iser, W. (1987). *El acto de leer. Teoría del efecto estético*. Barcelona: Taurus.
- Jenkins, H. (2009). *The Revenge of the Origami Unicorn: Seven Principles of Transmedia Storytelling (Well, Two Actually. Five More on Friday)*. Available online at: http://henryjenkins.org/blog/2009/12/the_revenge_of_the_origami_uni.html (accessed January 5, 2021).
- Jenkins, H., Ford, S., and Green, J. (2015). *Cultura Transmedia. La creación de contenido y valor en una cultura en red*. Barcelona: Gedisa.
- Johnson, B., and Rosenbaum, J. (2017). (Don't) Tell me how it ends: spoilers, enjoyment, and involvement in television and film. *Media Psychol.* 21:4. doi: 10.1080/15213269.2017.1338964
- Khaled Comics (2020). *The Video Has Been Removed Due to a Copyright Claim, I Don't Know Why Specially That It Should be Under a Fair Use Act, Here It Is Again*. Available online at: <https://twitter.com/KhaledComics/status/1235622077190877184> (accessed December 15, 2020).
- Knight2king (2004). *Weasley is Our King. Unplottables. Knight2King and other Harry Potter Theories*. Available online at: <https://unplottables.wordpress.com/weasley-is-our-king/> (accessed December 2, 2020).
- Landow, G. (2009). *Hipertexto 3.0. Teoría crítica y nuevos medios en la era de la globalización*. Barcelona: Paidós.
- Lankshear, C., and Knobel, M. (2010). *Nuevos alfabetismos. Su práctica cotidiana y el aprendizaje en el aula*. Madrid: Morata.
- Lévy, P. (2004). *Inteligencia colectiva: por una antropología del ciberespacio*. Washington, DC: Organización Panamericana de la Salud.
- Lluch, G. (2017). "Los jóvenes y adolescentes comparten la lectura," in *¿Cómo leemos en la sociedad digital? Lectores, booktubers y prosumidores*, ed F. Cruces (Madrid: Ariel), 28–51.
- Lluch, G. (2018). *Claves para promocionar la lectura en la red*. Madrid: Síntesis.
- Lluch, G., Tabernero-Sala, R., and Calvo-Valios, V. (2015). Epitextos virtuales públicos como herramientas para la difusión del libro. *Prof. Inf.* 2:6. doi: 10.3145/epi.2015.nov.11
- Mallori, D. (2014). *This Guy On An Old Harry Potter Forum Says Dumbledore Is A Time-Traveling Ron Weasley And I Want To Hear Him Out. The Toast*. Available online at: <https://the-toast.net/2014/01/02/let-the-man-speak/> (accessed March 2, 2020).

- Marcos, J. (2018). *Por qué el verdadero nombre de Jon Nieve es Aegon Targaryen también en los libros [video]*. YouTube. Available online at: https://www.youtube.com/watch?v=CcZ-_rr7Sww (accessed May 12, 2021).
- Marina, J. A. (2007). *Las arquitecturas del deseo*. Barcelona: Anagrama.
- Marina, J. A. (2011). *Los secretos de la motivación*. Barcelona: Ariel.
- Mittell, J. (2009). "Lost in a great story: evaluation in narrative television (and television studies)," in *Reading Lost: Perspectives on a Hit Television Show*, ed R. Pearson (London: Bloomsbury Publishing), 119–138.
- Naseri, Z., and Noruzi, A. (2016). The effect of digital age on youth reading habits and practices. *Res. Inf. Sci. Public Lib.* 22:2.
- Paladines-Paredes, L., and Margallo, A. (2020). Los canales booktuber como espacio de socialización de prácticas lectoras juveniles. *Ocnos J. Read. Stud.* 19:1. doi: 10.18239/ocnos_2020.19.1.1975
- Pink, D. H. (2010). *La sorprendente verdad sobre qué nos motiva*. Barcelona: Gestión 2000.
- Roning Ikari (2016). *[Harry Potter] Hogwarts is a madhouse (literally)*. San Francisco, CA: Reddit. Available online at: https://www.reddit.com/r/FanTheories/comments/4z7gh8/harry_potter_hogwarts_is_a_madhouse_literally/ (accessed December 7, 2020).
- Rovira-Collado, J. (2017). Booktrailer y Booktuber como herramientas LIJ 2.0 para el desarrollo del hábito lector. *Res. Read.* 7, 55–72. doi: 10.24310/revistaisl.vi7.10981
- Rowling, J. K. (2015). *Dumbledore as Death. It's a Beautiful Theory and It Fits*. Available online at: https://twitter.com/jk_rowling/status/634666937990152192 (accessed December 10, 2020).
- Scolari, C. A. (2013). *Narrativas transmedia: cuando todos los medios cuentan*. Barcelona: Deusto.
- Scolari, C. A., Masanet, M. J., Guerrero-Pico, M., and Establés, M. J. (2018). Transmedia literacy in the new media ecology. *Prof. Inf.* 27:4. doi: 10.3145/epi.2018.jul.09
- Silva, P., Amer, Y., Tsikerdanos, W., Shedd, J., Restrepo, I., and Murray, J. (2015). "A game of thrones companion: orienting viewers to complex storyworlds via synchronized visualizations," in *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (Brussels), 197–172.
- Thomas, P. (2018). Canon wars: a semiotic and ethnographic study of a Wikipedia talk page debate concerning the canon of Star Wars. *J. Fandom Stud.* 6:3. doi: 10.1386/jfs.6.3.279_1
- Völcker, M. (2017). "Spoiler!? I'm completely painless, I read everything": fans and spoilers – results of a mixed-method study. *Participations. J. Audience Recept. Stud.* 14:1.
- Whoofph (2019). *Hagrid is a Death Eater*. San Francisco, CA: Reddit. Available online at: https://www.reddit.com/r/FanTheories/comments/cmb746/hagrid_is_a_death_eater/ (accessed March 20, 2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Amo and García-Roca. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Validation of the Writing Strategies Questionnaire in the Context of Primary Education: A Multidimensional Measurement Model

Olga Arias-Gundín¹, Sara Real², Gert Rijlaarsdam³ and Paula López^{1*}

¹ Department of Psychology, Sociology and Philosophy, University of Leon, Leon, Spain, ² Ponferrada Associated Centre, National University of Distance Education (UNED), Leon, Spain, ³ Research Institute for Child Development and Education, University of Amsterdam, Amsterdam, Netherlands

OPEN ACCESS

Edited by:

María Isabel de Vicente-Yagüe Jara,
University of Murcia, Spain

Reviewed by:

Inmaculada Clotilde Santos Díaz,
University of Malaga, Spain
Ester Trigo-Ibáñez,
Universidad de Cádiz, Spain

*Correspondence:

Paula López
plogp@unileon.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 26 April 2021

Accepted: 17 May 2021

Published: 05 July 2021

Citation:

Arias-Gundín O, Real S, Rijlaarsdam G
and López P (2021) Validation of the
Writing Strategies Questionnaire in the
Context of Primary Education: A
Multidimensional Measurement
Model. *Front. Psychol.* 12:700770.
doi: 10.3389/fpsyg.2021.700770

Research has shown that writers seem to follow different writing strategies to juggle the high cognitive demands of writing. The use of writing strategies seems to be an important cognitive writing-related variable which has an influence on students' writing behavior during writing and, therefore, on the quality of their compositions. Several studies have tried to assess students' writing preferences toward the use of different writing strategies in University or high-school students, while research in primary education is practically non-existent. The present study, therefore, focused on the validation of the Spanish Writing Strategies Questionnaire (WSQ-SP), aimed to measure upper-primary students' preference for the use of different writing strategies, through a multidimensional model. The sample comprised 651 Spanish upper-primary students. Questionnaire data was explored by means of exploratory (EFA) and confirmatory (CFA) factor analysis. Through exploratory factor analysis four factors were identified, labeled thinking, planning, revising, and monitoring, which represent different writing strategies. The confirmatory factor analysis confirmed the adequacy of the four-factor model, with a sustainable model composed of the four factors originally identified. Based on the analysis, the final questionnaire was composed of 16 items. According to the results, the Spanish version of the Writing Strategies Questionnaire (WSQ-SP) for upper-primary students has been shown to be a valid and reliable instrument, which can be easily applied in the educational context to explore upper-primary students' writing strategies.

Keywords: writing strategies, questionnaire, upper-primary education, psychometrics, validity

INTRODUCTION

Writing has been defined as a problem-solving task that places multiple cognitive demands on the writer (Hayes, 1996). As Flower and Hayes indicated in the first cognitive model of writing (Flower and Hayes, 1980), writers have to manage several cognitively costly processes such as planning what to say, translating and transcribing those plans into written text, and revising either the plans or the written text (Alamargot and Chanquoy, 2001; Hayes, 2012). The use of these processes, especially

in young writers, in whom basic transcription skills are not yet automated (Pontart et al., 2013; Alves et al., 2016; Limpo et al., 2017; Llauro and Dockrell, 2020), consumes much of the capacity of their working memory as these processes recursively interact during composition (McCutchen, 2011).

Following a comprehensive literature review, Graham and Harris (2000) concluded that writing development seems to depend on the automation of transcription skills and the acquisition of high-levels of self-regulation in order to handle high-level processes such as planning and revision. Self-regulation, represented by the use of writing strategies, is a critical aspect of writing as it enables writers to achieve their writing goals (Zeidner et al., 2000; Santangelo et al., 2016; Puranik et al., 2019). These strategies may reduce cognitive overload as they allow writers to divide, sequence, and regulate the attention paid to the different writing processes (Kieft et al., 2006; Beauvais et al., 2011). Empirical research has shown that writers' strategic behavior during composition strongly predicts the quality of "novices" and "experts" texts (Beauvais et al., 2011; Graham et al., 2017a, 2019; Wijekumar et al., 2019). Accordingly, the use of writing strategies has been generally considered to be a critical individual writing-related variable (Kieft et al., 2008), and is a major focus of research in writing instruction (Harris et al., 2010; Graham and Harris, 2018) from the earliest stages of education (Arrimada et al., 2019). Exploring students' use of different writing strategies during composition seems to be a critical aspect and should be considered in the fields of writing and writing instructional research.

Several studies have attempted to explore how writers differ in the use of different writing strategies (Torrance et al., 1994, 1999, 2000; Biggs et al., 1999; Lavelle et al., 2002; Kieft et al., 2006, 2007, 2008). These studies identified two main writing strategies, related with the processes identified in the first seminal cognitive model of writing (Flower and Hayes, 1980), such as planning and revising. According to these studies, writers who follow a planning strategy tend to plan before beginning to write, whereas writers who prefer the revising strategy tend to plan by writing a rough draft first and then revising it. Despite the high-value of these studies, it is important to note that they only focused on analyzing the writing strategies in undergraduate (Torrance et al., 1994, 1999, 2000; Biggs et al., 1999; Lavelle et al., 2002; Arias-Gundín and Fidalgo, 2017; Robledo Ramón et al., 2018) and secondary-school students (Kieft et al., 2006, 2008). To our knowledge, just one study has explored the use of different writing strategies with upper-primary Flemish students (De Smedt et al., 2018). In this study, the authors implemented the Writing Strategies Questionnaire initially developed by Kieft et al. (2006, 2008) and identified four factors by means of exploratory and confirmatory factor analysis which were labeled thinking, planning, revising and controlling. The planning and revising strategies were consistent with those identified in previous studies with secondary school students (Kieft et al., 2006, 2008). However, in that study the authors found two additional factors. The controlling factor was defined as students' tendency to check the content or structure of their text, whereas the thinking factor make reference to the extent to which students first think about the content of their text and about their

writing approach before they start writing. Thus, according to this study, it seems to be that the questionnaire assesses writing strategies in a more comprehensive way than initially intended by Kieft et al. (2006, 2008).

Additionally, it is important to consider that in all the previously reported studies, data were collected independently of the writing task through questionnaires, which may have led to biases due to self-reported estimates of writing strategies (Fidalgo and García, 2009). However, it is difficult to think of a feasible alternative for exploring writing strategies which would allow researchers to collect data from a representative sample size. Therefore, it is vitally important to conduct studies to explore the psychometric properties and the validity of these questionnaires. The advantages of exploring these aspects of the Writing Strategies Questionnaire would be the possibility of capturing students' strategy preferences non-intrusively, exploring some aspects that remain unclear about writing style (i.e., stability), and the possibility of comparing student outcomes according to their writing strategy preference in intervention studies as one key individual feature of writers at different ages (Kieft et al., 2008).

Therefore, the main goal of the present study is to analyze the factor structure and validity of a Spanish version of the Writing Strategies Questionnaire (WSQ-SP) (Kieft et al., 2006, 2008) implemented with Spanish upper-primary students, analyzing the adjustment of the factorial model proposed based on the scientific literature (De Smedt et al., 2018), which consists of four interrelated factors taking into account the recursive nature of the writing process: Thinking, Planning, Revision, and Monitoring (see **Figure 1**). Additionally, the traditional two-factor model initially found (Kieft et al., 2006, 2008) will also be explored to test which is the most appropriate scale structure for the questionnaire.

Moreover, a second goal of the study is to analyze the factorial invariance of the proposed model by considering different variables such as gender and grade.

MATERIALS AND METHODS

Participants

The sample comprised 651 Spanish primary school students in 16 fourth-grade ($N = 178$, 27%), 16 fifth-grade ($N = 246$; 38%), and 14 sixth-grade classes ($N = 227$; 35%). Students' ages ranged from 9 to 13 (Mage = 9.5 years, $SD = 0.55$ for fourth graders; Mage = 10.4 years, $SD = 0.52$ for fifth graders; Mage = 11.5 years, $SD = 0.54$ for sixth graders) and with similar proportions of boys and girls (47.19% girls in 4th grade; 48.37% girls in 5th grade; 55.07% girls in 6th grade). The students came from seven public and four semi-private schools in the city of Ponferrada, finding students from families with a high diversity of socio-economic status. However, it should be noted that most students came from families with medium to high incomes.

The criterion for choosing the participants of the study was that they should be students in 5th or 6th grade of elementary education and that Spanish should be their first language. Students in their final years of primary education were considered for developmental reasons. According to the studies

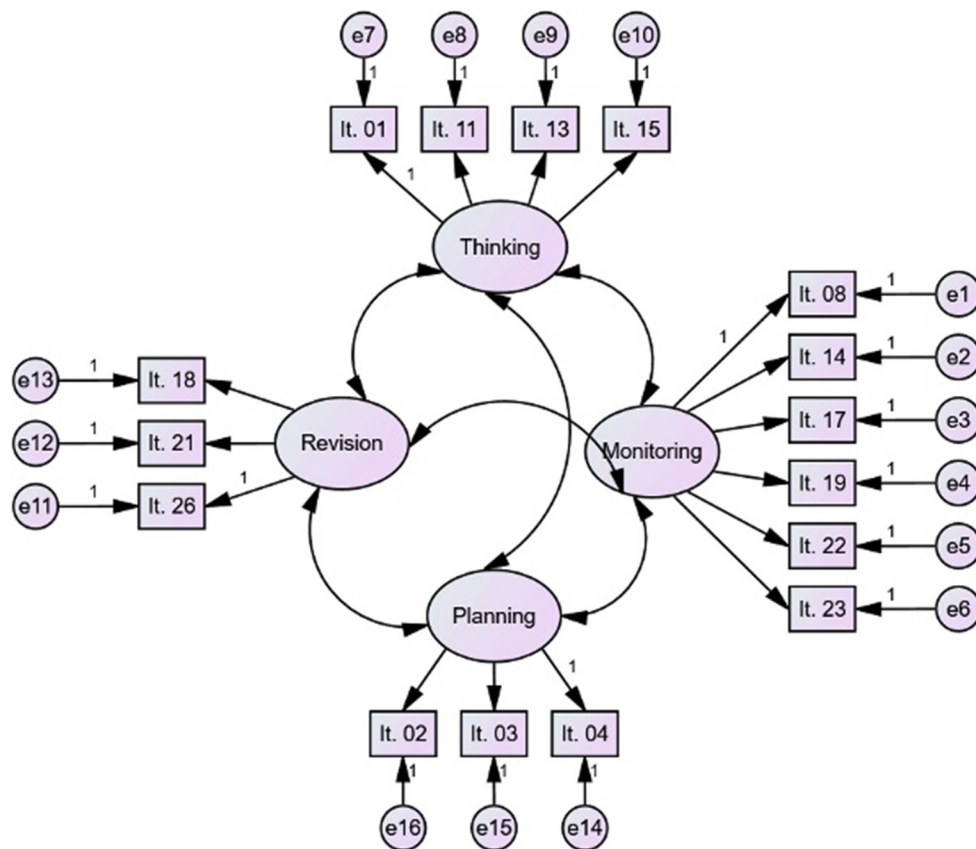


FIGURE 1 | Hypothesized model of the factor structure of the WSQ-SP, composed of four interrelated factors.

of Berninger et al. (1992, 1994, 1996), planning and revision skills appear progressively during the primary education stage, with the last processes appearing in the last grades (5th and 6th). Additionally, although students with learning disabilities participated in the study, their data was not considered for the analysis. This was done on the basis of previous studies, which have shown differences in the use of high-level cognitive processes between upper-primary students with and without learning disabilities (García and Fidalgo, 2008; Graham et al., 2017b).

Procedure

Prior to the implementation of the study, consent was requested from the Consejería de Educación de Castilla y León [Regional Department of Education of Castilla and Leon], the autonomous community in which the study was carried out. Once the study was approved by the expert committee of the regional department of Education, the researchers contacted all the schools in Ponferrada and surrounding areas. Subsequently, a meeting was held with the heads of the schools to inform them in detail about the study and the procedure to be followed during its execution. Those schools that decided to participate in the study sent the parents an information letter in which the research aims were presented, asking them for

informed consent for their children to participate in the study. They were given the opportunity to express concerns and to request that their children's data not be included in the study. Following that, the study was undertaken with participation from only those students whose parents had given informed consent. The study was conducted following the Code of Ethics of the World Medical Association (Declaration of Helsinki) (Williams, 2008).

Data was collected in a natural context within regular Spanish language classes. Students were asked to complete the Spanish WSQ and writing a narrative text in a 50-min session. The questionnaire was administered by one of the researchers in this study who has a degree in Psychology and experience in administering similar kind of tests. Additionally, she received specific training on the implementation of the questionnaire. Moreover, the assessment session was audio-recorded to make sure that the assessment procedure occurred as intended.

Measures

Students' Writing Strategies

In this study, we began with the 26-item questionnaire measuring students' writing strategies that has been used in previous studies

(Kieft et al., 2006, 2008). Students rate their agreement with each item on a five-point scale (1–5).

For the translation of the questionnaire, we combined direct and inverse translation of the items. The questionnaire was translated from Dutch to English by a Dutch researcher who was also fluent in English and Spanish. Then this researcher and a member of the Spanish team each separately translated the English version into Spanish, in order to compare the two versions. The two Spanish translations were compared and discussed, looking for possible discrepancies.

Following that, an expert-panel assessed the suitability of the questionnaire according to the age of the target population. This panel of experts was made up of five schoolteachers with extensive experience in education (three in primary education, one in early childhood education and one in special needs education). Some changes were made to the wording to improve the understanding of the meaning of some items.

The first version of the questionnaire was then trialed with a small sample of upper-primary students to identify possible mistakes and assess general understanding. Students had no issues with it, hence no changes were made, and this produced the final version of the questionnaire (see **Supplementary Material**).

Data Analysis

In order to explore the psychometric properties of the questionnaire, we first analyzed the normal distribution of each item, verifying that they gave kurtosis and skewness indices between ± 7 and ± 3 , respectively (Kline, 2011). The magnitude and direction of the relationship between items was also analyzed using Pearson's correlation coefficient.

The validity of the factor structure was analyzed in two steps. First, we conducted exploratory factor analysis (EFA) with the aim of determining whether the items saturated the two factors of the original version or the four factors proposed in the present study (see **Figure 1**). Second, we performed a confirmatory factor analysis (CFA).

The maximum likelihood method was used to estimate the model using the covariance matrix of the items in order to analyze the fit of the proposed model. In order to investigate the model's goodness of fit, a number of statistics were considered: (a) absolute indices such as the Chi-square ratio and degrees of freedom (X^2/df) and the goodness-of-fit index (GFI); (b) the comparative fit index (CFI) as an incremental fit index; (c) the adjusted goodness-of-fit index (AGFI) and the root mean square error of approximation (RMSEA) as parsimony adjustment indices. The goodness-of-fit of the model was assessed according to the following rules: (a) the X^2/df ratio is < 3 ; (b) values above 0.90 for the goodness-of-fit index (GFI), comparative goodness-of-fit index (CFI) and adjusted goodness-of-fit index (AGFI) are acceptable; (c) values below 0.08 for the root mean square error of approximation (RMSEA) indicate acceptable model fit (Browne and Cudeck, 1993; Hoyle, 1995; Kline, 1998; Hu and Bentler, 1999; Valdés et al., 2019).

Finally, the factorial invariance of the proposed model was analyzed by testing the fit of the model using confirmatory factor analysis (CFA) and composite reliability considering the variables gender and school year.

RESULTS

First, the results of the exploratory factor analysis (EFA) of the WSQ-SP are provided in order to check the factor structure of the proposed model, as well as the loading of the items on each of the factors. Second, the results of the confirmatory factor analysis (CFA) are presented showing the fit of the proposed model, as well as a comparison with the traditional two-dimensional model. Finally, the results are presented with respect to the factorial invariance of the WSQ-SP questionnaire considering gender and grade.

Exploratory Factor Analysis (EFA)

All of the items exhibited values within the range of normal distribution (asymmetry: ranging between -1.37 and 1.20 ; kurtosis: ranging between -0.96 and 0.82), hence the hypothesis of univariate normality was rejected in all cases (Kline, 2011).

An exploratory factor analysis (EFA) was carried out using the Maximum Likelihood extraction method and Oblimin rotation. The data showed a good fit for this kind of model, evidenced by Bartlett's sphericity test ($\chi^2(171) = 2216.68, p < 0.000$) and the Kaiser-Meyer-Olkin (KMO) value of 0.83 (Lloret-Segura et al., 2014). As a criterion for item inclusion, factor weights > 0.30 were considered for only one of the factors, reflecting the theoretical soundness of the scale (Hair et al., 1999). Ten items were excluded because they did not match the different factors (items 5, 6, 7, 9, 10, 12, 15, 16, 20, and 24). The results showed that the 16 items of the scale are grouped into four factors which were theoretically identified and retained. These factors were labeled revising, monitoring, thinking, and planning and together explain 32.08% of the variance. The first factor, *monitoring*, corresponds to how much students checked the content or structure of their text during composition. This factor consisted of six items explaining 18.0% of the variance and had a composite reliability of 0.82 . The second factor, *revising*, is related to how much students revised the content of their text once the text was written. This factor included three items explaining 7.9% of the variance and had a composite reliability of 0.85 . The third factor, *planning*, is related to how much students thought about the content of their text in advance, using external planning devices such as a draft sheet. This factor included three items explaining 3.8% of the variance and had a composite reliability of 0.75 . Finally, the fourth factor, *thinking*, corresponds to how much students needed to have a clear idea of the content or structure of the text in their minds before they started to write. This factor consisted of four items explaining 2.4% of the variance and a composite reliability of 0.79 (see **Table 1**).

Confirmatory Factor Analysis (CFA)

We performed CFA for the 16 items in the WSQ-SP using Amos software in SPSS. We used Maximum Likelihood (ML) factor analysis with the CFA command. The results of the CFA suggest that overall, the model had a good fit to the data according to the indices ($\chi^2/df = 2.23$; GFI = 0.96 ; AGFI = 0.95 ; CFI = 0.93 ; RMSEA = 0.04 CI $(0.03-0.05)$).

The values of the regression coefficients suggest that the factors explained an acceptable part of the variance of the items

TABLE 1 | Exploratory factor analysis (EFA) of the WSQ-SP.

Items	Factor loadings				Commonality
	F1	F2	F3	F4	
8. While writing, I regularly check whether my text doesn't contain sentences that are too long or incorrect.	0.50				0.29
^R 14. When writing, I sometimes write paragraphs of which I know that they are not yet correct, but I prefer to continue writing.	0.38				0.18
^R 17. I usually hand in my text without checking whether the paragraphs are well arranged.	0.70				0.46
19. Before I hand in my text, I check whether it is structured logically.	0.70				0.58
^R 22. When I reread my texts, sometimes they are very chaotic.	0.33				0.12
23. I have to reread the texts I wrote, to prevent redundancies.	0.33				0.27
18. When I reread and rewrite my text, the structure of my text changes a lot.		0.71			0.51
21. When I rewrite my texts, the content often changes a lot.		0.73			0.50
26. When I finished writing, I reread and improve a lot: there might change a lot in my text.		0.56			0.40
2. I always use a diagram before I start to write.			−0.39		0.16
3. Before writing a text, I jot down some notes on a scribbling paper. Later, I elaborate these notes.			−0.83		0.62
4. Before I start to write a text, I prefer to write down some thoughts on a scribbling paper to discover what I think about the topic.			0.65		0.43
1. When I write a text, I spend a lot of time thinking on how to approach it.				0.33	0.19
11. I need to have my thoughts clear, before I can start to write.				0.60	0.32
13. Before I write down a sentence, I have it clear in my mind.				0.41	0.19
15. Writing helps me to clarify my thoughts.				0.65	0.17
				Correlations	
Monitoring	−				
Revision	0.03	−			
Planning	0.27*	0.21*	−		
Thinking	0.45*	0.22*	0.27*	−	

* $p < 0.01$. ^RItems recoded in the analyses.

(see **Figure 2**). The correlation between the factors indicated that the factors were related but did not present problems of collinearity.

Considering that the proposed model was corroborated by the results, it was compared with the traditional two-dimensional structure identified in previous studies (Kieft et al., 2006, 2008). The model proposed in this study exhibited the best factorial fit (see **Table 2**).

Factor Invariance Analysis

To check that the effectiveness of the model was not significantly affected by the features of the sample, the proposed model was subjected to CFA by selecting the sample based on gender and grade. These two variables were chosen for the following reasons. Gender was considered because some studies have shown it to be a variable that can influence student learning and achievement in general (e.g., Reilly et al., 2019) and specifically in the use of cognitive writing strategies (e.g., Berninger et al., 1992; Jones, 2011). Additionally, grade was chosen because it is during this period of schooling that higher-level cognitive processes related to textual planning and revision appear following different rates of development (Berninger et al., 1992, 1994, 1996). The aim was to ensure that the questionnaire is reliable regardless of gender or grade.

As **Table 3** shows, the composite reliability of each factor in all of the proposed models, based on the characteristics of the sample and their combinations, is high (ranging between: 0.81 and 0.92 for the monitoring factor; 0.70 and 0.93 for the thinking factor; 0.70 and 0.82 for the planning factor; and 0.81 and 0.91 for the revising factor). The model shows a good overall fit for the gender and grade variables, with the indicators meeting the established parameters. There was just one exception for the adjusted goodness-of-fit index (AGFI) in the case of 4th grade students (0.88), which was very close to the desired value (0.90). When the model was analyzed based on the interaction of gender and grade, the absolute index, Chi-square ratio and degrees of freedom, and the RMSEA as the parsimony adjustment index, demonstrated acceptable model fit, with the remaining indicators being close to the desired value (0.90). However, it is important to note that when the model was analyzed based on gender-grade interaction, the sample shrank considerably. This influenced the results, given that CFA is sensitive to sample size. The literature recommends performing CFA analysis with samples of more than 200 participants (Valdés et al., 2019). In all of the cases analyzing the model with samples smaller than 200 students, some indicators did not give the desired values, as **Table 3** shows.

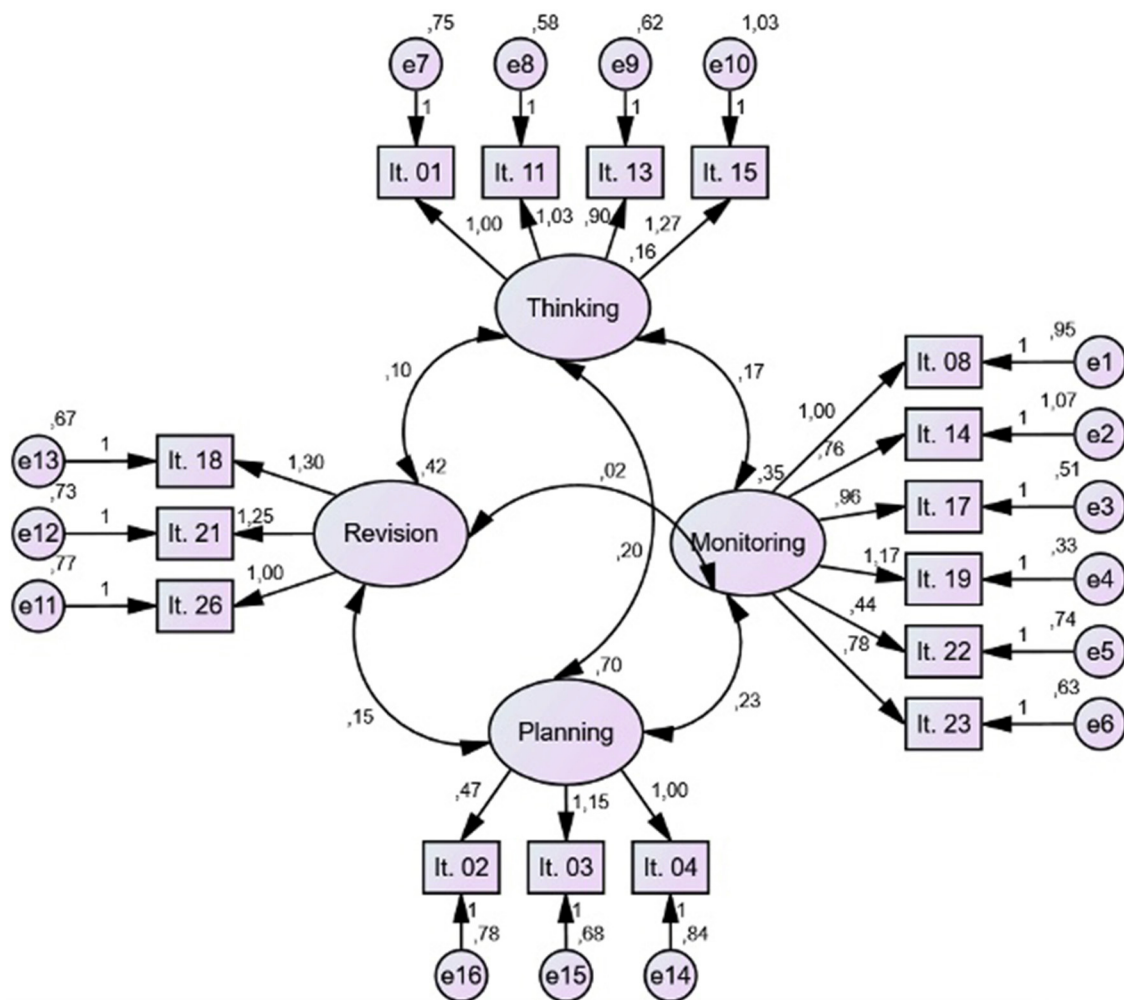


FIGURE 2 | Path diagram of the hypothesized model. Confirmatory Factor Analysis of the questionnaire.

TABLE 2 | Goodness of fit indices for each model of the CFA of the WSQ-SP ($N = 651$).

Model	χ^2/df	GFI	AGFI	CFI	RMSEA (IC at 90%)
4 Factors*	2.23	0.96	0.95	0.93	0.04 (0.03–0.05)
2 Factors**	6.83	0.86	0.81	0.65	0.09 (0.08–0.10)

*Proposed model with 4 factors: Monitoring, Revision, Planning, and Thinking.

**Traditional two-dimensional model: Planning and Revision.

DISCUSSION

The main goal of the present study was to analyze the factor structure and validity of the Spanish WSQ-SP with upper-primary students. An additional goal was to analyze the factorial invariance of the proposed model by considering different variables such as gender and grade.

With regard to the first goal of the study, the results relating to the questionnaire's factor structure were in line with the previous study carried out with Flemish upper-primary students

(De Smedt et al., 2018) in which four factors were identified; planning, revising, monitoring and thinking. In addition, on comparing this model with the two-factor model (i.e., planning and revising), generally identified in previous studies with more expert writers (Kieft et al., 2006, 2008), the four-factor model demonstrated a better match with the questionnaire structure.

This four-factor model is consistent with the differentiation of planning and revision processes that have generally been considered in terms of their occurrence during the process of writing a text (Berninger et al., 1994). As planning and revision

TABLE 3 | Goodness-of-Fit Indices for the proposed model of the questionnaire based on sample features.

Model (n)	χ^2	p	χ^2/df	GFI	AGFI	CFI	RMSEA (IC al 90%)	Composite reliability			
								Monitoring	Thinking	Planning	Revising
Gender											
Girls (<i>n</i> = 328)	176.92	0.000	1.80	0.94	0.91	0.90	0.05 (0.04–0.06)	0.86	0.76	0.72	0.88
Boys (<i>n</i> = 323)	150.64	0.001	1.54	0.94	0.92	0.93	0.04 (0.03–0.06)	0.86	0.84	0.80	0.86
Grade											
4th (<i>n</i> = 178)	147.36	0.001	1.50	0.91	0.88	0.91	0.05 (0.03–0.07)	0.86	0.85	0.74	0.81
5th (<i>n</i> = 246)	151.01	0.000	1.54	0.93	0.90	0.91	0.05 (0.03–0.06)	0.87	0.73	0.77	0.86
6th (<i>n</i> = 227)	118.67	0.076	1.21	0.94	0.92	0.97	0.03 (0.00–0.05)	0.86	0.74	0.76	0.89
Gender and Grade											
4th Girls (<i>n</i> = 82)	117.23	0.090	1.19	0.85	0.80	0.91	0.05 (0.00–0.08)	0.81	0.74	0.76	0.89
5th Girls (<i>n</i> = 119)	136.63	0.006	1.39	0.88	0.83	0.89	0.06 (0.03–0.08)	0.85	0.70	0.79	0.87
6th Girls (<i>n</i> = 125)	146.16	0.001	1.49	0.88	0.83	0.85	0.06 (0.04–0.08)	0.92	0.85	0.73	0.91
4th Boys (<i>n</i> = 93)	111.56	0.165	1.14	0.88	0.84	0.95	0.04 (0.00–0.07)	0.86	0.93	0.79	0.83
5th Boys (<i>n</i> = 97)	161.09	0.000	1.64	0.83	0.78	0.73	0.08 (0.06–0.10)	0.86	0.85	0.70	0.84
6th Boys (<i>n</i> = 100)	102.62	0.355	1.05	0.90	0.86	0.99	0.02 (0.00–0.06)	0.86	0.71	0.82	0.90

can occur before or during translating, a distinction was made between advanced and online planning, post-translation and online revision. In this way, the *thinking* and *planning* factors were related to the two different, but complementary, ways of planning. According to previous studies, writers differ in how they plan. While some writers make an outline in note form before drafting, others plan without producing an outline. This latter form of planning has been called “mental planning” (Kellogg, 1988; Torrance et al., 2000). Thus, the *thinking* factor would correspond to mental planning while the *planning* factor would correspond to outline planning. Similarly, the *revising* and *monitoring* factors can be interpreted according to when revision occurs. According to Berninger and Swanson (1994) considering the timing of revision it is possible to differentiate between online revision (i.e., revision that takes place during composition) and post-translation revision (i.e., revision that takes place after composition). Thus, the revising factor would correspond with post-translation revision while the monitoring factor would correspond with online revision. In other words, the results of the present study indicate that the questionnaire is not only exploring students’ use of planning and revising strategies in a general way, but rather also assessing different types of planning and revision strategies depending on when they take place when students are writing a text. These results are in line with the arguments presented by Kieft et al. (2007) and Tillema et al. (2011), who pointed out that the revising scale was composed not only of items related to post-translation revision but also to monitoring. Moreover, the better fit of the four-factor model can be explained based on the fact that these processes seem to have different rates of development (Berninger et al., 1992, 1994, 1996). Based on the implementation of cross-sectional studies with students aged between 6 and 15 years old, the authors found that online planning and revision seems to appear at around ages 6–9 (1st–3rd grades). The authors also found that advanced planning and post-translation revision were

the last processes to appear around the last years of primary school (ages 9–12; 4th–6th grades). This would clearly explain why the four-factor model has a better fit to the data from primary school pupils. Here, it is also important to consider that the four factors were shown to exhibit correlation but no problems of collinearity were found. This result is in line the view of writing as a recursive activity in which one process may interrupt others during composition (Flower and Hayes, 1980).

In terms of the second goal of the study, analyzing the factorial invariance of the proposed model by considering different variables such as gender and grade, the results showed that the questionnaire structure was independent of the feature of the sample. The results of the present study seem to be generalizable to upper-primary students regardless of gender or grade.

In summary, the major contribution of this study is the validation of the WSQ-SP with upper-primary students, as validation is a critical step for the development of reliable measurement tools in all scientific domains (Muñiz and Fonseca-Pedrero, 2019). From this study, we can conclude that the questionnaire provides more precise information than initially expected and it is a suitable tool for easily, reliably assessing upper-primary students’ writing.

The validation of this questionnaire is a first step toward a reliable analysis of this variable, which will continue with analyzing aspects that have not yet been investigated, such as its stability, the moderating effect it has on writing intervention in upper-primary students (Kieft et al., 2006, 2008), and the effect of instruction itself on writing. Having a validated questionnaire will also make it possible to analyze the relationship between students’ use of strategies and other important writing-related variables such as reading (Fidalgo et al., 2014; Qin and Liu, 2021), motivation (Rocha et al., 2019), and students’ knowledge (Wijekumar et al., 2019). It would also be interesting to analyze the relationship between the results provided by this scale and

the writing processes students follow through the use of online measures such as the triple task (García and Fidalgo, 2008; Fidalgo et al., 2014) and thinking aloud (López et al., 2019).

Finally, as an educational contribution, this instrument may be a useful tool that will help provide teachers with information about their students' strategies and consequently help them to adapt the writing instruction according to their needs. All of this, without a doubt, will have a positive impact on students' writing performance, not only in initial educational levels (e.g., López et al., 2017), but also in later educational stages, such as at University, where students often find it difficult to write academic texts (Connelly et al., 2005).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, under request to the corresponding author, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Consejería de Educación de Castilla y León. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

REFERENCES

- Alamargot, D., and Chanquoy, L. (2001). *Studies in Writing Series: Vol. 9. Through the Models of Writing*. Dordrecht: Kluwer Academic. doi: 10.1007/978-94-010-0804-4
- Alves, R. A., Limpo, T., Fidalgo, R., Carvalhais, L., Pereira, L. Á., and Castro, S. L. (2016). The impact of promoting transcription on early text production: effects on bursts and pauses, levels of written language, and writing performance. *J. Educ. Psychol.* 108, 665–679. doi: 10.1037/edu0000089
- Arias-Gundín, O., and Fidalgo, R. (2017). El perfil escritor como variable moduladora de los procesos involucrados en la composición escrita en estudiantes universitarios [Writer profile as a modulating variable of processes involved in written composition in undergraduate students]. *EJIHPE* 7, 59–68. doi: 10.30552/ejihpe.v7i1.195
- Arrimada, M., Torrance, M., and Fidalgo, R. (2019). Effects of teaching planning strategies to first-grade writers. *Br. J. Educ. Psychol.* 89, 670–688. doi: 10.1111/bjep.12251
- Beauvais, C., Olive, T., and Passerault, J. M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *J. Educ. Psychol.* 103, 415–428. doi: 10.1037/a0022545
- Berninger, V. W., Cartwright, A. C., Yates, C. M., Swanson, H. L., and Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: shared and unique functional systems. *Read. Writ.* 6, 161–196. doi: 10.1007/BF01026911
- Berninger, V. W., and Swanson, H. L. (1994). "Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing," in *Children's Writing: Toward a Process Theory of the Development of Skilled Writing*, Vol. 2, ed E. C. Butterfield (Greenwich, CT: JAI Press), 57–81.
- Berninger, V. W., Whitaker, D., Feng, Y., Swanson, H. L., and Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *J. Sch. Psychol.* 34, 23–52. doi: 10.1016/0022-4405(95)00024-0
- Berninger, V. W., Yates, C. M., Cartwright, A. C., Rutberg, J., Remy, E., and Abbott, R. D. (1992). Lower-level developmental skills in beginning writing. *Read. Writ.* 4, 257–280. doi: 10.1007/BF01027151

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the design of the work, analysis and data interpretation, drafting and revising it critically, and approved it for publication.

FUNDING

This research has been supported by the Spanish Ministry of Economy and Competitiveness through a project EDU2015-67484-P (MINECO/FEDER).

ACKNOWLEDGMENTS

We would like to thank staff and students at the Peñalba, Flores del Sil, Espíritu Santo, San Antonio, Jesús Maestro, La Borreca, Navaliegos, Valentín García Yebra, Concepcionistas, Asunción, and San Ignacio schools for their assistance in completing this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.700770/full#supplementary-material>

- Biggs, J., Lai, P., Tang, C., and Lavelle, E. (1999). Teaching writing to ESL graduate students. A model and an illustration. *Br. J. Educ. Psychol.* 69, 293–306. doi: 10.1348/000709999157725
- Browne, M., and Cudeck, R. (1993). "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, eds K. A. Bollen and J. S. Long (Newbury Park, CA: SAGE Publications), 136–162.
- Connelly, V., Dockrell, J. E., and Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educ. Psychol.* 25, 99–107. doi: 10.1080/0144341042000294912
- De Smedt, F., Merchie, E., Barendse, M., Rosseel, Y., De Naeghel, J., and Van Keer, H. (2018). Cognitive and motivational challenges in writing: studying the relation with writing performance across students' gender and achievement level. *Read. Res. Q.* 53, 249–272. doi: 10.1002/rrq.193
- Fidalgo, R., and García, J. N. (2009). La evaluación de la metacognición en la composición escrita [Evaluating metacognition in written composition]. *Estud. Psicol.* 30, 51–72. doi: 10.1174/021093909787536290
- Fidalgo, R., Torrance, M., Arias-Gundín, O., and Martínez-Cocó, B. (2014). Comparison of reading-writing patterns and performance of students with and without reading difficulties. *Psicothema* 26, 442–448. doi: 10.7334/psicothema2014.23
- Flower, L., and Hayes, J. R. (1980). "The dynamics of composing: making plans and juggling constraints," in *Cognitive Processes in Writing*, eds L. W. Gregg and E. R. Steinberg (Hillsdale, NJ: Lawrence Erlbaum Associates), 31–49.
- García, J. N., and Fidalgo, R. (2008). Orchestration of writing processes and writing products: a comparison of sixth-grade students with and without learning disabilities. *Learn. Disabil. Contemp. J.* 6, 77–98.
- Graham, S., Collins, A. A., and Rigby-Wills, H. (2017a). Writing characteristics of students with learning disabilities and typically achieving peers: a meta-analysis. *Except. Child.* 83, 199–218. doi: 10.1177/0014402916664070
- Graham, S., and Harris, K. (2018). "Evidence-based writing practices: a meta-analysis of existing meta-analysis," in *Design Principles for Teaching Effective Writing: Theoretical and Empirical Grounded Principles*, eds R. Fidalgo, K. Harris, and M. Braaksma (Leiden: Brill Editions), 13–37. doi: 10.1163/9789004270480_003

- Graham, S., and Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educ. Psychol.* 35, 3–12. doi: 10.1207/S15326985EP3501_2
- Graham, S., Harris, K. R., Fishman, E., Houston, J., Wijekumar, K., Lei, P. W., et al. (2019). Writing skills, knowledge, motivation, and strategic behavior predict students' persuasive writing performance in the context of robust writing instruction. *Element. Sch. J.* 119, 487–510. doi: 10.1086/701720
- Graham, S., Harris, K. R., Kiuahara, S. A., and Fishman, E. J. (2017b). The relationship among strategic writing behavior, writing motivation, and writing performance with young, developing writers. *Element. Sch. J.* 118, 82–104. doi: 10.1086/693009
- Hair, J., Anderson, R., Tatham, R., and Black, W. (1999). *Análisis multivariante*, 5 Edn. Madrid: Prentice Hall.
- Harris, K. R., Santangelo, T., and Graham, S. (2010). "Metacognition and strategies instruction in writing," in *Metacognition, Strategy Use, and Instruction*, eds H. S. Waters and W. Schneider (New York, NY: The Guilford Press), 226–256.
- Hayes, J. R. (1996). "A new framework for understanding cognition and affect in writing," in *The Science of Writing: Theories, Methods, Individual Differences, and Applications*, eds C. M. Levy and S. Ransdell (Mahwah, NJ: Lawrence Erlbaum Associates), 1–27.
- Hayes, J. R. (2012). "Evidence from language bursts, revision, and transcription for translation and its relation to other writing processes," in *Translation of Thought to Written Text While Composing: Advancing Theory, Knowledge, Research, Methods, Tools, and Applications*, eds M. Fayol, D. Alamargot, and V. W. Berninger (New York, NY: Psychology Press), 15–25.
- Hoyle, R. H. (1995). *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks, CA: Sage.
- Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equat. Model.* 6, 1–55. doi: 10.1080/10705519909540118
- Jones, S. (2011). Mapping the landscape: gender and the writing classroom. *J. Writ. Res.* 3, 161–179. doi: 10.17239/jowr-2012.03.03.2
- Kellogg, R. T. (1988). Attentional overload and writing performance: effects of rough draft and outline strategies. *J. Exp. Psychol.* 14, 355–365. doi: 10.1037/0278-7393.14.2.355
- Kieft, M., Rijlaarsdam, G., Galbraith, D., and Van den Bergh, H. (2007). The effects of adapting a writing course to students' writing strategies. *Br. J. Educ. Psychol.* 77, 565–578. doi: 10.1348/096317906X120231
- Kieft, M., Rijlaarsdam, G., and Van den Bergh, H. (2006). Writing as a learning tool: testing the role of students' writing strategies. *Eur. J. Psychol. Educ.* 12, 17–34. doi: 10.1007/BF03173567
- Kieft, M., Rijlaarsdam, G., and van den Bergh, H. (2008). An aptitude-treatment interaction approach to writing-to-learn. *Learn. Instruct.* 18, 379–390. doi: 10.1016/j.learninstruct.2007.07.004
- Kline, P. (1998). *The New Psychometrics: Science, Psychology, and Measurement*. London: Psychology Press.
- Kline, R. B. (2011). "Convergence of structural equation modeling and multilevel modeling," in *Handbook of Methodological Innovation in Social Research Methods*, eds M. Williams and W. P. Vogt (London: Sage), 562–589. doi: 10.4135/9781446268261.n31
- Lavelle, E., Smith, J., and O'Ryan, L. (2002). The writing approaches of secondary students. *Br. J. Educ. Psychol.* 72, 399–418. doi: 10.1348/000709902320634564
- Limpo, T., Alves, R. A., and Connelly, V. (2017). Examining the transcription-writing link: effects of handwriting fluency and spelling accuracy on writing performance via planning and translating in middle grades. *Learn. Individ. Diff.* 53, 26–36. doi: 10.1016/j.lindif.2016.11.004
- Llaurado, A., and Dockrell, J. E. (2020). The impact of orthography on text production in three languages: catalan, English, and Spanish. *Front. Psychol.* 11:878. doi: 10.3389/fpsyg.2020.00878
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., and Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Ann. Psychol.* 30, 1151–1169. doi: 10.6018/analesps.30.3.199361
- López, P., Torrance, M., and Fidalgo, R. (2019). The online management of writing processes and their contribution to text quality in upper-primary students. *Psicothema* 31, 311–318. doi: 10.7334/psicothema2018.326
- López, P., Torrance, M., Rijlaarsdam, G., and Fidalgo, R. (2017). Effects of direct instruction and strategy modeling on upper-primary students' writing development. *Front. Psychol.* 8:1054. doi: 10.3389/fpsyg.2017.01054
- McCutchen, D. (2011). From novice to expert: implications of language skills and writing-relevant knowledge for memory during the development of writing skill. *J. Writ. Res.* 3, 51–68. doi: 10.17239/jowr-2011.03.01.3
- Muñiz, J., and Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test [Ten steps for test development]. *Psicothema* 31, 7–16. doi: 10.18682/pd.v11i.854
- Pontart, V., Bidet-Ildei, C., Lambert, E., Morisset, P., Flouret, L., and Alamargot, D. (2013). Influence of handwriting skills during spelling in primary and lower secondary grades. *Front. Psychol.* 4:818. doi: 10.3389/fpsyg.2013.00818
- Puranik, C. S., Boss, E., and Wanless, S. (2019). Relations between self-regulation and early writing: domain specific or task dependent?. *Early Child. Res. Q.* 46, 228–239. doi: 10.1016/j.jecresq.2018.02.006
- Qin, J., and Liu, Y. (2021). The influence of reading texts on L2 reading-to-write argumentative writing. *Front. Psychol.* 12:655601. doi: 10.3389/fpsyg.2021.655601
- Reilly, D., Neumann, D. L., and Andrews, G. (2019). Gender differences in reading and writing achievement: evidence from the National Assessment of Educational Progress (NAEP). *Am. Psychol.* 74:445. doi: 10.1037/amp0000356
- Robledo Ramón, P., Arias-Gundín, O., Palomo, M., Andina, E., and Rodríguez, C. (2018). Perfil escritor y conocimiento metacognitivo de las tareas académicas en los estudiantes universitarios. *Publicaciones* 48, 243–270. doi: 10.30827/publicaciones.v48i1.7335
- Rocha, R. S., Filipe, M., Magalhães, S., Graham, S., and Limpo, T. (2019). Reasons to write in grade 6 and their association with writing quality. *Front. Psychol.* 10:2157. doi: 10.3389/fpsyg.2019.02157
- Santangelo, T., Harris, K., and Graham, S. (2016). "Self-regulation and writing," in *Handbook of Writing Research, 2nd Edn*, eds C. A. MacArthur, S. Graham, and J. Fitzgerald (New York, NY: Guilford Press), 174–193.
- Tillema, M., van den Bergh, H., Rijlaarsdam, G., and Sanders, T. (2011). Relating self-reports of writing behaviour and online task execution using a temporal model. *Metacogn. Learn.* 6, 229–253. doi: 10.1007/s11409-011-9072-x
- Torrance, M., Thomas, G. V., and Robinson, E. J. (1994). The writing strategies of graduate research students in the social sciences. *High. Educ.* 27, 379–392. doi: 10.1007/BF03179901
- Torrance, M., Thomas, G. V., and Robinson, E. J. (1999). Individual differences in the writing behaviour of undergraduate students. *Br. J. Educ. Psychol.* 69, 189–199. doi: 10.1348/000709999157662
- Torrance, M., Thomas, G. V., and Robinson, E. J. (2000). Individual differences in undergraduate essay writing strategies. A longitudinal study. *High. Educ.* 39, 181–200. doi: 10.1023/A:1003990432398
- Valdés, A. A., García, F. I., Torres, G. M., Urías, M., and Grijalva, C. S. (2019). *Medición en Investigación Educativa con Apoyo del SPSS y el AMOS [Measurement in Educational Research with Support of SPSS and AMOS]*. Clave Editorial.
- Wijekumar, K., Graham, S., Harris, K. R., Lei, P. W., Barkel, A., Aitken, A., et al. (2019). The roles of writing knowledge, motivation, strategic behaviors, and skills in predicting elementary students' persuasive writing from source material. *Read. Writ.* 32, 1431–1457. doi: 10.1007/s11145-018-9836-7
- Williams, J. R. (2008). Revising the declaration of Helsinki. *World Med. J.* 54, 120–125.
- Zeidner, M., Boekaerts, M., and Pintrich, P. R. (2000). "Self-regulation: directions and challenges for future research," in *Self-Regulation: Theory, Research, and Applications*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (Orlando, FL: Academic Press), 749–768. doi: 10.1016/B978-012109890-2/50052-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Arias-Gundín, Real, Rijlaarsdam and López. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Beyond the Educational Context: Relevance of Intrinsic Reading Motivation During COVID-19 Confinement in Spain

Raquel De Sixte^{1*}, Inmaculada Fajardo², Amelia Mañá², Álvaro Jáñez¹, Marta Ramos¹, María García-Serrano¹, Federica Natalizi³, Barbara Arfé³ and Javier Rosales¹

¹ Department of Developmental and Educational Psychology, Universidad de Salamanca, Salamanca, Spain, ² Estructura de Recerca Interdisciplinària de Lectura, Universitat de València, Valencia, Spain, ³ Department of Developmental Psychology and Socialization, Università degli Studi di Padova, Padua, Italy

OPEN ACCESS

Edited by:

María Isabel
de Vicente-Yagüe Jara,
University of Murcia, Spain

Reviewed by:

Juan Ramón Moreno-Vera,
University of Murcia, Spain
Jose Gomez Galan,
Ana G. Mendez University System,
Puerto Rico
Benito León Del Barco,
University of Extremadura, Spain
Raquel Gilar-Corbi,
University of Alicante, Spain

*Correspondence:

Raquel De Sixte
rsixte@usal.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 30 April 2021

Accepted: 08 June 2021

Published: 12 July 2021

Citation:

De Sixte R, Fajardo I, Mañá A,
Jáñez Á, Ramos M,
García-Serrano M, Natalizi F, Arfé B
and Rosales J (2021) Beyond
the Educational Context: Relevance
of Intrinsic Reading Motivation During
COVID-19 Confinement in Spain.
Front. Psychol. 12:703251.
doi: 10.3389/fpsyg.2021.703251

What role could have intrinsic motivation toward reading in an extraordinary situation like the recent confinement? This research examines the relationship between intrinsic reading motivation (IRM) and reading habits in an adult population considering types of reading (for leisure, work/study, social networks, and news), gender, and distress generated by the coronavirus disease 2019 (COVID-19) pandemic. Participants were 3,849 adults from Spain who were surveyed about their reading practices: before, during the first weeks, and after several weeks of confinement. Linear mixed effects models (LMMs) were used to analyze data. Results showed a three-way interaction between reading frequency, IRM, and type of reading. Also, distress seems to pose a differential impact depending on the type of reading. The higher the IRM, the lesser the time devoted to study/work reading and the more to social and news reading (at the beginning of confinement). In this sense, IRM can function as a protective factor of reading behavior but only for leisure reading. Results support previous findings of the importance of consciously promoting this type of motivation in all individuals beyond educational contexts, since it seems to be positively related to well-being. Other results and implications are discussed.

Keywords: intrinsic reading motivation, reading behavior, gender, distress, COVID-19

INTRODUCTION

Due to coronavirus disease 2019, the Spanish population was subjected to strict lockdown since March 13, 2020. Recent research has shown how this situation caused many changes in the habits of people, such as eating (Pérez-Rodrigo et al., 2020), use of drugs, TV consumption, time for videogames, and physical exercise (Balluerka et al., 2021). It seems clear that from that moment on, most of the work tasks and whole leisure activities were moved inside the houses, something that had not happened before, greatly altering them in many cases. Reading was one of the daily activities that were most affected by this radical change in habits (Salmerón et al., 2020). Reading habits and the amount of time devoted to different reading activities (or types of reading) are influenced by several individual factors (Scales and Rhee, 2001; Schutte and Malouff, 2007; Garcés-Bacsal and Yeo, 2017). For example, there is a positive relationship between intrinsic motivation and time spent reading, a relationship that is stronger for leisure reading (Schiefele et al., 2012). An important research question is whether motives underlying reading habits are maintained in a

situation of confinement. This exceptional context offers a unique opportunity to explore the role of intrinsic reading motivation, keeping in mind that the pandemic situation could cause a triggered or emerging situational interest (Linnenbrink-García and Patall, 2016) on certain types of reading. Research on this issue could be beneficial to support the self-determination theory (SDT) and to help us reflect on its potential implications beyond academic contexts. Another individual factor associated with reading behaviors is gender: we know that females have higher levels of intrinsic motivation to read (Wigfield and Guthrie, 1997; Swalander and Taube, 2007; Vansteenkiste et al., 2009) and, therefore, read more than males (Scales and Rhee, 2001). A third individual factor related to changes in habits during the pandemic is individual distress. A situation of confinement and pandemic as the one experienced in 2020 and again in 2021 can generate anxiety and psychological distress (Casagrande et al., 2020), factors that can significantly affect our behaviors (Ingram et al., 2020; Stanton et al., 2020), such as reading habits. For instance, the greater psychological distress experienced during the recent coronavirus disease 2019 (COVID-19) pandemic in China was associated, especially among women, with increased time spent daily on social media searching for and reading information on COVID-19 (Guo et al., 2021). To the best of the knowledge of the authors, the three-way interaction between reading habits, intrinsic reading motivation, and distress has not yet been explored. Thus, the main objective of this study is to understand the nature of such interaction, since this could lead to comprehension of how distress and motivational factors are related to changes in reading habits in highly distressful situations and whether this relationship is different for each gender.

Motivation (SDT) and Reading

Schiefele et al. (2012) reviewed the construct of reading motivation, which can be understood as current or habitual reading motivation (cf. Pekrun, 1993). In this research, we focused on the habitual reading motivation which "... denotes the relatively stable readiness of a person to initiate particular reading activities" (Schiefele et al., 2012, p. 429). Regardless of this distinction, Schiefele et al. (2012) also provided a comprehensive review of the dimensionality of the construct of reading motivation. From a psychological perspective, it has been studied from different lines of research, for instance, through the value beliefs of reading tasks (e.g., Durik et al., 2006; Solheim, 2011), from the self-concept and self-efficacy of reading (e.g., Chapman and Tunmer, 1995; Wigfield and Guthrie, 1997; MRQ reading efficacy subscale) or *via* orientations of objectives related to reading (e.g., Graham et al., 2008). However, according to Schiefele et al. (2012), the most important distinction is the one that exists between intrinsic and extrinsic reading motivation. Intrinsic reading motivation is defined as the willingness to read because the activity is satisfying in itself for an individual. In contrast, extrinsic reading motivation refers to when reading is motivated by its expected consequences, either to obtain positive outcomes or to avoid negative ones (Wigfield and Guthrie, 1997; Becker et al., 2010; de Naeghel et al., 2012, 2014; Schiefele et al., 2012, 2016). These types of

reading motivation can be described as higher order categories subsuming several dimensions.

The SDT (Ryan and Deci, 2019) is an exceptional framework for approaching the study of these motivational categories. As an organismic theory, SDT posits that its representation of motivation is universal (Pintrich and Schunk, 2006; Ryan and Deci, 2020). This is justified by considering the existence of three basic psychological needs that are innate to human behavior: autonomy, competence, and relatedness. Autonomy refers to the basic psychological need of feeling as the initiators of our own actions, so it is promoted by experiences of personal interest and value (Ryan and Deci, 2020). Externally controlled experiences, such as rewards and punishments, may undermine autonomy. Individuals also need to feel competent in their interactions with other people, in the tasks and activities they perform, and in adapting to their environment in a broad sense. This need is best satisfied when tasks or activities afford optimal challenges and opportunities for growth. Finally, relatedness refers to the need of being part of a group, so it has commonly been denominated as a need for belonging (Pintrich and Schunk, 2006; Ryan and Deci, 2020). It is promoted by respect and affection. Taking these needs into account, intrinsic motivation can be defined as "the human need of being competent and self-determined in relation with the environment" (Deci, 1980, p. 27). Autonomy and competence needs are, thus, directly related to intrinsic motivation (Garon-Carrier et al., 2016).

From this theoretical perspective, most of the general outcome categories (i.e., achievement, persistence, well-being) treated in motivational research are not exclusive to the education domain. Hence, "...current results in this research field can be probably generalized to domains beyond education such as workplaces, health-related behavior and interpersonal relationships" (Howard et al., 2021, p. 19). For this reason, SDT can serve as a framework for studying adult reading habits.

According to SDT, the types of reading motivation can also be ordered on a continuum of self-determination (Howard et al., 2017; Ryan and Deci, 2020), varying from an absence of self-determination (i.e., motivation), to partially self-determined behaviors (e.g., an extrinsic motivation referred to as introjected regulation), to the most self-determined behaviors (i.e., intrinsic motivation). Thus, intrinsic motives will always be self-determined, whereas extrinsic motives can be categorized as more or less self-determined: external, introjected, identified, and integrated regulation (Ryan and Deci, 2019, 2020; Howard et al., 2021). External regulation concerns behaviors driven by externally imposed rewards and punishments (e.g., reading to avoid punishment). Introjected regulation is characterized by ego-involvement motives, because the goal is to gain and maintain approval from the self and others (Howard et al., 2021), driven by feelings of "I should..." and guilt. Identified and integrated regulations are the most self-determined of these extrinsic motives. Individuals are driven by an identified regulation when their behaviors are consistent with perceived personal values and meaning, regardless of the enjoyment derived from enacting those behaviors. Finally, when driven by an

integrated regulation, individuals assimilate the enactment of a behavior into their sense of self such that the behavior becomes a fully congruent element of their identity. One of the core hypotheses of SDT is that supporting basic psychological needs facilitates the most autonomous types of motivation (i.e., intrinsic motivation and integrated and identified regulation), whereas thwarting those needs undermines it (Ryan and Deci, 2020).

This study focuses on the most self-determined type of motivation: intrinsic reading motivation. This decision is based on the strength of this type of motivation to predict performance and well-being, making it possible to consider it as the most optimal type of motivation (Garon-Carrier et al., 2016). A very recent meta-analysis on SDT by Howard et al. (2021) has reinforced this argument and has provided revealing data linking intrinsic motivation with some individual variables related to well-being (e.g., anxiety and distress), a relationship that we intend to explore in this study.

Intrinsic motivation can be defined as a psychological desire to execute behaviors (reading, in this case) for the only purpose of obtaining the satisfaction, pleasure, or excitement associated with enacting the behavior itself (Ryan and Deci, 2019). In other words, reading is intrinsically motivating as long as it satisfies the psychological needs of competence and autonomy (Ryan and Deci, 2009).

In this study, we expected the confinement to have an impact on how people met their basic needs through their reading habits. We also explored the association between intrinsic reading motivation (IRM) and reading behavior (i.e., reading frequency) and how it varied with the type of reading, gender, and distress in such an exceptionally stressful context. The results obtained can offer valuable information about their relevance beyond an educational context in the adult population.

Intrinsic Reading Motivation, Reading Behavior, Type of Reading, Gender, and Distress

Studies have consistently found a positive relationship between intrinsic reading motivation and reading behavior (Wang and Guthrie, 2004; Unrau and Schlackman, 2006; Law, 2008, 2009; Becker et al., 2010; Retelsdorf et al., 2011). Hence, people who read for their own enjoyment read more often than those who read for other external motivations, such as to avoid a punishment or to obtain a reward. However, this association has been mainly observed in educational contexts. Schiefele et al. (2012) suggested how in educational settings the association between intrinsic motivation and reading frequency is stronger for leisure reading than for academic reading, finding that the higher the intrinsic motivation to read, the higher the frequency of leisure reading. Furthermore, de Naeghel et al. (2012) found that, regardless of the context in which motivation is measured (i.e., recreational versus academic), autonomous motivation is positively related to leisure reading frequency. Studies focused on adult reading motivation are rare. Schutte and Malouff (2007), although did not specifically analyze intrinsic motivation, found that understanding reading as part of the self (reading as part of one's identity or,

according to SDT, self-determined motivation) was strongly related to reading frequency (self-reported measure). This was moderately related to hours spent on leisure reading and not significantly associated to hours spent on mandatory reading.

Previous studies have also found that gender plays an important role in the relationship between reading motivation and frequency and/or type of reading. Studies focused on reading motivation find that girls are more intrinsically motivated to read, and this is positively associated to reading frequency (de Naeghel et al., 2012). Moreover, higher intrinsic motivation toward academic tasks of girls is positively related to their achievement and learning (Ratelle et al., 2007; Vansteenkiste et al., 2009). Also, the studies focused on adult reading (Schutte and Malouff, 2007) have found that women scored significantly higher than men on reading motivation scales and on self-determined motivation subscales. In addition, studies and surveys on reading habits systematically show that women read more than men, and especially read more for leisure (Scales and Rhee, 2001; Mellard et al., 2007; Federación de Gremios de Editores de España (FGEE), 2020) even in confinement (Salmerón et al., 2020; Federación de Gremios de Editores de España (FGEE), 2021).

The results on motivation to read and gender effects seem quite robust. However, in such an exceptional situation as confinement, there might be other factors that could affect the relationship between motivation, gender, and reading habits. Casagrande et al. (2020), for example, have found that the COVID-19 outbreak has had a psychological impact in terms of anxiety symptomatology and psychological distress on a population of 2,291 Italian respondents to a survey. Similarly, Alzueta et al. (2021) analyzed the psychological impact of the pandemic across 59 countries around the world. They found that a significant proportion of respondents reported moderate to severe symptoms of depression (25.4%) and anxiety (19.5%), and that European citizens, women, and young adults were the ones who obtained higher scores. Experiencing psychological distress is significantly associated with changes in health and physical behavior (Ingram et al., 2020; Stanton et al., 2020). Moreover, Guo et al. (2021) suggested an association between distress and reading behaviors. Concretely, they found that the amount of time people spent daily on social media searching for and reading information on COVID-19 is, among other things, a predictor of psychological distress. Also, Alzueta et al. (2021) suggested that younger adults may be more vulnerable to the psychological effects of the COVID-19 pandemic as a consequence of greater exposure to media. However, a systematic analysis of how distress generated by confinement might affect reading habits is, to the best of the knowledge of the authors, still lacking. Such association may also be influenced by motivational factors: we know from previous studies that people who have high intrinsic reading motivation toward learning tasks (i.e., musical or academic content) experience less distress in task performance than those with less self-determined motivation (Stoeber and Eismann, 2007). That is, it seems that being intrinsically motivated toward a learning task protects against the occurrence of distress associated with the performance of that

task. Furthermore, Levine et al. (2020) found that students with autonomous motivation read more books recreationally, which was associated with lower levels of psychological distress. They conclude that recreational reading mitigates the frustration of one's basic psychological needs by improving the feeling of being autonomous, competent, and socially connected. Thus, exploring the association between distress, reading habits, gender, and motivation is important to better understand the role of intrinsic motivation on reading. It is possible that people with high levels of intrinsic motivation toward reading do not vary their reading frequency even though they experience a high degree of distress generated by the confinement situation (Levine et al., 2020). However, according to Guo et al. (2021), the time spent on social media and reading news about COVID is a predictor of distress, so we wonder whether the IRM would play some role in this relation, as reasons of people for reading social media and news might be independent of IRM.

Therefore, one of the challenges we sought to address in this correlational study is to explore whether these relationships between intrinsic reading motivation and reading frequency are maintained in the adult population from prior to confinement, at the start of confinement, and after a month of confinement.

Hypotheses

Based on prior studies, we made the following hypotheses:

H1: IRM will be related to reading frequency (RF). Specifically, it is expected:

H1.1: A positive relationship between IRM and RF, so that those people who report higher IRM will show higher RF than those who report low intrinsic motivation (Schiefele et al., 2012).

H1.2: This positive relationship between IRM and RF will occur at the three points in time investigated: prior to confinement, at the start of confinement, and during the continuation of confinement.

H1.3: According to previous studies (de Naeghel et al., 2012), we did expect to find a positive relationship between IRM and RF in the types of academic or work and leisure reading, a relationship that will be stronger in the latter (de Naeghel et al., 2012, 2014). On the other hand, we did not expect a specific relationship between IRM and social reading and reading news. It can be hypothesized that news and social reading may not be the main means of satisfying the pleasure associated with enacting the reading behavior itself.

H2: IRM and RF will be different for males and females. Differences will be the following:

H2.1: Based on previous studies, we expect females to have higher IRM than males (e.g., Wigfield and Guthrie, 1997; Swalander and Taube, 2007; Vansteenkiste et al., 2009). Therefore, and in line with hypothesis 1.1, females will spend more hours reading (RF).

H2.2: The superiority of women in terms of intrinsic motivation and reading frequency will be constant across the three-time period: before, in the beginning, and few weeks of confinement.

H3: According to previous studies (Casagrande et al., 2020), situations of confinement can generate stress and overloads in individuals. We expected that these stressors will condition to a lesser extent the RF in people with higher IRM (Howard et al., 2021), especially in the type of leisure reading (Levine et al., 2020).

DESIGN

Participants

The sample for this study was formed by 4,181 adults from Spain. From that initial sample, 332 participants were excluded because of one or several of the following reasons: they did not accept the use of their data for the study, they did not currently live in Spain, they showed an incoherent pattern of responses, or they were not in the 18–65 age range (see Salmerón et al., 2020, for more information on the database used for this study). After removing all these cases, the final sample was composed of 3,849 Spanish adults between 18 and 65 years old ($M = 33.52$ years, $SD = 13.93$). As shown in **Table 1**, the final sample was predominantly female, young, and medium-educated.

A non-probabilistic sample was used. A link to the survey was published online and sent to educational institutions, friends, and family members asking to forward it. In order to facilitate the generalization of the results, a weighting adjustment method was performed to account for misrepresented groups. First, the reference values (i.e., frequencies by age and gender), as provided

TABLE 1 | Descriptive statistics of the participants.

N		3849
Gender	Female	2724 (70.8%)
	Male	1125 (29.2%)
Age range	From 18 to 24 years	1620 (42.1%)
	From 25 to 34 years	683 (17.7%)
	From 35 to 44 years	548 (14.2%)
	From 45 to 54 years	567 (14.7%)
	From 55 to 65 years	431 (11.2%)
Occupation	Students	1782 (46.3%)
	Workers	1944 (50.5%)
	Unemployed	222 (5.8%)
	Retired	79 (2.1%)
	Other	28 (0.7%)
Completed studies	Primary education	32 (0.8%)
	Secondary education	1858 (48.3%)
	Undergraduate degree	1200 (31.2%)
	Master degree	534 (13.9%)
	Ph.D. degree	220 (5.7%)
	Other	5 (0.1%)

Percentages in brackets are shown.

by the Spanish Nation Institute of Statistics (Instituto Nacional de Estadística, 2020), were computed. Then, the observed frequencies by age range and gender in the database were computed. The weight value was $W = (n_p/n_s) \times (N_s/N_p)$, where n_p is the frequency by age and gender in the population, n_s is the frequency by age and gender in the sample, N_s is the sample size, and N_p is the population size.

Procedure

All data were extracted from the database READ-COGvid (Salmerón et al., 2020), collecting from April 11, 2020 to April 19, 2020 in Spain.

Data for the READ-COGvid reading habits survey were collected by an unrestricted self-selected survey. It published a link to the survey on social media and sent links to the survey to educational associations, undergraduate students from several universities in Spain, and to social networks of the researchers with a request to spread it. Responses were collected *via* the tool LimeSurvey, and data were stored in the servers of the University of Valencia, following the GDPR Compliance. The study was designed following the ethical principles of the Declaration of Helsinki. Before their participation, the participants were informed about the goals of the study and about the ethical guidelines followed in the design and data treatment.

Measures

Intrinsic Reading Motivation

An adapted version of the SRQ -Reading Motivation Questionnaire (de Naeghel et al., 2012) was used, developed to capture intrinsic reading motivation (e.g., “I read because I enjoy reading”). Four items were selected from the original questionnaire. Items were scored on a five-point Likert scale, ranging from 1 (completely disagree) to 5 (completely agree). The four-item questionnaire had good internal consistency: McDonald's omega ($\omega = 0.89$) and average variance extracted (AVE = 0.65). A confirmatory factor analysis (CFA) was performed to assess the reliability of the instrument more deeply. Because of the ordinal nature of the scale, the diagonally weighted least squares (DWLS) estimation was selected. The factor model was significant ($\chi^2 = 37.879$; $p < 0.001$). Additional fit measures were appropriate: IFI = 0.984, GFI = 0.993, CFI = 0.984, RMSEA = 0.068.

Reading Frequency on Type of Reading

This scale assessed how much daily time the participants recalled spending on different reading activities: reading for leisure, reading for work or study, reading news to keep up with current events, and social reading (Scales and Rhee, 2001; Torppa et al., 2020). For each reading activity, they answered using the following scale: nothing, ~30 min a day, 1, 2, 3, 4 h a day.

Reading Habits Times

The participants fulfilled the previous scale (reading frequency on type of reading) three times: first, they recalled the last

time (before the confinement) they had spent few days at home (e.g., holiday, weekend...), and that report was used for the “Before” measure. Then, they fulfilled the scale again recalling the first 2 weeks of confinement (this was the “Beginning” measure). Finally, for the “Month” measure, they reflected about the current period, after few weeks of confinement had passed. Therefore, the survey was sent just once to each participant, and they fulfilled the reading frequency on the type reading scale three times, just recalling different points in time.

Distress

The personal distress (PD) subscale of the Interpersonal Reactivity Index (IRI) (Davis, 1980) was used to assess the distress of the participants during the lockdown. This is a seven-item five-point Likert scale tapping feelings of anxiety and self-control in tense situations. The respondents rated statements such as “In emergency situations, I feel apprehensive and ill-at-ease” from 1 (does not describe me well) to 5 (describes me very well). Validated adaptation in Spanish (Escrivá et al., 2004) was used. The PD scale showed good internal reliability in the sample ($\omega = 0.82$; AVE = 0.42). Although AVE is a bit low, it is still considered appropriate if composite reliability is higher than 0.6 (Fornell and Larcker, 1981). Composite reliability for the subscale is 0.84. A CFA was performed to assess the reliability of the instrument more deeply. Because of the ordinal nature of the scale, the DWLS estimation was selected. The factor model was significant ($\chi^2 = 548.527$; $p < 0.001$). Additional fit measures were appropriate: IFI = 0.912, GFI = 0.973, CFI = 0.912, RMSEA = 0.1. Although RMSEA is a bit high, this test is not very reliable on its own, and previous studies have suggested that there is no support for universal RMSEA cut-off points, recommending the use of the other goodness of fit measures to make decisions (Chen et al., 2008). Overall, the scale reliability is fair.

Data Analysis

In order to test H1 and H2, we used a linear mixed effects model (LMM) with the lmer and emmeans packages in R (R Core Team, 2020) for each type of reading (leisure vs. news vs. social vs. study/work), which is four in total. The categorical fixed factors were time (before vs. beginning vs. month) and gender (female vs. male), and the continuous fixed factor was IRM. The participants were the random effect in the model. The dependent variable was the number of reading hours per day (reading frequency, RF). As RF was not normally distributed, before performing any statistical analysis, we log transformed it, which resulted in more normal distributions. The rest of the assumptions for the analysis were met, the *F*- and *p*-values for the analyses. For each significant effect, we also report in-text, the β values (standardized coefficients) as a measure of effect size ($\beta < 0.2$ = weak, $\beta > 0.2$ and < 0.5 = moderate, $\beta > 0.5$ = strong effect, Acock, 2014). The effect size was calculated with the function “effectsize: standardize_parameters()” in R.

TABLE 2 | Effects of IRM, time, and gender on RF for each type of reading.

	Study/Work					
	Sum sq	Mean sq	NumDF	DenDF	F value	Pr(>F)
Time	0.08	0.04	2	8305.9	0.37	0.692
IRM	2.64	2.64	1	3991.8	24.18	0.000***
Gender	0.03	0.03	1	4078.4	0.25	0.614
Time:IRM	0.4	0.2	2	8305.9	1.82	0.162
Time:Gender	0.47	0.24	2	8305.9	2.18	0.114
IRM:Gender	0.03	0.03	1	3991.8	0.26	0.607
TIME:IRM:Gender	0.82	0.41	2	8305.9	3.75	0.024*
Leisure						
Time	0.25	0.13	2	8189.6	1.56	0.210
IRM	46.48	46.48	1	3793.5	576.63	<2.2e-16***
Gender	0.04	0.04	1	3886	0.53	0.465
Time:IRM	1.52	0.76	2	8189.6	9.41	0.000***
Time:Gender	0.27	0.13	2	8189.6	1.67	0.189
IRM:Gender	0.06	0.06	1	3793.5	0.76	0.384
TIME:IRM:Gender	0.17	0.09	2	8189.6	1.07	0.342
Social						
Time	0.6	0.1	2	7902	8.91	0.000***
IRM	2.41	2.41	1	3766.2	71.87	<2.2e-16***
Gender	0.79	0.79	1	3822.2	23.49	0.000***
Time:IRM	0.14	0.07	2	7902	2.07	0.126
Time:Gender	0.01	0.01	2	7902	0.21	0.812
IRM:Gender	0.22	0.22	1	3766.2	6.57	0.010*
TIME:IRM:Gender	0.01	0.00	2	7902	0.08	0.924
News						
Time	0.55	0.28	2	8023.8	5.16	0.006**
IRM	0.86	0.86	1	3596.8	15.97	0.000***
Gender	0.03	0.03	1	3689.4	0.57	0.449
Time:IRM	0.08	0.04	2	8023.8	0.75	0.474
Time:Gender	0.04	0.02	2	8023.8	0.38	0.685
IRM:Gender	0.03	0.03	1	3596.8	0.63	0.427
TIME:IRM:Gender	0.02	0.01	2	8023.8	0.16	0.849

Sig. codes. 0***; 0.001**; 0.01*. The data reported in this table have been calculated with transformed reading frequency (RF).

RESULTS

Effect of Intrinsic Reading Motivation, Time of Confinement, and Type of Reading on Reading Frequency

A table with all the effects for each type of reading can be seen in **Table 2**. For clarity purposes, results will be commented for each type of reading next.

Leisure Reading

Table 2 shows that the main effects of time of confinement and gender were not significant. However, the main effect of IRM was significant. As predicted, the higher the IRM, the higher the RF ($\beta = 0.29$). The interaction between time and IRM was also significant: although the relationship between intrinsic motivation and RF was positive at any moment of confinement, it was stronger after a month of the confinement than at the beginning and before ($\chi^2 = 46.477$, $p < 0.001$); that

is, the slope of IRM was higher at the month of confinement ($\beta = 0.06$).

Study

In the case of reading for study/work, only the main effect of IRM was significant ($\beta = 0.06$); that is, the higher the IRM, the higher the RF, so H1.1 was confirmed. The effect of time was only significant in interaction with IRM and gender, so we will analyze it in the next section.

News

In the case of news reading, the main effect of IRM was again significant; that is, again, the higher the IRM, the higher the RF ($\beta = 0.05$), confirming H1.1. The main effect of time was also significant, but in this case, the RF was significantly higher at the beginning of the confinement ($\beta = 0.39$) than before and after a month ($\beta = 0.33$) (see **Figure 1**) regardless of the level of

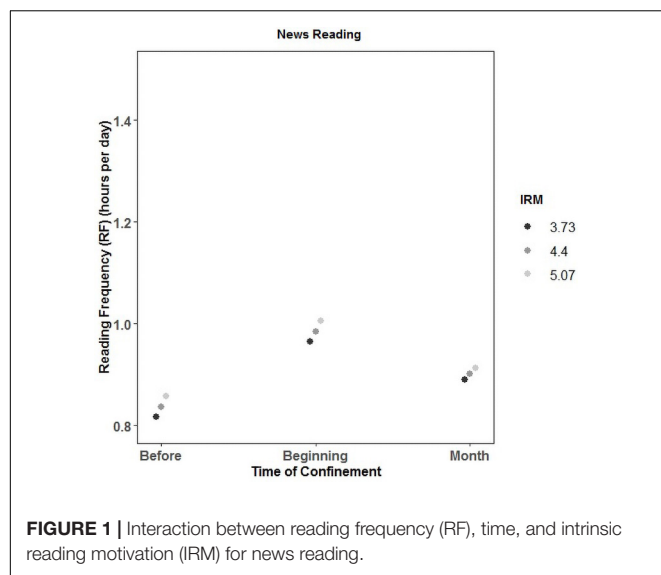


FIGURE 1 | Interaction between reading frequency (RF), time, and intrinsic reading motivation (IRM) for news reading.

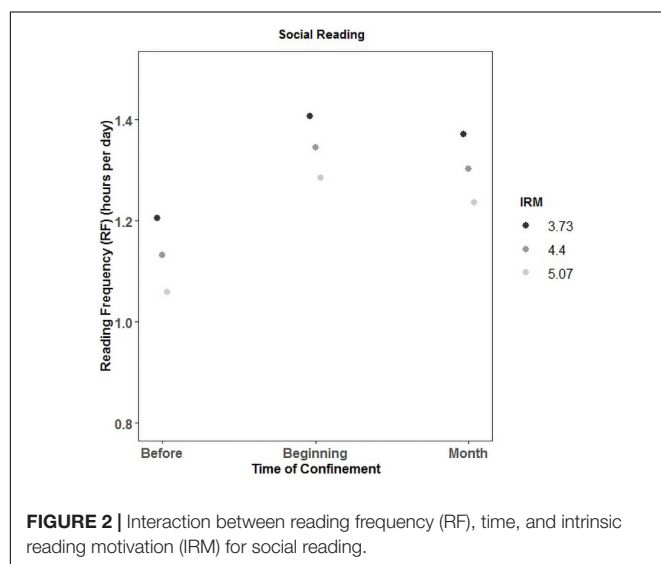


FIGURE 2 | Interaction between reading frequency (RF), time, and intrinsic reading motivation (IRM) for social reading.

intrinsic motivation. The interaction between time and IRM was not significant.

Social

Finally, in the case of social reading, the main effect of both IRM and time was again significant. However, in the case of IRM ($\beta = -0.18$), contrary to the pattern in the rest of types of reading, the higher the IRM, the lower the frequency of social reading (before: $\chi^2 = 70.146$, $p < 0.001$; beginning: $\chi^2 = 47.8$; $p < 0.001$; month: $\chi^2 = 61$, $p < 0.001$). The effect of time followed the same pattern than in the case of news, that is, the RF was significantly higher at the beginning of the confinement ($\beta = 0.33$) than before and after a month ($\beta = 0.33$) (see **Figure 2**) regardless of the level of intrinsic motivation.

TABLE 3 | Means and SE of reading frequency (hours per day) per type of reading, time, and gender (untransformed data).

Type of reading	Time	Gender			
		Female		Male	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Leisure	Before	0.93	0.02	0.94	0.03
	Beginning	1.27	0.02	1.25	0.03
	Month	1.42	0.02	1.30	0.03
News	Before	0.93	0.02	0.94	0.03
	Beginning	1.27	0.02	1.18	0.03
	Month	1.10	0.02	1.03	0.03
Socialize	Before	1.38	0.02	1.10	0.03
	Beginning	1.95	0.02	1.43	0.03
	Month	1.82	0.02	1.38	0.03
Study/Work	Before	1.88	0.02	1.81	0.03
	Beginning	2.07	0.02	2.15	0.03
	Month	2.14	0.02	2.14	0.03

M, mean; *SE*, standard errors.

Effect of Intrinsic Reading Motivation and Gender on Reading Frequency

Descriptives for RF for each type of reading considering gender and each time of confinement are shown in **Table 3**.

The main effect of gender was significant only in the case of social reading, with females showing higher RF of social media than males ($M_{\text{females}} = 1.89$, $SE = 0.02$; $M_{\text{males}} = 1.4$, $SE = 0.03$). Also, for this type of reading, the interaction between IRM and gender was significant, such that the slope of IRM was higher for females than for males ($\chi^2 = 6.574$; $p = 0.01$) (see **Figure 3**).

In the case of study/work reading, the three-way interaction between IRM, gender, and time was significant. As can be seen in **Figure 4**, females showed higher RF for study/work than males before the confinement, especially if they presented higher IRM. However, males only showed higher RF for work/study than females at the beginning of the confinement provided they presented high IRM.

Effect of Distress, Intrinsic Reading Motivation, and Time of Confinement on Reading Frequency by Type of Reading

In order to explore H3, we split the participants into two groups according to the median of IRM (more IRM participants > 4.5 , less IRM participants ≤ 4.5). We then conducted LMM analyses introducing distress and time of confinement as continuous factors, and RF as dependent variable (log-transformed in order to normalized it) per each type of reading and group of IRM (eight LMM analyses in total). Again, the participants were the random effect in each one of the eight models. The results of these analyses are shown in **Table 4**.

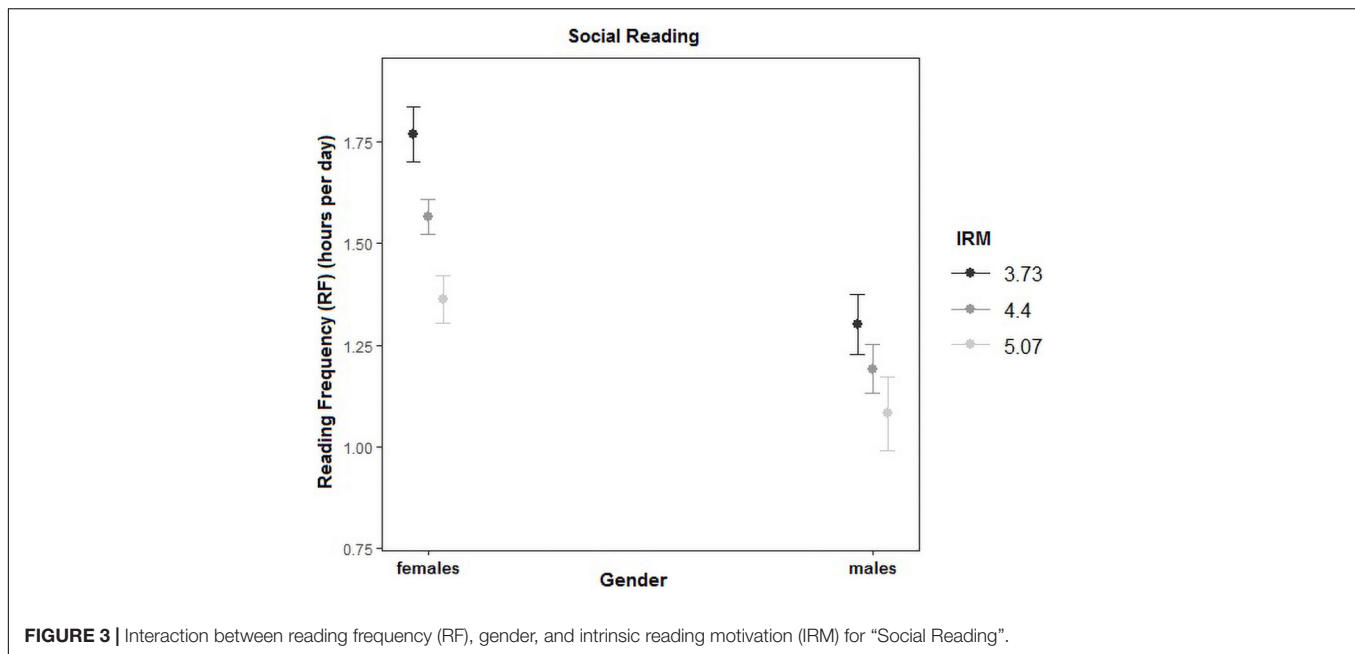


FIGURE 3 | Interaction between reading frequency (RF), gender, and intrinsic reading motivation (IRM) for "Social Reading".

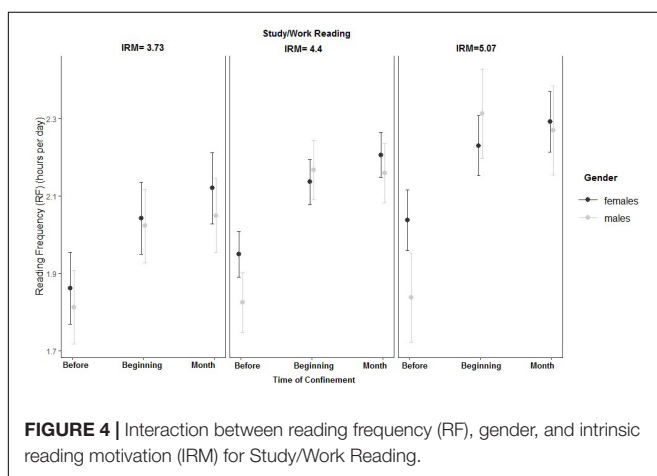


FIGURE 4 | Interaction between reading frequency (RF), gender, and intrinsic reading motivation (IRM) for Study/Work Reading.

Study/Work Reading

For higher IRM participants, the effect of time was again significant, such that RF was significantly lower before of the confinement than at the beginning ($z.ratio = -7.246$; $p < 0.001$; $\beta = -0.264$) and at the month ($z.ratio = -7.205$; $p < 0.001$; $\beta = -2.62$) of confinement. However, the difference in RF was not significant between the beginning and the month of confinement ($z.ratio = 0.04$; $p = 0.999$). The effect of distress was also significant in such a way that the higher the distress, the lower the RF ($\beta = -0.03$).

For lower IRM participants, neither the effect of time nor the effect of distress was significant.

Leisure Reading

For higher IRM participants, the effect of time was significant again. In this type of reading, however, the difference in RF between the beginning and the month of confinement was also

significant ($z.ratio = -6.698$; $p < 0.001$, $\beta = -0.601$), as in the case of before ($z.ratio = -14.548$; $p < 0.001$, $\beta = -0.315$) vs. beginning and before vs. month ($z.ratio = -21.245$; $p < 0.001$, $\beta = -0.845$). However, there was no main effect of distress for the higher IRM participants.

For lower IRM participants, the effect of time was significant again, such that RF was significantly lower before the confinement than at the beginning ($z.ratio = -13.288$; $p < 0.001$, $\beta = -0.46$) and at the month ($z.ratio = -15.226$; $p < 0.0001$; $\beta = -0.527$) of the confinement. However, the difference in RF was not significant between the beginning and the month of the confinement ($z.ratio = -1.938$; $p = 0.128$). As in the case of higher IRM, there was no main effect of distress for lower IRM.

Social Reading

For higher IRM participants, both the main effect of time ($\beta = 0.42$) and distress ($\beta = 0.08$) were significant. As it can be seen in **Figure 5A**, higher IRM participants showed more social RF at the beginning and after a month of the confinement than before and those participants more distressed showed higher RF of social media in general. However, the interaction between distress and time on RF was also significant, so the effect of distress was significant before ($\chi^2 = 10.413$, $p < 0.004$) and at the beginning of the confinement ($\chi^2 = 19.144$, $p < 0.001$) but not after a month ($\chi^2 = 4.596$, $p = 0.096$).

For lower IRF participants (**Figure 5B**), the main effect of time ($\beta = 0.32$) and distress ($\beta = 0.12$) were again significant. With regard to time, social RF was higher at the beginning and at the month of the confinement than before. Again, the higher the distress, the higher the RF for social media. The interaction time \times distress was not significant.

TABLE 4 | Effects of distress and time on RF for each type of reading and higher and lower IRM.

Study/Work						
	Sum sq	Mean sq	NumDF	DenDF	F value	Pr(>F)
Higher IRM						
Time	1.34	0.67	2	3903.1	6.09	0.002**
Distress	0.46	0.46	1	1973	4.17	0.041*
Time:distress	0.14	0.07	2	3903.1	0.65	0.520
Lower IRM						
Time	0.20	0.10	2	4400.7	0.93	0.393
Distress	0.03	0.03	1	2278	0.31	0.575
Time:distress	0.31	0.15	2	4400.7	1.42	0.241
Leisure						
	Sum sq	Mean sq	NumDF	DenDF	F value	Pr(>F)
Higher IRM						
Time	1.86	0.93	2	3896.4	11.99	6.45E-06***
Distress	0.04	0.04	1	1893.5	0.52	0.471
Time:distress	0.34	0.17	2	3896.4	2.21	0.110
Lower IRM						
Time	0.61	0.30	2	4265	3.6	0.0275*
Distress	0.29	0.29	1	2129	3.44	0.064.
Time:distress	0.36	0.18	2	4265	2.14	0.118
Social						
	Sum sq	Mean sq	NumDF	DenDF	F value	Pr(>F)
Higher IRM						
Time	2.49	1.24	2	3698.1	36.12	2.914E-16***
Distress	0.44	0.44	1	1844.5	12.91	0.000***
Time:distress	0.31	0.16	2	3698.1	4.55	0.011*
Lower IRM						
Time	1.59	0.8	2	4164.2	23.91	4.751E-11***
Distress	1.34	1.34	1	2122.4	40.29	2.669E-10***
Time:distress	0.07	0.03	2	4164.2	1.05	0.351
News						
	Sum sq	Mean sq	NumDF	DenDF	F value	Pr(>F)
Higher IRM						
Time	0.36	0.18	2	3771.7	3.37	0.034*
Distress	0.06	0.06	1	1789.7	1.18	0.278
Time:distress	0.54	0.27	2	3771.7	5.08	0.006**
Lower IRM						
Time	0.29	0.15	2	4255	2.73	0.065.
Distress	0.07	0.07	1	2052.2	1.27	0.260
Time:distress	0.22	0.11	2	4255	2.06	0.128

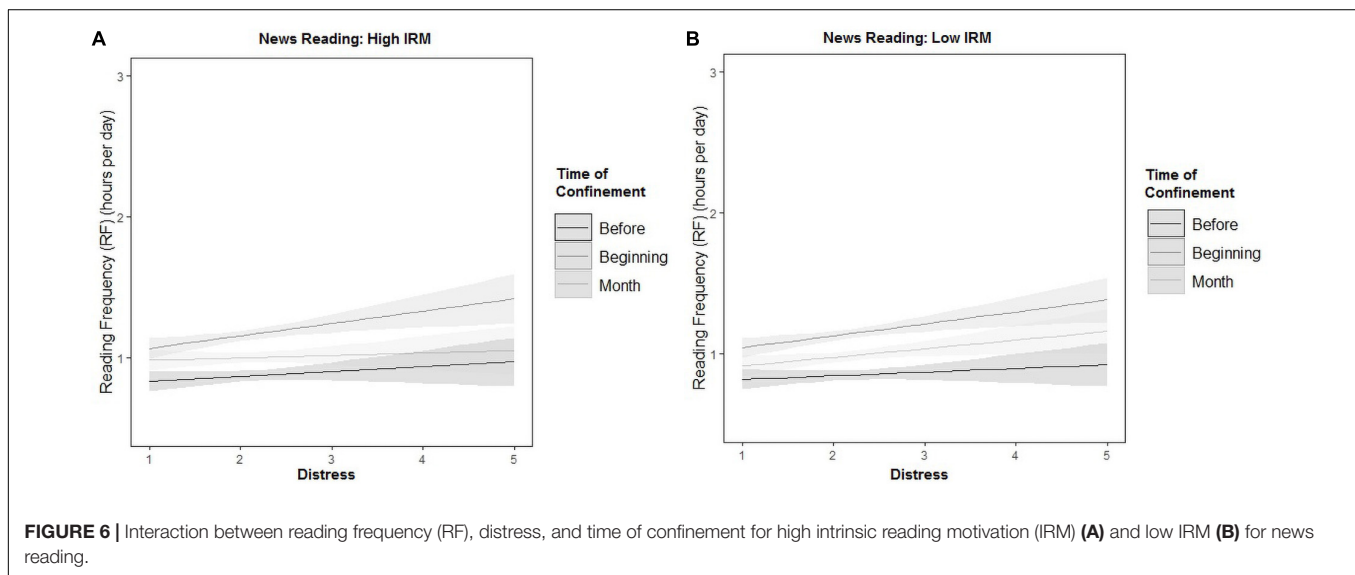
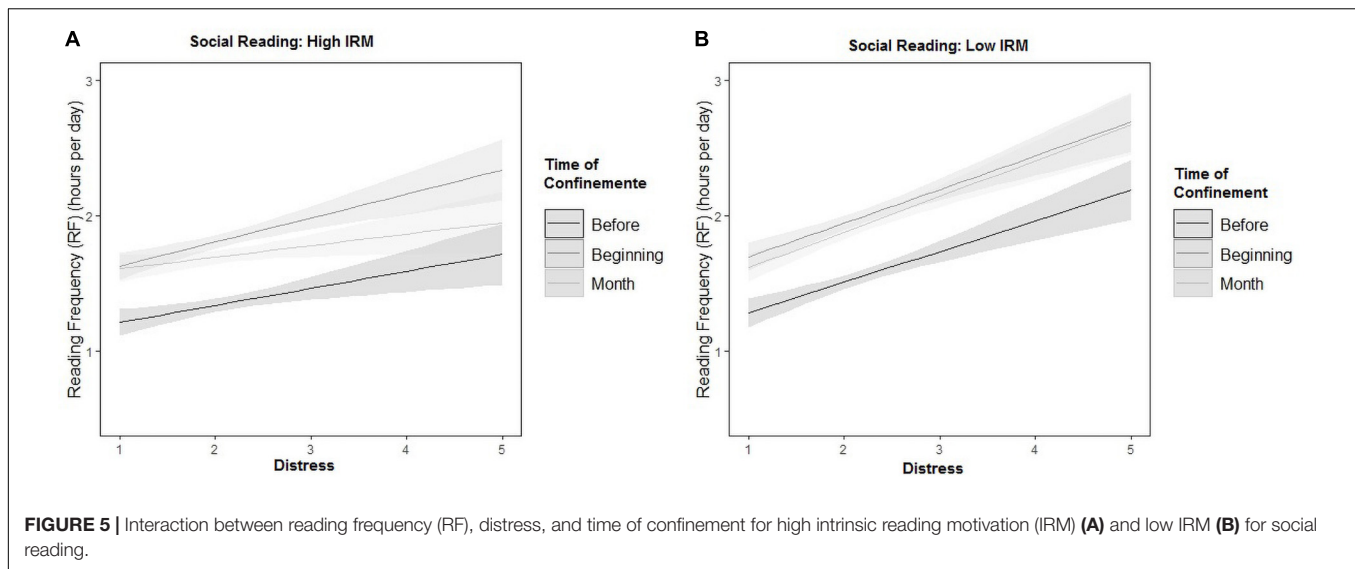
Sig. codes. 0***; 0.001**; 0.01*. The data reported in this table have been calculated with transformed reading frequency (RF).

News Reading

In the case of higher IRM participants, the effect main effect of Time ($\beta = 0.18$) was significant but not the effect of distress. However, distress and time interacted significantly, so the effect of distress was significant at the beginning of the confinement ($\chi^2 = 6.51$, $p < 0.01$) but not before ($\chi^2 = 10.358$,

$p > 0.1$) and after a month ($\chi^2 = 0.183$, $p > 0.1$) (see **Figure 6A**).

For lower IRM participants, neither the effect of distress nor the effect of time on RF was significant (see **Figure 6B**). The interaction Distress \times Time was also not significant.



DISCUSSION

In this research, we used a correlational design to examine the relationship between IRM (as defined within SDT) and reading habits in an adult population taking into account types of reading (i.e., for leisure, for work or study, social networks, and news), gender, and distress generated by the exceptional situation of confinement in Spain caused by the COVID-19 pandemic. In line with theoretical expectations and previous studies, IRM was significantly associated with higher RF. This relationship was stronger for females and for leisure reading, and it was maintained despite the increased distress associated with confinement. Despite the exceptional nature of the pandemic and lockdown situation, this type of motivation seems to support the hypothesis of universal effects of intrinsic motivation, which are independent of contingencies (Ryan and Deci, 2019, 2020).

With regard to the first hypotheses, the main effect of IRM was significant, showing, as expected, that participants with higher IRM invested more hours reading (RF) than participants with lower IRM (Schiefele et al., 2012). This relationship held true at all time points: before, at the beginning, and after a month of confinement for leisure reading. The stronger relationships between IRM and RF were obtained for leisure, work, and social reading, although RF for each type of reading was different: more time invested for work reading, followed by social, and finally news and leisure at a similar level. Also, the relationship between IRM and RF was stronger for leisure than for work reading. This may mean that these types of reading could better satisfy competence and autonomy needs, that is why they are considered intrinsically motivating for the people who enjoy them (de Naeghel et al., 2012). Specifically, leisure reading seems to fit perfectly with SDT assumptions. People can initiate autonomously their reading behaviors even in more

formal or compulsory reading contexts like study/work reading, meeting their needs from an intrinsic motivation. It is not totally surprising that those basic needs most positively related to IRM (autonomy and competence) are better satisfied by leisure and work/study reading. However, this effect was the weakest for news reading, at least in the confinement situation analyzed. The fact that news reading was not influenced that much by IRM could imply that different motives might be in play, for example, reading news may satisfy the person just for the feeling of being informed, not for reading itself.

Contrary to the expectations of the authors, the relationship between social reading and IRM was negative. It is possible that social reading (i.e., social media such as Instagram) facilitates a feeling of belonging in some situations (Pintrich and Schunk, 2006), but it may not be the most appropriate method of satisfying that need. In fact, all the participants reduced their social reading after a month of confinement. For individuals enjoying reading in itself, it seems logical to devote their reading time to leisure reading (which is the one with the strongest positive relation with IRM in this study), because leisure reading might be the most appropriate type of reading to satisfy the motivation of just reading, with no other purposes (like being informed or to connect with other people). Since time to read is somewhat limited in any given day, the more time a person spends in leisure reading, the less time s/he has left for other types of reading, which may explain why the higher IRM (implying longer leisure reading time), the shorter time devoted to social reading. However, this is just a speculative hypothesis that needs to be confirmed or falsified in future studies. Also, SDT recognizes that most intentional behaviors can be multiply motivated (Litalien et al., 2017). In this sense, people who did social reading may have been simultaneously motivated by several types of regulations at the same time (Ryan and Deci, 2020), for example, intrinsic and identified. Social reading may satisfy less self-determined needs, or it may satisfy those basic needs in a different way than leisure reading.

Regarding the second hypothesis, although reading frequencies were higher for females in general, the only significant main effect of type of reading and gender was for social reading. Females were more intrinsically motivated to read than men, as also demonstrated by other studies (Ratelle et al., 2007; Vansteenkiste et al., 2009). In this regard, previous research has suggested the possibility that women with higher autonomous motivation could be more efficient using their time and environment to focus on their studies (e.g., Vallerand et al., 1997). This can explain why women showed higher RF for study/work than men before the confinement. However, males showed higher RF for work/study than females at the beginning of the confinement provided they presented high IRM. This is very interesting, since it suggests that gender differences in reading frequency may be influenced by the context (like a confinement situation).

Finally, intrinsic reading motivation seems to play a different role in reading behavior when the interaction with distress and type of reading is included. Thus, when individuals have high intrinsic motivation, this type of motivation seems to protect reading behavior only in the case of leisure reading. The results

reveal in this sense that it is necessary to consider the interaction between distress and the content of the different readings in order to understand the changes in reading behavior despite high intrinsic motivation. Thus, in the case of study/work reading, distress seems to affect the decrease in reading frequency even if individuals have high intrinsic motivation. In this case, it is possible to think that the interaction between distress and the content of these readings somehow overrides the weight of intrinsic motivation as far as reading behavior is concerned.

In social and news reading, people with high intrinsic motivation read more when distress came into play. In the case of social reading, this interaction can be explained by the potential that social media can become addictive in stressful situations like the COVID pandemic (Zhao and Zhou, 2021). With regard to news reading, at the same level of motivation and stress, timing also seems to have an impact on reading behavior, especially understandable in cases like the one experienced during the past confinement. Thus, it is possible that, at the beginning of the confinement, people were more motivated to read the news because of the need to perceive some control or autonomy in a context in which they were beginning to lack it. This effect ceases to be observed as the confinement progresses.

Therefore, distress seems to pose a differential impact depending on the type of reading: when the reading activity does not pose a demand or requirement (e.g., social or study/work reading) and is not related to stressful stimuli (e.g., news reading), intrinsic reading motivation is a protector against distress and facilitates the satisfaction of personal basic needs. However, when the type of reading is more demanding, distress has a negative impact, and paradoxical events may happen, such as the results obtained showing how the higher IRM participants for study/work reading devote less time to that type of reading if they are distressed, probably because the association of work/study with distress does not allow to satisfy the basic needs associated with that type of reading.

These results support the argument of the need for studying these variables beyond an educational context (Howard et al., 2021) and a young population. Reading motivation should be very present in educational research, but it is now undeniable that we should expand the scope to adult populations in other contexts, since there might be important effects for distress and well-being that may help different interventions or screening strategies. This also suggests that developing intrinsic motivation to read since early childhood may be important not only to promote reading *per se* and for the well-known positive impact on learning and performance but also for the effects beyond the educational context related to the health and well-being of individuals.

Limitations and Further Research

The non-probabilistic sampling procedure and the correlational design are the main limitations of this study. The great majority of the readers involved in this sample are young adults with a medium to high education level. The findings of this study, thus, cannot be generalized to adult readers in general and do not allow us to establish causal relations between variables. It would be interesting to investigate how this stressful environment

affected the reading habits of children, too, or if those habits were influenced or mediated by the behaviors of parents. Also, educational level is a critical factor, which did not receive sufficient attention in this study. In addition, socio-economic status might be a critical variable modulating the effects of motivation to read on reading habits.

Because of the limitations of the confinement, all measures collected in this, as well as in other studies on the recent pandemic, are based on self-reports. Objective reading measures could yield different results and, in general, could be useful to validate the findings of this study. Given the exceptional circumstances of this pandemic, replicating these findings in other studies may be difficult. However, further studies can be conducted in situations in which individuals experience a temporary condition of distress. Exploring how intrinsic reading motivation relates to non-academic behaviors, such as well-being, may lead to interesting findings. It would be desirable to perform more controlled studies exploring all motivational dimensions and how each of them (and their combinations) relates to different types of reading.

CONCLUSION

Intrinsic motivation behaviors (as such, self-determined) have a positive impact during mandatory lockdown periods. It is important to highlight that this confinement was known to be only temporary, but there was also uncertainty at some points about the specific duration. In this sense, people who showed higher IRM managed to protect their reading habits regardless of a stressful context.

Intrinsic reading motivation seems to maintain its positive impact even in extreme situations, such as the global pandemic. However, not every type of reading should be used to satisfy the basic needs associated with IRM. As this study shows, in some cases, a high IRM is related to a higher RF but also to higher distress. When reading involves demands or requirements from the reader and it is performed in a stressful environment like the recent confinement, the relationship between IRM and RF may not be necessarily positive, since increased RF can be associated with higher distress and, therefore, with lower wellbeing. This has been the case for social reading.

REFERENCES

- Acocck, A. C. (2014). *A Gentle Introduction to Stata*, 4th Edn, College Station, TX: Stata Press.
- Alzueta, E., Perrin, P., Baker, F. C., Caffarra, S., Ramos-Usuga, D., Yuksel, D., et al. (2021). How the COVID-19 pandemic has changed our lives: a study of psychological correlates across 59 countries. *J. Clin. Psychol.* 77, 556–570. doi: 10.1002/jclp.23082
- Balluerka, N., Gómez, J., Hidalgo, M. D., Gorostiaga, A., Espada, J. P., Padilla, J. L., et al. (2021). *Las Consecuencias Psicológicas de la covid-19 y el Confinamiento Universidad del País Vasco*. Leioa: Universidad del País Vasco.
- Becker, M., McElvany, N., and Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: a longitudinal study. *J. Educ. Psychol.* 102, 773–785. doi: 10.1037/a0020084
- Casagrande, M., Favieri, F., Tambelli, R., and Forte, G. (2020). The enemy who sealed the world: effects quarantine due to the COVID-19 on sleep quality,

These results contribute to the field and to those people interested in motivation and its relationship with the adaptive behaviors of individuals. Results also support previous findings of the importance of consciously promoting this type of motivation from an early age not only because of the benefits for learning and academic performance but also for the impact beyond the educational context on the management of difficult or stressful situations of individuals.

DATA AVAILABILITY STATEMENT

Publicly Available datasets were analyzed in this study. This data can be found here: https://osf.io/24et3/?view_only=68613c73dd71499bbdadbad93d4ca79a.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RD, IF, and AM implemented the study. IF conducted the statistical analyses. RD, AJ, and AM drafted the manuscript. MR, MG-S, and FN formatted the study. RD, AJ, AM, BA, and JR reviewed the study and performed substantial suggestions. All the authors participated in the conception of the study, reviewed the manuscript critically for relevant intellectual content, and approved the submitted version.

ACKNOWLEDGMENTS

We thank all respondents of the survey, and the people and institutions who helped to distribute the survey.

- anxiety, and psychological distress in the Italian population. *Sleep Med.* 20, 12–20. doi: 10.1016/j.sleep.2020.05.011
- Chapman, J. W., and Tunmer, W. E. (1995). Development of young children's reading self-concepts: an examination of emerging subcomponents and their relationship with reading achievement. *J. Educ. Psychol.* 87, 154–167. doi: 10.1037/0022-0663.87.1.154
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., and Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol. Methods Res.* 36, 462–494. doi: 10.1177/0049124108314720
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catal. Select. Documents Psychol.* 10, 1–17.
- de Naeghel, J., van Keer, H., and Vanderlinde, R. (2014). Strategies for promoting autonomous reading motivation: a multiple case study research in primary education. *Frontline Lear. Res.* 2:83–101. doi: 10.14786/flr.v2.i1.84

- de Naeghel, J., van Keer, H., Vansteenkiste, M., and Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: a self-determination theory perspective. *J. Educ. Psychol.* 104, 1006–1021. doi: 10.1037/a0027800
- Deci, E. L. (1980). *The Psychology of Self-Determination*. Lanham, MD: Lexington books.
- Durik, A. M., Vida, M., and Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: a developmental analysis. *J. Educ. Psychol.* 98, 382–393. doi: 10.1037/0022-0663.98.2.382
- Escrivá, V. M., Navarro, M. D. F., and García, P. S. (2004). La medida de la empatía: análisis del interpersonal reactivity index. *Psicothema* 16, 255–260.
- Federación de Gremios de Editores de España (FGEE) (2020). *Hábitos de Lectura y Compra de Libros en España 2019*. Available online at: <https://www.gremieditores.cat/wp-content/uploads/2020/02/200221PRE-FGEE-Ha%CC%81bitos-lectura-presentacio%CC%81n.pdf> (accessed June 21, 2021).
- Federación de Gremios de Editores de España (FGEE) (2021). *Hábitos de Lectura y Compra de Libros en España 2020*. Available online at: <https://www.editoresmadrid.org/wp-content/uploads/2021/02/Barometro-de-Habitos-de-Lectura-2020-FGEE.pdf> (accessed June 21, 2021).
- Fornell, C., and Larcker, C. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.* 18, 382–388.
- Garces-Bacsal, R. M., and Yeo, S. (2017). Why and what they read when they Don't have to: factors influencing the recreational reading habits of gifted students in Singapore. *J. Educ. Gifte* 40, 1–19. doi: 10.1177/0162353217717035
- Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J. P., et al. (2016).). Intrinsic motivation and achievement in mathematics in elementary school: a longitudinal investigation of their association. *Child Dev.* 87, 165–175. doi: 10.1111/cdev.12458
- Graham, J., Tisher, R., Ainley, M., and Kennedy, G. (2008). Staying with the text: the contribution of gender, achievement orientations, and interest to students' performance on a literacy task. *Educ. Psychol.* 28, 757–776. doi: 10.1080/01443410802260988
- Guo, Y. Q., Sims, O. T., Qin, W. D., and Yang, F. (2021). Factors associated with symptoms of depression and psychological distress during the COVID-19 pandemic. *Behav. Sci.* 11:13. doi: 10.3390/bs11020013
- Howard, J. L., Gagné, M., and Bureau, J. S. (2017). Testing a continuum structure of self determined motivation: a meta-analysis. *Psychol. Bull.* 143, 1346–1377. doi: 10.1037/bul0000125
- Howard, J. L., Bureau, J. S., Guay, F., Chong, J. X. Y., and Ryan, R. M. (2021). Student motivation and associated outcomes: a meta-analysis from self-determination theory. *Perspect. Psychol. Sci.* doi: 10.1177/1745691620966789
- Ingram, J., Maciejewski, G., and Hand, C. J. (2020). Changes in diet, sleep, and physical activity are associated with differences in negative mood during COVID-19 Lockdown. *Front. Psychol.* 11:2328. doi: 10.3389/fpsyg.2020.588604
- Instituto Nacional de Estadística (2020). *Población de 16 y más Años por Nivel de Formación Alcanzado, Sexo y Grupo de Edad [Dataset]*. Available online at: <https://www.ine.es/jaxiT3/Tabla.htm?t=6347> (accessed June 21, 2021).
- Law, Y. K. (2008). The relationship between extrinsic motivation, home literacy, classroom instructional practices, and reading proficiency in second-grade chinese children. *Res. Educ.* 80, 37–51. doi: 10.7227/RIE.80.4
- Law, Y. K. (2009). The role of attribution beliefs, motivation and strategy use in Chinese fifth-graders' reading comprehension. *Educ. Res.* 51, 77–95. doi: 10.1080/00131880802704764
- Levine, S. L., Holding, A. C., Milyavskaya, M., Powers, T. A., and Koestner, R. (2020). Collaborative autonomy: the dynamic relations between personal goal autonomy and perceived autonomy support in emerging adulthood results in positive affect and goal progress. *Motivat. Sci.* doi: 10.1037/mot0000209
- Linnenbrink-García, L., and Patall, E. A. (2016). "Motivation," in *Encyclopedia of Personality and Individual Differences*, eds V. Zeigler-Hill and T. K. Shackelford (Berlin: Springer International Publishing), 91–103.
- Litalien, D., Morin, A. J. S., Gagné, M., Vallerand, R. J., Losier, G. F., and Ryan, R. M. (2017). Evidence of a continuum structure of academic self-determination: a two-study test using a bifactor-ESEM representation of academic motivation. *Contemp. Educ. Psychol.* 51, 67–82. doi: 10.1016/j.cedpsych.2017.06.010
- Mellard, D. F., Patterson, M. B., and Prewett, S. (2007). Reading practices among adult education participants. *Read. Res. Q.* 42, 188–213. doi: 10.1598/RRQ.42.2
- Pekrun, R. (1993). "Facets of adolescents' academic motivation: a longitudinal expectancy-value approach," in *Advances in Motivation and Achievement*, Vol. 8, eds M. L. Maehr and P. R. Pintrich (Greenwich, CT: JAI Press), 139–189.
- Pérez-Rodrigo, C., Citores, M. G., Bárbara, G. H., Litago, F. R., Sáenz, L. C., Aranceta Bartrina, J., et al. (2020). Changes in eating habits during lockdown period due to the COVID-19 pandemic in Spain. *Rev. Española Nutr. Comunitaria* 26, 101–111. doi: 10.14642/RENC.2020.26.2.5213
- Pintrich, P. R., and Schunk, D. H. (2006). *Motivación en Contextos Educativos*. London: Pearson.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., and Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: a person-oriented analysis. *J. Educ. Psychol.* 99, 734–746. doi: 10.1037/0022-0663.99.4.734
- Retelsdorf, J., Köller, O., and Möller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learn. Instr.* 21, 550–559.
- Ryan, R. M., and Deci, E. L. (2009). "Promoting self-determined school engagement: motivation, learning, and well-being," in *Educational Psychology Handbook Series. Handbook of Motivation at School*, eds K. R. Wenzel and A. Wigfield (London: Routledge/Taylor & Francis Group), 171–195.
- Ryan, R. M., and Deci, E. L. (2019). "Brick by brick: the origins, development, and future of self-determination theory," in *Advances in Motivation Science*, Vol. 6, ed. A. J. Elliot (Amsterdam: Elsevier), 111–156. doi: 10.1016/bs.adms.2019.01.001
- Ryan, R. M., and Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp. Educ. Psychol.* 61:101860. doi: 10.1016/j.cedpsych.2020.101860
- Salmerón, L., Arfé, B., Ávila, V., Cerdán, R., De Sixte, R., Delgado, P., et al. (2020). READ-COVID: a database from reading and media habits during COVID-19 confinement in Spain and Italy. *Front. Psychol.* 11:575241. doi: 10.3389/fpsyg.2020.575241
- Scales, A. M., and Rhee, O. (2001). Adult reading habits and patterns. *Read. Psychol.* 22, 175–203. doi: 10.1080/027027101753170610
- Schiefele, U., Schaffner, E., Moller, J., and Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behaviour and competence. *Read. Res. Q.* 47, 427–463. doi: 10.1002/RRQ.030
- Schiefele, U., Stutz, F., and Schaffner, E. (2016). Longitudinal relations between reading motivation and reading comprehension in the early elementary grades. *Learn. Individ. Differ.* 51, 49–58. doi: 10.1016/j.lindif.2016.08.031
- Schutte, N. S., and Malouff, J. M. (2007). Dimensions of reading motivation: development of an adult reading motivation scale. *Read. Psychol.* 28, 469–489. doi: 10.1080/02702710701568991
- Solheim, O. J. (2011). The impact of reading self-efficacy and task value on reading comprehension scores in different item formats. *Read. Psychol.* 32, 1–27. doi: 10.1080/02702710903256601
- Stanton, R., To, Q. G., Khalesi, S., Williams, S. L., Alley, S., Thwaite, T. L., et al. (2020). Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. *Int. J. Environ. Res. Public Health* 17:4065. doi: 10.3390/ijerph17114065
- Stoeber, J., and Eismann, U. (2007). Perfectionism in young musicians: relations with motivation, effort, achievement, and distress. *Personal. Individ. Differ.* 43, 2182–2192. doi: 10.1016/j.paid.2007.06.036
- Swalander, L., and Taube, K. (2007). Influences of family-based prerequisites, reading attitude, and self-regulation on reading ability. *Contemp. Educ. Psychol.* 32, 206–230. doi: 10.1016/j.cedpsych.2006.01.002
- Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M. K., Tolvanen, A., and Poikkeus, A. M. (2020). Leisure reading (but not any kind) and reading comprehension support each other—a longitudinal study across grades 1 and 9. *Child Dev.* 91, 876–900. doi: 10.1111/cdev.13241
- Unrau, N., and Schlackman, J. (2006). Motivation and its relationship with reading achievement in an urban middle school. *J. Educ. Res.* 100, 81–101.
- Vallerand, R. J., Fortier, M. S., and Guay, F. (1997). Self-determination and persistence in a real-life setting: toward a motivational model of high school

- drop up. *J. Personal. Soc. Psychol.* 72, 1161–1176. doi: 10.1037/0022-3514.72.5.1161
- Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., and Lens, W. (2009). Motivational profiles from a self-determination perspective: the quality of motivation matters. *J. Educ. Psychol.* 101, 671–688. doi: 10.1037/a0015083
- Wang, J. H. Y., and Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Read. Res. Q.* 39, 162–186. doi: 10.1598/RRQ.39.2.2
- Wigfield, A., and Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *J. Educ. Psychol.* 89, 420–432. doi: 10.1037/0022-0663.89.3.420
- Zhao, N., and Zhou, G. (2021). COVID-19 stress and addictive social media use (SMU): mediating role of active use and social media flow. *Front. Psychiatry* 12:635546. doi: 10.3389/fpsyt.2021.635546
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 De Sixte, Fajardo, Mañá, Jáñez, Ramos, García-Serrano, Natalizi, Arfé and Rosales. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Comparing the Effects of Stroke-Appearing and Stroke-Disappearing on Learning the Order of Strokes in Chinese Characters

Jon-Chao Hong¹, Kai-Hsin Tai^{1*}, Ming-Yueh Hwang¹ and Pei-Hsin Lin²

¹ Chinese Language and Technology Center, National Taiwan Normal University, Taipei City, Taiwan, ² Department of Business Administration, National Taiwan University, Taipei City, Taiwan

OPEN ACCESS

Edited by:

Pedro García Guirao,
University of Ostrava, Czechia

Reviewed by:

Ariel Laurencio,
University of Sassari, Italy
Yingyi Liu,
Manchester Metropolitan University,
United Kingdom

*Correspondence:

Kai-Hsin Tai
star99xin@gmail.com

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 03 May 2021

Accepted: 02 July 2021

Published: 17 August 2021

Citation:

Hong J-C, Tai K-H, Hwang M-Y and
Lin P-H (2021) Comparing the Effects
of Stroke-Appearing and
Stroke-Disappearing on Learning the
Order of Strokes in Chinese
Characters.
Front. Psychol. 12:704457.
doi: 10.3389/fpsyg.2021.704457

Different approaches to stimulating perceptions in learning can be easily designed with technology-enhanced learning systems. This study aimed to explore how different approaches can influence learners' perceptions that may negatively or positively affect their learning performance of writing Chinese characters using the correct Chinese order of strokes (COS). We therefore designed an e-learning system which was subdivided into two modes: stroke-appearing (i.e., using red to mark incorrect strokes) and stroke-disappearing (i.e., using blanks to mark incorrect strokes) to indicate strokes written in the incorrect order. We then investigated the modes that would facilitate a higher level of attention and better learning outcomes. A total of 10 third-grade elementary school students participated in the experiment, divided into two test groups. Their EEG data were collected, and time series analysis and *t*-tests were utilized to analyze the differences. The results indicated that: (1) there was a significant difference in the attention levels of the students practicing with the stroke-appearing and stroke-disappearing modes when learning COS, and (2) there was a significant difference in the learning outcomes of the students practicing with the stroke-appearing and stroke-disappearing modes when learning COS. These findings support the specific role of stroke order knowledge in learning Chinese characters and the need for the design of an effective method for teaching children to learn Chinese characters.

Keywords: Chinese learning attention, Chinese order of strokes, human-computer interface, learning performance, visual perception

INTRODUCTION

There are many different modes of practice which can contribute to performance gains (Ossmy and Mukamel, 2018), and there is ongoing scientific research on what actually constitutes optimal practice. Researchers have argued that action perception stimulates activity in motor pathways, while also modifying behavior and facilitating learning. It has been found that observing an action leads to the observer activating the corresponding motor plan (Heyes, 2011; Cook et al., 2014). This motor mirror system phenomenon plays a critical role in the promotion of imitation learning and action understanding (Rizzolatti and Craighero, 2004; Bastiaansen et al., 2009; Cook et al., 2014).

Writing exercises have been found to help improve Chinese memorization of Chinese characters (Hsiung et al., 2017), and hastening writing tasks that provide accurate Chinese character writing (Jaganathan and Lee, 2014). By characterizing sensory-evoked activity in the stroke order when writing Chinese characters, we designed two modes of sensory-evoked activities for participants to practice Chinese order of strokes (COS), namely stroke-appearance (i.e., marking the forgotten strokes in red) and stroke-disappearance (i.e., marking the forgotten strokes with blanks) to explore how two modes of action perception would implicitly modify memory. We could then highlight the effects of perception on Chinese stroke acquisition.

Heyes (2001) proposed the associative sequence learning (ASL) theory, according to which the perception-action links which mediate the mirror system arise mainly from the experience of observing and executing the same actions (Cook et al., 2014). Consequently, the mirror system is directly affected by visuo-motor training with incompatible mappings, allowing learners to acquire counter-mirror properties (Heyes et al., 2005; Catmur et al., 2007). In line with this, the mechanisms underlying the modulation effects of COS have rarely been examined with respect to learning by imitation that directs visuo-motor matching learning. Thus, for this study we designed a COS system with two modes (i.e., stroke-appearing and stroke-disappearing) for third-grade students to practice writing Chinese characters. The parameters of the system were designed to enhance students' attention and learning outcomes.

Stroke order is regarded as one of the efficient ways to recognize Chinese characters. There are two main methods of recognizing Chinese characters, namely stroke order sequence and the definition of between-primitive distance measurement (Tung and Jean, 2018). Writing production processing which involves writing execution may evoke the orthographic process of reading through a kinesthetic gesture orthographic code system or the connection between visual orthography and writing motion (Yin et al., in press). The writing process without visual feedback can modulate orthographic processing when reading Chinese characters (Yin et al., in press). It has been proposed that the processes of learning Chinese might establish a motor gesture decoding system for recognizing Chinese characters (Yin and Zhang, 2021). An empirical study revealed that writing exercises help foreign language learners learn Chinese characters, but stroke order learning may not significantly improve the recognition of Chinese characters (Hsiung et al., 2017). This result may be explained by the stroke-number effect, which proposed that the absence of a stroke-number effect may reflect a parallel or holistic strategy (Jiang et al., 2020). However, there are no consistent results showing that stroke-number effect may affect the recognition of Chinese characters (Jiang et al., 2020). A meta-analysis study compared the learning outcomes of typing and handwriting in Chinese, the results of which revealed that handwriting had positive effects on Chinese learners' orthography recognition and orthography-semantic mapping at both the character and lexical levels (Lyu et al., 2021). The most conventional way to learn Chinese characters is to provide cross-grid lines and to ask learners to trace the strokes and characters. The interactive function of tracing fade-out strokes could be a scaffold for Chinese character writing (Xu

et al., 2020). Learners' attention seems to be a critical factor that affects their learning outcomes. A previous study proposed that sustained attention and inhibition does not significantly support the relation to task performance (Guo et al., 2021). However, few studies have made a comparison of two approaches of attracting attention; therefore, the present study utilized EEG to explore the difference in the attention of the two test groups.

THEORETICAL BACKGROUND

Attention

Comparable with goal-directed attention, orienting from memory is dependent on internal representations; these representations can, however, guide attention reflexively without volitional control (Hutchinson and Turk-Browne, 2012). Central vision will process any objects for perceptual recognition, or as targets for action. This makes higher acuity information regarding that object available for any behavioral purposes. According to the above theory, spatial attention is instantiated within the motor system (Similä and McIntosh, 2015). Guitart et al. (2019) also pointed out that visual attention is an important factor of the understanding process. Therefore, the planning of a goal-directed action is both necessary and sufficient to result in a shift of visual attention to those cues. A range of practical strategies have been developed by research psychologists with the aim of improving performers' concentration skills (Greenlees and Moran, 2003). These strategies aim to help performers achieve a focused state of mind in which there is no difference between what they are thinking about and what they are doing (Kremer and Moran, 2012). This raises the question of whether the strategies of stroke-appearing and stroke-disappearing depend on shared or separate selective attention mechanisms. This study aimed to answer this question.

Learning Outcomes

The "focus of attention" mechanism enables the use of cue stimuli to improve readiness and engage motor preparation processes (Handy et al., 2003). The research result of Lin et al. (2021) also verify that attention plays an important role of Chinese character recognition. Yantis and Serences (2003) suggested that the appearance of a new object might have the advantage of automatically triggering selective attention. Learners' visual attention will track by what they have seen (Wang et al., 2019). However, other studies (e.g., Pratt and McAuliffe, 2001) have proposed that selective processing can be triggered by any salient transient, including the disappearance of visual features. Since reflexive mechanisms of attention could be triggered by both the appearance and disappearance of objects (Hopfinger and Maxwell, 2005), a question that arises is how the visually guided cues in the COS system with the stroke-disappearing and stroke-appearing modes could affect learning effectiveness. Further, we aimed to explore which design modes would significantly improve the students' learning outcomes.

RESEARCH HYPOTHESES

Although e-learning has been shown to have many advantages for students' learning, students may still lose attention and focus.

Such negative effects on students' engagement in e-learning could well be the result of their needs for multiple modes of support (McCombs and Vakili, 2005). They may also be due to the failure to represent learning content with effective design strategies (Botturi et al., 2006; Hwang et al., 2008; Li et al., 2021). Hedges et al. (2013) stated that timing is important for play activities and for the development of attention and learning, and that almost everything has a temporal component; in particular, neuronal activity changes over time. Based on the above studies, we considered time as an independent variable and the level of attention as a dependent variable, and hypothesized that students' attention would be affected as time passes. A research hypothesis is therefore proposed as follows:

Hypothesis 1: There is a linear relationship between attention level and time.

Burns et al. (2011) argued that not all instructional materials have a significant impact on learners' understanding. However, a number of studies have suggested that the structure of learning content can have an important influence on the level of attention that learners pay to it (Bartsch and Cobern, 2003; Hosam et al., 2010). Based on the above argument, we aimed to investigate which mode of missing strokes in COS learning would facilitate attention. Thus, we propose the following hypothesis:

Hypothesis 2: There is a significant difference in the attention level of the stroke-appearing and stroke-disappearing groups when practicing COS.

According to Islam's (2013) findings, the use of e-learning systems can have a weak influence on students' academic performance, while Shih et al. (2008) argued that e-learning systems may not be beneficial in all learning situations. However, Stettler and Francis (2018) found that the classification of images requires the design of a human visual system that promotes good learning. Pituch and Lee (2006) suggested that all those involved in developing, designing, and purchasing e-learning systems should take the needs and values of the system users into careful consideration, and ensure that the system is able to meet those demands. In general, if learners perceive a high degree of system functionality and content features, there would be a higher performance level in e-learning. Based on the above literature, it is worth considering testing the learning outcomes of different modes of an e-learning system (e.g., stroke-appearing and stroke-disappearing). Hence, our third research hypothesis is proposed as follows:

Hypothesis 3: There is a significant difference in the learning outcomes of the stroke-appearing and stroke-disappearing groups when practicing COS.

RESEARCH INSTRUMENT

Materials

Hong et al. (2009) suggested that Drill and Practice with time pressure can be used to encourage players to work on a task by correctly applying knowledge, and can give them more opportunities to exercise strategies other than memorization.

In addition, Plass et al. (2014) examined the two design factors of color and shape to investigate which may evoke positive emotions. Accordingly, in this study we designed a COS e-learning system to be used as Drill and Practice material. The system was designed with two perception modes: stroke-appearing and stroke-disappearing. When practicing with the appearing mode, if a mistake is made in the order of writing a stroke, the wrong stroke will appear in red on the screen. When practicing with the disappearing mode, the incorrect stroke will disappear. **Table 1** shows how to play the Chinese character stroke order game.

How to Play

The research instrument, the COS e-learning system, is a computer adapted system. The administrator can add Chinese characters to the COS platform. For the purpose of this research, we added Chinese characters based on the Chinese textbook used by the participants. For third-grade elementary school students, there are eight lessons introducing 100 characters in the first semester. The administrator ensured that the stroke order of each character was correct. In this study, we chose Lesson 1 as the target content. All of the participants used a mouse to write the Chinese characters, for both the appearing and disappearing modes, while practicing COS. **Table 1** compares the two modes of COS practice. When learners practiced COS in the stroke-appearing mode, a stroke written in the wrong order was shown in red. In contrast, in the stroke-disappearing mode, if a mistake was made in the order of a stroke, the incorrect stroke disappeared. After practicing, the block window showed learners their score, the practicing time, and the percentage of correct and incorrect answers. The scoring formula in COS is as follows: $100 - (\text{number of incorrect strokes} \times \text{number of trials}) / (\text{total number of strokes} \times 100)$.

METHOD

Participants and Measuring Apparatus

Ten third-grade students from an elementary school in Taipei participated in this study. There were five boys and five girls, all between 7 and 8 years old. The students were native Chinese speakers with standard accent, and they were unaware of the purpose of the experiment.

Participants' attention was recorded with an electroencephalogram (EEG) apparatus. EEG is defined as alternating electrical activity that is detected and recorded by metal electrodes and conductive media placed on the surface of the scalp (Niedermeyer and Lopes da Silva, 1993). Human EEG studies posit that the alpha oscillations play a key role in visual attention (Thut et al., 2006; Sauseng et al., 2009). It has been found in previous EEG studies that sustained modulations of the oscillatory α -band (8–14 Hz) activity reflect changes related to attention due to the anticipation of visual events (Babiloni et al., 2002). In teaching, if applied to language, mathematics, and other e-learning materials, the EEG apparatus can be employed to monitor students' attention level to determine whether the mode enhances the students' learning attention (Aziz-Zadeh

TABLE 1 | Comparison of the COS game modes.

Appearing mode

When the player writes a stroke in the incorrect order, the wrong stroke will be shown in red.



After the learner finishes writing each Chinese character, the screen will display the strokes written in the wrong order in red.



When learners finish the game, a window will display the score, the number of times the character has been tested, and the percentage of correct and incorrect stroke orders. Learners can click on the bottom to re-play the game.



Disappearing mode

When the player writes a stroke in the incorrect order, the wrong stroke will become invisible.



After the learner finishes each Chinese character, the strokes written in the wrong order will become invisible. The invisible stroke is marked by a blue circle below.



et al., 2006). Thus, in this study we used a simple EEG apparatus to investigate participants' visual attention.

Procedure

We used the COS e-learning system to test students' performance of writing Chinese characters with corrective feedback. Five of the participants were randomly assigned to the stroke-appearing group and five to the stroke-disappearing group. They practiced

three times for 10 min each in an experiment session (i.e., one session a week over a period of 3 weeks). After the three sessions were completed, we obtained the participants' COS scores. We obtained the "Learning outcome" for each group (stroke-appearing and stroke-disappearing) by subtracting the scores of the first experiment from the scores of the second experiment. The result was used to analyze the difference between the two groups. Moreover, all of the EEG data collected during the

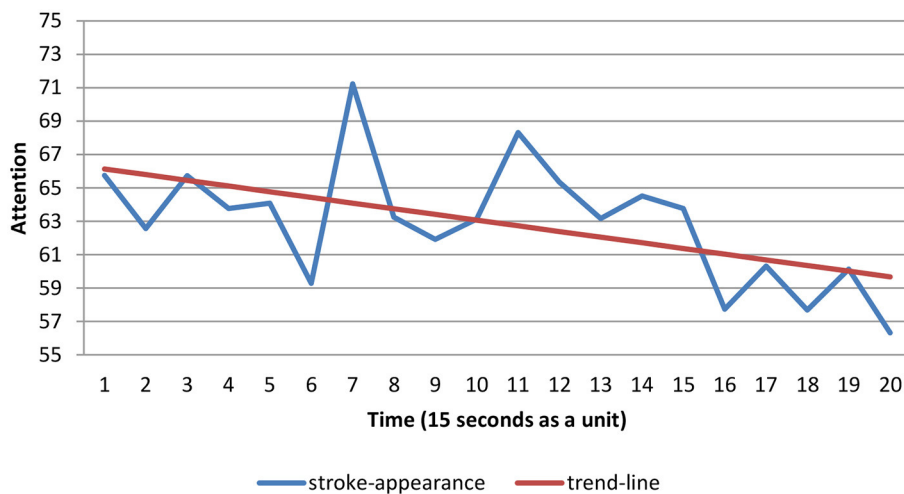


FIGURE 1 | Time series chart of stroke-appearance.

first experiment were realized in the frequency domain between 8 and 14 Hz using the EEG processing software from which the values of students' attention were collected while they played the first 5 min of COS. These attention values were used to examine the trend and analyze the differences between the two groups.

RESEARCH FINDINGS

The data analysis was conducted in two steps. Initially, we employed a time series analysis based on data collected from the EEG for attention level. Next, an independent-samples *t*-test was conducted to analyze the differences in attention level and learning outcomes.

Relationship Between Attention Level and Practice Modes

The EEG data while students played the first 5 min of COS were collected as a unit every 15 s to compute the mean scores. Twenty mean scores were computed to construct the time series charts. The attention level detected by the EEG device was retrieved from the database (Chen and Huang, 2014). A time series was employed to observe the variation in students' attention levels, and the trend line was calculated as a linear regression of the observed data using the least squares method, of which the forecast regression equation is as follows:

$$T_t = b_0 + b_1 t$$

where *t* is the unit of time, T_t is the value of forecast for the t^{th} observation, $b_0 = \bar{Y} - b_1 \bar{t}$ is the intercept of the trend line, and $b_1 = \frac{\sum tY_t - (\sum t \sum Y_t)/n}{\sum t^2 - (\sum t)^2/n}$ is the slope of the trend line.

As evaluated over the available record of the stroke-appearing mode (Figure 1), time and attention are fairly correlated ($r = -0.55^{***}$, $p < 0.001$). It should be noted, however, that this

correlation was mainly driven by the time series' strong linear trends; that is, the least squares method can be utilized to compute the regression equation, which was $T_t = 66.47 - 0.34t$ ($F = 42.50^{***}$, $p < 0.001$), and the slope (b_1) was equal to -0.34 , which indicated that students' attention levels decreased over time. As evaluated using the available record of the stroke-disappearing mode (Figure 2), time and attention are fairly correlated ($r = 0.54^{***}$, $p < 0.001$). Once again, the correlation was mainly driven by the strong time series' linear trends. That is, the least squares method can be utilized to compute the regression equation, which was $T_t = 44.97 + 0.36t$ ($F = 40.34^{***}$, $p < 0.001$), and the slope (b_1) was equal to 0.36 , indicating that the students' attention levels increased over time. The participants in the stroke-appearing group showed decreasing attention levels over time, whereas those in the stroke-disappearing group had increasing attention levels. Furthermore, the intercept of the trend line of the stroke-appearing mode was greater than that of the stroke-disappearing mode ($66.47 > 44.97$). It revealed that the stroke-appearing mode raised more attention than the stroke-disappearing mode.

Attention Levels With Different COS Learning Modes

One aim of this study was to determine whether there were significant differences in the attention levels of the stroke-appearing and stroke-disappearing modes of COS learning. An independent-samples *t*-test was conducted to examine the differences between the two modes in terms of attention levels. Table 2 shows that the results of the independent-samples *t*-test were $t = 2.632^*$ ($F = 0.070$, $p < 0.05$) revealing that there were significant differences in the attention levels of the stroke-appearing and stroke-disappearing groups. The mean score for the stroke-appearing group ($M = 62.45$, $SD = 5.76$) was higher than that of the stroke-disappearing group ($M = 49.21$, $SD = 9.67$), indicating that participants in the stroke-appearing group

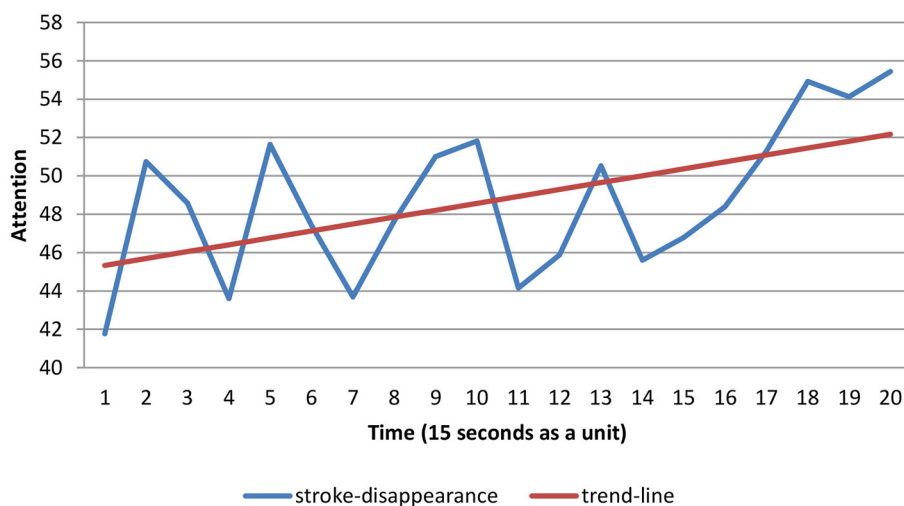


FIGURE 2 | Time series chart of stroke-disappearance.

TABLE 2 | Attention levels of the different COS modes.

Dimension modes		N	M	SD	F	t	p	R ²
Attention	Stroke-appearing	5	62.450	5.757	0.070	2.632*	0.030	0.464
	Stroke-disappearing	5	49.201	9.673				

* $p < 0.05$.

were more engaged in learning COS than those in the stroke-disappearing group. The result also showed that the effect size was $R^2 = 0.464$, meaning that 46.4% of the data could be explained.

Learning Outcomes of Different COS Modes

We also aimed to determine whether there were significant differences in the learning outcomes of the stroke-appearing and stroke-disappearing modes in the COS learning. An independent-samples t -test was conducted to examine the differences in the learning outcomes of the two modes. **Table 3** shows that the results of the independent-samples t -test were $t = 2.593^*$ ($F = 0.363$, $p < 0.05$), revealing that there were significant differences in the learning outcomes of the stroke-appearing and stroke-disappearing groups. The mean score for the stroke-appearing group ($M = 1093.67$, $SD = 207.95$) was higher than that of the stroke-disappearing group ($M = 530.71$, $SD = 438.70$), indicating that participants in the stroke-appearing group performed better while practicing COS than did those in the stroke-disappearing group. The result also indicated that the effect size was $R^2 = 0.457$, meaning that 45.7% of the data could be explained.

DISCUSSION

According to associative sequence learning (ASL) theory, the perception-action links allow us to explore the mechanisms underlying the modulation effects of COS by imitation that directs visuo-motor matching learning. We designed a COS system which was subdivided into two modes (stroke-appearing and stroke-disappearing) for students to practice writing characters, and we then examined the attention and outcomes of participants when practicing COS with these two different modes.

In examining hypothesis 1, the results of this study revealed that there was a linear relationship between attention level and time spent on COS learning. We found that under the stimulus of different modes, it was clear that participants' attention levels altered over time. Interestingly, students' attention decreased in the stroke-appearing group as time passed, but increased in the stroke-disappearing group; however, the stroke-appearing mode attracted more attention than the stroke-disappearing mode. The maximum of individuals' sustained visual attention is ~ 5 min (Nuechterlein et al., 1983), which may be caused by mental fatigue (Van Cutsem et al., 2017). Jollie et al. (2016) revealed that individuals tend to be affected by location-based expectancies which are generated by predictive visual cues. These findings were consistent with prior research. Hedges et al. (2013)

TABLE 3 | Learning outcomes of the different COS modes.

Dimension modes		N	M	SD	F	t	p	R ²
Learning outcomes	Stroke-appearing	5	1093.670	207.952	0.363	2.593*	0.032	0.457
	Stroke-disappearing	5	530.710	438.696				

* $p < 0.05$.

noted that timing is important for play activities and for the development of attention and learning. It has also been shown by several studies that adverse effects may result from a failure to adopt appropriate learning content design strategies (Botturi et al., 2006; Hwang et al., 2008). Our results support the previous studies which found that there is a linear relationship between attention level and time spent on COS learning.

In examining hypothesis 2, the results of this study revealed that there was a significant difference in the attention levels of students using the stroke-appearing and stroke-disappearing modes when practicing COS. We found that the attention levels of the stroke-appearing mode students were higher than those of the stroke-disappearing group. Thus, stroke-appearing is a better mode of learning than stroke-disappearing. Jiang (2018) proposed a multi-level framework of spatial attentional control, which considered that goals, perceptual salience, and selection history are the major sources to maintain highly efficient spatial attention. In short, the allocation and transfer of spatial attention may be affected by visual cues. Handy et al. (2003) proposed using cue stimuli that can enhance readiness and engage motor preparation processes. Burns et al. (2011) found that not all instructional materials can have a significant effect on learners' understanding. Many other researchers have also suggested that the structure of learning content can have an impact on the level of attention learners pay to it (e.g., Bartsch and Cobern, 2003; Hosam et al., 2010). The results of our study support those of previous research as they indicate that there was a significant difference in the attention of the stroke-appearing and stroke-disappearing groups when practicing COS.

In examining hypothesis 3, the results of this study revealed that there was a significant difference in the learning outcomes of the students using the stroke-appearing and stroke-disappearing modes when practicing COS. We found that the learning outcomes of the group using the stroke-appearing mode were higher than those of the group which practiced with the stroke-disappearing mode. Thus, stroke-appearing is a better mode of learning than stroke-disappearing; using an appropriate design for learning content may assist learners and positively affect their learning outcomes. Pituch and Lee (2006) suggested that researchers should carefully consider the needs and values of the system users and ensure that the system is well-developed and appropriate for the learner. Even though Islam's (2013) study found evidence that e-learning systems can have a weak influence on students' academic performance, the results of our study support those of other research (Stettler and Francis, 2018) as they indicate that there was a significant difference in the learning outcomes of students using the different modes of learning content design, that is, the stroke-appearing

and stroke-disappearing modes for COS, and suggest that an appropriate design will lead to a higher level of performance in e-learning. Conversely, an inappropriate design will lead to a lower level of performance.

CONCLUSIONS

As stroke exercises helped improve Chinese character memorization (Hsiung et al., 2017), and are perceived to be useful for producing accurate Chinese writing (Jaganathan and Lee, 2014), we designed two modes of sensor-evoked activities for participants to practice Chinese order of strokes (COS), namely stroke-appearance (i.e., marking the forgotten strokes in red) and stroke-disappearance (i.e., marking the forgotten strokes with blanks) to explore how perception of these two modes of action would implicitly modify memory, and then to highlight the effects of perception on Chinese stroke acquisition. The results of this study showed that the appearance of a cue would promote the (Chinese) learners' cognitive and affective development. This suggests that although the simulated stroke-appearing mode may have distracted the learners' attention as time passed, their learning outcomes were still better than those of the students who practiced with the stroke-disappearing mode, indicating that rich cues allow learners to focus their attention, and consequently enhance their learning.

To summarize, in this study we adopted the theoretical perspective of attention levels in order to study the cues of learning effects on the learning of Chinese character writing. Our findings indicate that the COS e-learning system may be an important and effective mode for enhancing students' Chinese character acquisition, and especially that of elementary school students. The findings of our study provide insights into Chinese character processing during learning, and have implications for the design of an effective method for teaching children to learn Chinese characters. Most importantly, educators may use COS e-learning in their classes as a method to simulate immersion to bridge the gap between the classroom and the real use environment.

LIMITATIONS AND FUTURE RESEARCH

The use of e-learning systems for teaching and learning Chinese strokes has become a common phenomenon in recent years. Although investigations of the different effects of stroke-appearing and stroke-disappearing have focused on attention levels and learning outcomes, there are other variables that have also been shown to influence students'

perceptions and performance in Chinese character learning tasks. Where this study used the two modes of stroke-appearing and stroke-disappearing to examine students' attention and learning outcomes, future studies may attempt to employ effective equipment or devices to explore participants' emotional data and investigate participants of different genders and ages.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

This study involving human participants were reviewed and approved by Research Ethics Committee of National Taiwan

Normal University. The participants provided written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

J-CH: original draft. K-HT: data collection and review and editing. M-YH: review and editing. P-HL: data analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the Chinese Language and Technology Center of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

REFERENCES

- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr. Biol.* 16, 1818–1823. doi: 10.1016/j.cub.2006.07.060
- Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Ciozza, G., Del Percio, C., et al. (2002). Human cortical electroencephalography (EEG) rhythms during the observation of simple aimless movements: a high resolution EEG study. *Neuroimage* 17, 559–572. doi: 10.1006/nimg.2002.1192
- Bartsch, R. A., and Cobern, K. M. (2003). Effectiveness of PowerPoint presentations in lectures. *Comput. Educ.* 41, 77–86. doi: 10.1016/S0360-1315(03)00027-7
- Bastiaansen, J. A., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philos. Trans. R. Soc. B* 364, 2391–2404. doi: 10.1098/rstb.2009.0058
- Botturi, L., Cantoni, L., Lepori, B., and Tardini, S. (2006). “Fast prototyping as a communication catalyst for e-learning design,” in *Making the Transition to E-Learning: Strategies and Issues*, eds M. Bullen and D. Janes (Hershey, PA: IGI Global), 266–283.
- Burns, J., Clift, J., and Duncan, J. (2011). Understanding of understanding: implications for learning and teaching. *Br. J. Educ. Psychol.* 61, 276–289. doi: 10.1111/j.2044-8279.1991.tb00985.x
- Catmur, C., Walsh, V., and Heyes, C. M. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531. doi: 10.1016/j.cub.2007.08.006
- Chen, C. M., and Huang, S. H. (2014). Web-based reading annotation system with an attention-based self-regulated learning mechanism for promoting reading performance. *Br. J. Educ. Technol.* 45, 959–980. doi: 10.1111/bjet.12119
- Cook, R., Bird, G., Catmur, C., Press, C., and Heyes, C. (2014). Mirror neurons: from origin to function. *Behav. Brain Sci.* 37, 177–192. doi: 10.1017/S0140525X13000903
- Greenlees, I., and Moran, A. P. (2003). *Concentration Skills Training in Sport*. London: British Psychological Society.
- Guitart, I. A., Hervet, G., and Hildebrand, D. (2019). Using eye-tracking to understand the impact of multitasking on memory for banner ads: the role of attention to the ad. *Int. J. Advert.* 38, 154–170. doi: 10.1080/02650487.2018.1473023
- Guo, C., Tsegaye, A., Arató, J., and Logemann, H. A. (2021). The role of attention, inhibition, and statistical learning in Chinese character recognition by novices. *Curr. Res. Behav. Sci.* 2:100012. doi: 10.1016/j.crbeha.2020.100012
- Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S., and Gazzaniga, M. S. (2003). Graspable objects grab attention when the potential for action is recognized. *Nat. Neurosci.* 6, 421–427. doi: 10.1038/nn1031
- Hedges, J. H., Adolph, K. E., Amso, D., Bavelier, D., Fiez, J. A., Krubitzer, L., et al. (2013). Play, attention, and learning: how do play and timing shape the development of attention and influence classroom learning? *Ann. N. Y. Acad. Sci.* 1292, 1–20. doi: 10.1111/nyas.12154
- Heyes, C. (2001). Causes and consequences of imitation. *Trends Cogn. Sci.* 5, 253–261. doi: 10.1016/S1364-6613(00)01661-2
- Heyes, C., Bird, G., Johnson, H., and Haggard, P. (2005). Experience modulates automatic imitation. *Brain Res. Cogn. Brain Res.* 22, 233–240. doi: 10.1016/j.cogbrainres.2004.09.009
- Heyes, C. M. (2011). Automatic imitation. *Psychol. Bull.* 137, 463–483. doi: 10.1037/a0022288
- Hong, J. C., Hwang, M. Y., Lu, C. H., Cheng, C. L., Lee, Y. C., and Lin, C. L. (2009). Playfulness-based design in educational games: a perspective on an evolutionary contest game. *Interact. Learn. Environ.* 17, 15–35. doi: 10.1080/10494820701483615
- Hopfinger, J. B., and Maxwell, J. S. (2005). Appearing and disappearing stimuli trigger a reflexive modulation of visual cortical activity. *Brain Res. Cogn. Brain Res.* 25, 48–56. doi: 10.1016/j.cogbrainres.2005.04.010
- Hosam, A. S., Abbas, M., and Naufal, I. (2010). “The design and development of exceptional representation based on domain ontology and multi-agent systems for e-learning purposes,” in *Mathematical/Analytical Modelling and Computer Simulation (AMS), 2010 Fourth Asia International Conference* (Kota Kinabalu: IEEE), 517–520.
- Hsiung, H. Y., Chang, Y. L., Chen, H. C., and Sung, Y. T. (2017). Effect of stroke-order learning and handwriting exercises on recognizing and writing Chinese characters by Chinese as a foreign language learners. *Comput. Hum. Behav.* 74, 303–310. doi: 10.1016/j.chb.2017.04.022
- Hutchinson, J. B., and Turk-Browne, N. B. (2012). Memory-guided attention: control from multiple memory systems. *Trends Cogn. Sci.* 16, 576–579. doi: 10.1016/j.tics.2012.10.003
- Hwang, G. J., Tsai, C. C., and Yang, S. J. (2008). Criteria, strategies, and research issues of context-aware ubiquitous learning. *J. Educ. Technol. Soc.* 11, 81–91. Available online at: https://www.jstor.org/stable/jeductechsoci.11.2.81?seq=1#metadata_info_tab_contents
- Islam, A. K. M. N. (2013). Investigating e-learning system usage outcomes in the University context. *Comput. Educ.* 69, 387–399. doi: 10.1016/j.compedu.2013.07.037
- Jaganathan, P., and Lee, P. L. (2014). Knowledge and perception of stroke order among Chinese-as-a-foreign language students in a Malaysian university. *Int. J. Educ. Res.* 2, 147–160. Available online at: <https://www.ijern.com/journal/2014/November-2014/13.pdf>
- Jiang, N., Hou, F., and Jiang, X. (2020). Analytic versus holistic recognition of Chinese words among L2 learners. *Mod. Lang. J.* 104, 567–580. doi: 10.1111/modl.12662

- Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex* 102, 107–120. doi: 10.1016/j.cortex.2017.06.018
- Jollie, A., Ivanoff, J., Webb, N. E., and Jamieson, A. S. (2016). Expect the unexpected: a paradoxical effect of cue validity on the orienting of attention. *Atten. Percept. Psychophys.* 78, 2124–2134. doi: 10.3758/s13414-016-1164-x
- Kremer, J., and Moran, A. P. (2012). *Pure Sport: Practical Sport Psychology*. London: Routledge.
- Li, C., Zhang, J., and Yao, J. (2021). Streamer action recognition in live video with spatial-temporal attention and deep dictionary learning. *Neurocomputing* 453, 383–392. doi: 10.1016/j.neucom.2020.07.148
- Lin, G., Guo, Z., Chao, F., Yang, L., Chang, X., Lin, C. M., et al. (2021). Automatic stroke generation for style-oriented robotic Chinese calligraphy. *Future Gen. Comput. Syst.* 119, 20–30. doi: 10.1016/j.future.2021.01.029
- Lyu, B., Lai, C., Lin, C. H., and Gong, Y. (2021). Comparison studies of typing and handwriting in Chinese language learning: a synthetic review. *Int. J. Educ. Res.* 106:101740. doi: 10.1016/j.ijer.2021.101740
- McCombs, B., and Vakili, D. (2005). A learner-centered framework for e-learning. *Teach. Coll. Record* 107, 1582–1600. doi: 10.1111/j.1467-9620.2005.00534.x
- Niedermeyer, E., and Lopes da Silva, F. H. (1993). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 3rd Edn. Philadelphia, PA: Williams and Wilkins.
- Nuechterlein, K. H., Parasuraman, R., and Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science* 220, 327–329. doi: 10.1126/science.6836276
- Ossmy, O., and Mukamel, R. (2018). Perception as a route for motor skill learning: perspectives from neuroscience. *Neurosci. Frontiers Rev.* 382, 144–153. doi: 10.1016/j.neuroscience.2018.04.016
- Pituch, K. A., and Lee, Y. K. (2006). The influence of system characteristics on e-learning use. *Comput. Educ.* 47, 222–244. doi: 10.1016/j.compedu.2004.10.007
- Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., and Umd, E. (2014). Emotional design in multimedia learning: effects of shape and color on affect and learning. *Learn. Instruct.* 29, 128–140. doi: 10.1016/j.learninstruc.2013.02.006
- Pratt, J., and McAuliffe, J. (2001). The effects of onsets and offsets on visual attention. *Psychol. Res.* 65, 185–191. doi: 10.1007/s004260100058
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., et al. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852. doi: 10.1016/j.cub.2009.08.062
- Shih, M., Feng, J., and Tsai, C. C. (2008). Research and trends in the field of e-learning from 2001 to 2005: A content analysis of cognitive studies in selected journals. *Comput. Educ.* 51, 955–967. doi: 10.1016/j.compedu.2007.10.004
- Similä, S. S., and McIntosh, R. D. (2015). Look where you're going! Perceptual attention constrains the online guidance of action. *Vis. Res.* 110, 179–189. doi: 10.1016/j.visres.2014.06.002
- Stettler, M., and Francis, G. (2018). Using a model of human visual perception to improve deep learning. *Neural Netw.* 104, 40–49. doi: 10.1016/j.neunet.2018.04.005
- Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502. doi: 10.1523/JNEUROSCI.0875-06.2006
- Tung, C. H., and Jean, E. Y. (2018). Stroke-order-free on-line Chinese character recognition by stroke adjustment of two-layer bipartite weighted matching. *Future Gener. Comput. Syst.* 81, 219–234. doi: 10.1016/j.future.2017.09.074
- Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., and Roelands, B. (2017). The effects of mental fatigue on physical performance: a systematic review. *Sports Med.* 47, 1569–1588. doi: 10.1007/s40279-016-0672-0
- Wang, C. C., Hung, J. C., Chen, S. N., and Chang, H. P. (2019). Tracking students' visual attention on manga-based interactive e-book while reading: an eye-movement approach. *Multimed. Tools Appl.* 78, 4813–4834. doi: 10.1007/s11042-018-5754-6
- Xu, Y. J., Chiou, S. C., and You, M. (2020). Effects of improving the interactive design of a Chinese character learning system on the learning performance of Chinese as foreign language students. *Comput. Assist. Lang. Learn.* 33, 916–935. doi: 10.1080/09588221.2019.1599961
- Yantis, S., and Serences, J. (2003). Neural mechanisms of space-based and object based attentional control. *Curr. Opin. Neurobiol.* 13, 187–193. doi: 10.1016/S0959-4388(03)00033-3
- Yin, Y., Cai, X., Ouyang, M., and Zhang, Q. (in press). The N200 enhancement effect in reading Chinese is modulated by actual writing. *Neuropsychologia* 142. doi: 10.1016/j.neuropsychologia.2020.107462
- Yin, Y., and Zhang, Q. (2021). Chinese characters are read using not only visual but also writing motor information. *Psychophysiology* 58:e13696. doi: 10.1111/psyp.13696

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Hong, Tai, Hwang and Lin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cognitive Strategies and Textual Genres in the Teaching and Evaluation of Advanced Reading Comprehension (ARC)

Jesús-Nicasio García-Sánchez^{1,2*} and Judit García-Martín³

¹ Department of Psychology, Sociology and Philosophy, Universidad de León, León, Spain, ² Department of Humanities, Universidad De La Costa, Barranquilla, Colombia, ³ Department of Developmental and Educational Psychology, Universidad de Salamanca, Salamanca, Spain

OPEN ACCESS

Edited by:

Pedro García Guirao,
University of Ostrava, Czechia

Reviewed by:

José Manuel Aguilar Parra,
University of Almería, Spain
Magdalena Ramos Navas-Parejo,
University of Granada, Spain

*Correspondence:

Jesús-Nicasio García-Sánchez
jn.garcia@unileon.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 10 June 2021

Accepted: 13 September 2021

Published: 12 October 2021

Citation:

García-Sánchez J-N and
García-Martín J (2021) Cognitive
Strategies and Textual Genres in the
Teaching and Evaluation of Advanced
Reading Comprehension (ARC).
Front. Psychol. 12:723281.
doi: 10.3389/fpsyg.2021.723281

In the last decade, published data on the performance of Colombian students have concerned educators and researchers, making critical reading one of the priorities of Colombian education. That is why this article presents the results of a study carried out in a Latin American university in which the perceptions of students and professors are analyzed regarding the strategies and textual genres used to work and cross-evaluate the advanced reading comprehension (ARC). This study is materialized in the application of an *ad hoc* online questionnaire, in its two versions (students and teachers), designed through Survey Monkey. For this, it has the participation of 182 teachers and 2,775 students. There are several trends in the use of specific textual strategies and typologies to work and evaluate ARC, by both, depending on the department of assignment. The evidence found is provided and evaluated considering the implications for cross-curricular instruction and assessment in higher education in Latin America, including study limitations and prospects for overcoming them.

Keywords: advanced reading comprehension, instruction strategies, assessment systems, generic competences across curriculum, higher education

INTRODUCTION

Government agencies, in order to comply with the standards of the Ministry of Education, publish, every 3 years, the results obtained in the international PISA tests, regarding the performance of students in different areas of knowledge, among which are, reading (OCDE, 2016). In this sense, in the last decade, published data on the performance of Colombian students have concerned educators and researchers (Díaz et al., 2015; García J. N. et al., 2019; García J. R. et al., 2019; Inciarte et al., 2019; Marín et al., 2019; Rueda et al., 2019). That is why critical reading or advanced reading comprehension has become one of the priorities of Colombian education (Cassany, 2003; Cubides et al., 2017).

In this sense, in 2003, Cassany affirms that critical reading or advanced reading comprehension (hereinafter, ARC) is an explicit approximation of a demanding and complex reading typology that apart from the essential requirements of literal, inferential and intentional of the text, requires higher demands, as well as the interest expressed by the reader's understanding. All this implies implicitly the recovery of the connotations of the words used, the identification of the author's point of view, with special attention to the use of sarcasm and irony..., the distinction between the voices used and of course the identification of the textual genre employed in speech. In line

with the above, in the Colombian higher education space, the mandatory national test, *Saber Pro*, is the most important source of quality to examine five modules: (i) critical reading, (ii) quantitative reasoning, (iii) citizen competence, (iv) written communication and (v) English (García J. N. et al., 2019; García J. R. et al., 2019; Cabeza et al., 2020; Calderón et al., 2021). According to Calderón et al. (2021), the module related to critical reading has a total of 35 questions that evaluate three types of comprehension: (i) literal and explicit, in terms of the meaning of words and expressions used. and that represents 25% of the test, (ii) the global one that involves 40% of the questions and (iii) the criticism that covers 35% of the questions posed and that is based on the recognition of the strategies used, the identification of the different assumptions, the validity of the arguments. . . .

The role that advanced reading comprehension plays in the rest of the generic and specific competences is essential. For example, the quality of the written composition (written communicative competence) are designed from the planning processes that involve the generation of information and ideas, or the selection of ideas and review of the written message, so that reading comprehension becomes essential. (García and García-Martín, 2021a,b; Graham, 2021; López et al., 2021; Robledo-Ramón and García-Gutiérrez, 2021). The same happens with citizen competence, where critical thinking represents 70% of the assessment of competence, which implies the resolution of citizen problems, which require a high dose of advanced reading comprehension, as in the case of the development of lateral reading strategies and contrast of facts in the face of hoaxes and fakes (Brodsky et al., 2021; Cabrera et al., 2021). It is evident that advanced reading comprehension plays a role in generic quantitative reasoning competence. Solving computational problems, which involves quantification, owes much to the mastery of advanced reading comprehension. Evidence from the analysis of the semester tests of generic competences in the reference institution, as well as the analysis of more than a decade of the big data available in the ICFES of the MEN of Colombia, indicate that reading comprehension is essential for the quantitative competence (and general intelligence), and others such as English. It is evident that he also plays it for the different specific competences, typical of each profession and program or career (García and Jiménez, 2017; García-Martín and García-Sánchez, 2020). For example, the national *Saber Pro* test involves three types of components for the assessment of quantitative reasoning competence: interpretation and representation; formulation and execution; and argumentation. In all three components, advanced reading comprehension is essential; in addition to asking contextual or situational questions, both family and personal, work or occupational; social or community; and in that of scientific dissemination (Calderón et al., 2021). A competency, also transversal and generic, that is gaining more importance every day, is digital, for which high doses of advanced reading comprehension are also required and which plays a key role in enhancing the rest of the generic competences and specific, in addition to any higher learning (García-Martín and García-Martín, 2021).

Cognitive Strategies and Textual Genres for the Teaching and Assessment of ARC

Higher Education Institutions in Colombia focus on examining the quality of the processes used (Orozco, 2010; Arias et al., 2018). In this sense, Colombian universities, as the focus of this study, are presented as higher education institutions that promote cognitive strategies and textual genres that enhance the transversal development-through the curriculum-of critical reading or advanced reading comprehension of the students (Samper and Ospino, 2018), making them competent to respond to the needs and demands of today's multicultural, globalized, and dynamic society (Barnet, 2001; Castellar et al., 2021).

On the one hand, the review of previous studies and antecedents allowed to identify the different components, processes, strategies, genres, and instructional means involved in the evaluation and teaching of advanced reading comprehension. From the perspective of cognitive psychology, for example, the psychology of reading, it seems reasonable to consider processes, techniques and strategies related to the activation of previous knowledge, the relationship of meaning between elements, the ability to draw inferences, generation of mental models that allow the representation of the meaning of the text, the ability to apply meanings to other areas of knowledge and life; as well as the structure reflected through the most common textual genres in the university environment that are part of the textual macrostructure (argumentative, comparison and contrast, cause and effect, problem solution); and finally, the means deployed to access textual information, be it traditional or virtual.

From the model that is deployed in the *Saber Pro* tests (Calderón et al., 2021); and on the other, from the broad tradition of cognitive psychology, instructional psychology, and scientific psychology in general, many instruments, empirical investigations, contrasted evidence in instructional or intervention studies were analyzed, it was possible to construct the instruments used in this study research (García-Martín and García-Sánchez, 2020). The interest on the subject is reflected in the systematic reviews that analyze reading and writing through content (literacy across content) in which it seems necessary to answer questions about the moment, the procedures, the actions, in relation to the introduction from the beginning. assessment and teaching of advanced reading comprehension in the different disciplines and subjects of all university degrees. But, in addition, these reviews show the scarcity of studies that provide empirical validation. The example from the Scott and Washburn (2018) is paradigmatic. After analyzing 50 years of applications of literacy across content in the training of teachers in the USA, after contrasting its methodological quality, only twenty-nine studies can be included, and of these, only four allowed a quantitative comparison and therefore generalizable with other studies. The foci of what they perceive to be happening, of resistance to changes, and of the experience deployed in the training of new teachers, seem to be the key variables in this story. Along the same lines is the systematic review by Miller et al. (2018). This situation in Europe is not more encouraging, as the review by Uttl et al. (2017) refers to, making a meta-analysis of the previous meta-analyzes on the evaluation of teachers by the perceptions

of their students about their satisfaction with the evaluation and teaching, a common strategy in the different universities of the world, is disappointing, since, ultimately, the evidence provided by the empirical studies analyzed depends on the size of the samples, since with large samples no effects are found (that is why there are no only published studies), and yes with small samples (published studies). It is necessary to recognize the need to implement different strategies to verify the role of evaluation systems and instructional techniques, strategies, procedures, and processes in the deployment of generic competences in a transversal way in all subjects of the degree. The perception of the level of satisfaction with the evaluation (Reyes et al., 2020), or of the teaching methods used (Sánchez et al., 2019; Jiménez et al., 2020), or the types of actions deployed with advanced reading comprehension (Valero et al., 2015) or even the type of strategies such as the narratives deployed (Del Moral-Pérez et al., 2016) seem to mark relevant focuses and variables that must be analyzed. In the cultural field Asian, reflects the same trend. For example, the systematic review by Li et al. (2018) finds fifty-nine studies in Mainland China that meet the inclusion and exclusion criteria published during the previous 20 years, focused on constructs and focuses related to self-regulation of learning, finding as key variables the beliefs of capacity with learning, instructional strategies and the capacity for self-awareness and self-evaluation of instructional processes as a guarantee of self-regulatory capacity for productive and effective learning. Of these fifty-nine studies, only four provide data that allow comparative calculations, with two or more groups: the rest being only pre-post comparisons. The need to identify the variables responsible for the improvement of generic competences seems evident. And this is what contributes, also the impressive review of Peng et al. (2019) in which they find about fifty thousand studies, of which six hundred fifty meet the inclusion and exclusion criteria that they analyze with the meta-analysis methodology, referred to the role of fluid and crystallized intelligence in math instruction and reading comprehension. The above limitations are confirmed, in addition to the need to study evaluation and instruction variables, psychoeducational and instructional variables, such as those analyzed in this study.

Research Context and Theoretical Axes

It is a general project of analysis and promotion of the mainstreaming of generic competences or through subjects (García J. N. et al., 2019; García J. R. et al., 2019). The latest development has led to the creation of a skills observatory, initially focusing on generic skills, but with a view to including specific skills as well (García-Martín and García-Sánchez, 2020). On the one hand, this underlies, supports and covers the national final career tests of Saber Pro (Calderón et al., 2021), and on the other, the internal semester tests that are similar to the previous ones, but that affect the diagnosis from the first semester of each career, the state of the domain of generic and specific competences of all students, in order to strengthen and enhance them through actions designed in the evaluation and teaching of each subject (Sánchez et al., 2019). Within this framework, studies have been implemented on the teaching and evaluation of generic competences of advanced

reading comprehension, written communication and, last year, on civic competence (with very large components of advanced reading comprehension, in addition to critical thinking), along with other psychoeducational variables, such as self-efficacy, emotional intelligence and social performance, coping strategies in the face of problems, attitudes and anxiety in the face of these competences, among others, including mastery of the digital competence (García-Martín and García-Sánchez, Submitted). From the institutional perspective and high quality accreditation, the creation of the competences observatory is an unbeatable research context, in addition to serving as a catalyst and catalyst for the design of research, initiatives, innovations, promoting the quality of teaching, through the mainstreaming of competences in all subjects of all careers (CUC, 2018; García and García-Martín, 2021a,b).

The study's frame of reference focuses on five clearly interwoven and integrated axes (Marín et al., 2019). First, competency-focused teaching and the EHEA (European Higher Education Area), for the specific case of Colombia, the OCDE report is essential for understand it (OCDE, 2016). Second, active university methodologies toward a ubiquitous web 4.0. The latest advances regarding the introduction of active methodologies and innovations in university teaching are key, including advances and studies with digital competence, and the so-called web 4.0 (García-Martín and García-Sánchez, Submitted). In addition, a psychological and curricular instructional perspective. The instructional advances and curricular innovations are undeniable, both coming from the psychological, educational and contextual disciplines, which provide a first-line frame of reference for promoting the learning of competencies in the different subjects, including the psychoeducational variables, which are key in this process. (García-Martín and García-Sánchez, 2020; García-Martín and García-Martín, 2021). Fourth, the latest developments in instrument validation technology. The validation of instruments is key in any process of a didactic nature or of investigation. In the project of Evaluation and Methodologies and Incidents in Competences Generic (EMICOG) aimed at studying teaching methodologies and forms of evaluation cross-curricular subjects, from its rigorous and scientific nature, it is necessary to design, validation and application of instruments that in a standardized way allow the collection appropriate information (Chen and Lin, 2018; Cuesta et al., 2018; Habók and Magyar, 2018; Martínez-Ferrer et al., 2018; Puente-Martínez et al., 2018; Romera et al., 2018; Sinval et al., 2018; Zeng et al., 2019). And finally, the perspective of empirically-based interventions and quality indicators of innovations. The practices based on scientific evidence or empirically validated, have been the basis for selecting innovations that work, that are successful, and that should be promoted and promoted. This is the case of the best practices of the USOE (2021), or of the guidelines, protocols, validated programs promoted by the CEEDAR (2021).

Research Question

In line with the above, the *research question* of this study is: what does the analyze of techniques, strategies, cognitive processes contribute; as well as the textual genres used in teaching and

in the cross-sectional evaluation of critical reading or advanced reading comprehension, in a Latin American university, from a double perspective: that of the teaching staff and that of the students? This raises various hypotheses, on the one hand, it is expected that teachers show a greater knowledge of cognitive strategies and textual genres to teach and evaluate ARC than students, it is expected that there are differential patterns in the use of cognitive strategies and textual genres to teach and evaluate the ARC by teachers and students, according to the department to which they are attached.

METHOD

Participants

The sample of this study is made up of 203 teachers (125 men and 78 women) aged between 24 and 73 years with a mean age of 41 years and 2,775 students (1,336 men and 1,439 women) with a mean age of 22 years. The collaboration of all the students and all the professors of the institution was requested. For professors 525 permanent (40% international) and 204 non-permanent. 280 agreed to participate, although 203 answered all the questions. For the students, 3,645 agreed to collaborate, although with the complete answers 2,775, which represents ~30% of the total of students in all grades and semesters. With a simple random sampling, confidence level of 99% and margin of error of 1% (CUC, 2018; García-Martín and García-Sánchez, 2020; García and García-Martín, 2021a,b). This guarantees great diversity and representativeness to be considered a high-quality type institution in Latin America. All of them assigned to eight departments (Computer Science, Economics, Exact, Environmental, Law, Energy, Industrial and Humanities) of the focal university of study, which decide to participate in an informed and voluntary way, through the completion of the *ad hoc* online questionnaire, EMICOG. The departments included in the sampling and analysis are the eight academic and specific, since the University Extension is transversal and for the purposes of teacher training, as well as for the promotion of the mainstreaming of competences, among other functions. The data of all the participants are checked and verified (see **Table 1**), being representative of a high quality accredited Latin American institution.

In relation to the professional category of the participating teachers, they are mostly assistants, followed by adjuncts II and III, and to a lesser extent, holders II and III. On the other hand, in terms of non-university teaching experience, the average is 6 years, ranging from 0 to 45 years compared to university teaching experience, whose average is 10 years.

Instrument

An *ad hoc* online questionnaire is applied through the online survey tool, Survey Monkey, in its two versions, EMICOG-teachers (García-Martín et al., 2019a; García-Martín and García-Sánchez, 2020) and EMICOG-students (García-Martín et al., 2019b).

The EMICOG is previously validated and with the study samples, with KLM sampling adequacy data (>0.01), indicating McDonald's omega compound reliabilities above 0.90 as well as

Cronbach's alpha internal consistencies above 0.85; with a mean variance extracted >0.50 (convergent validity); the discriminant validity (square root of the mean variance extracted) is higher than the intercorrelations between the factors, as expected. For this reason, the construct validity is adequate with indices above what is desirable, like the rest of the indicators. Likewise, the confirmatory factor analysis (CFA) provides adequate measures, in accordance with what is recommended for a CFA, obtaining an NFI (Normed Fit Index), a TLI (Tucker-Lewis Index) and a CFI (Comparative Fit Index), above of 0.90 (as recommended); and with an RMSEA (Root Mean Square Error of Approximation), below 0.08 (as recommended).

The EMICOG consists of four sections: (i) demographic data; (ii) the level of knowledge about the mainstreaming of the ARC; (iii) the instructional strategies used to teach and cross-evaluate ARC and (iv) the textual genres used to teach and cross-evaluate ARC. On the one hand, in relation to the instructional strategies used, in the different type subjects, both to teach and to evaluate advanced reading comprehension or critical reading, the extraction of main ideas, the establishment of relationships between the ideas of some readings are examined. With the previous knowledge, the development of conclusions and inferences not explicit in the readings, the application of solutions suggested in the readings to other aspects, the explanation of the content of some reading and the recovery of the information previously read and its use to instruct. On the other hand, in the case of textual genres used for the teaching and evaluation of ARC, argumentative genres, comparison and contrast, essays, literary analysis, bibliographic reviews, cause-effect or problem-solution, definitions and reviews on the state of the question.

The questions included in the survey were of the Likert type, with five response options (see complete survey for teachers in García-Martín and García-Sánchez, 2020; that of students is very similar with the adapted questions). The duration of completion of the same was 30–40 mins. The responses were made through the Survey Monkey platform. Participations were requested on the occasion of the institutional semester tests of generic competencies in four waves, with all students of all grades participating. The request for participation came from the direction of the Center for Teaching Excellence and the Academic Vice-Rector's Office, urging their participation, both for all teachers and all students. The completion was carried out during the 4 weeks of application of the generic skills tests on the Survey Monkey platform, a professional survey, open 24 hours a day, 7 days a week. After voluntarily accepting to participate, and giving their informed consent, they completed the surveys. The advantages of Survey Monkey are many, registering all the responses, partial or complete, the start and end time, the potential for simultaneous response of thousands of participants, nested response options, among others. The order followed by the instruments was the same for all, according to the description of the different scales described.

Design and Procedure

An exploratory-descriptive design is followed that is materialized in the application of the two versions (teachers and students) of an *ad hoc* online questionnaire, the EMICOG, which analyses

TABLE 1 | Description of the participants.

	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	Total group
Teachers	16	27	33	22	36	25	26	18	203
Students	303	592	138	836	127	125	293	361	2,775
Total department	319	619	171	858	163	150	319	379	2,978

the cognitive strategies and textual genres used in teaching and in cross-sectional evaluation of CRA or advanced reading comprehension, from a double perspective.

Analysis of Data

In the first place, descriptive analysis of the participating teachers and students are carried out, as well as the verification of the normality of the variables (asymmetry and kurtosis), evidencing that these are normally distributed, so it is appropriate to carry out the parametric analysis. Immediately, reliability and validity analysis of the instrument were carried out in its two modalities (item-scale internal consistency, Cronbach's alphas, composite reliability, extracted mean variance, discriminant validity), using exploratory factor analysis with SPSS v26 (García-Martín et al., 2019a,b) and confirmatory with AMOS v26 with Gaskination's StatWiki plugins (<http://statwiki.gaskination.com/index.php?title=Plugins>) the Pattern Matrix Model Builder from the pattern matrices were used; as well as the Model Fit Measures, the Validity and Reliability Test and other functionalities. For the calculation of the composite reliabilities, extracted mean variances, convergent and discriminant validity, with Excel from the factor matrices, too. Subsequently, multivariate analysis is carried out based on the general linear model (GLM), considering the department as the grouping variable, the demographic variable common to the two versions of the instrument, and as dependent the rest of the measures on the instruction and evaluation of the critical reading or advanced reading comprehension, through the SPSS version 26, which shows statistically significant differences with high effect sizes.

RESULTS

As indicated above, the multivariate contrasts (λ Wilks = 0.003; $F = 1,277$; $p = 0.001$; $\eta^2 = 0.523$) of the analysis of variance (ANOVA), carried out through the General Linear Model (GLM), show statistically significant differences with large effect sizes when the department to which the participants are assigned is considered as a grouping variable. We can see the inter-subject effects of professors in the **Table 2** for the teaching and assessment of the advanced reading comprehension.

The inter-subject contrasts effects of the multivariate analyzes, referring to the students, are included in detail in **Table 3**.

For a better understanding of the results, different figures are included below that provide a global vision. In this sense, as can be seen in **Figure 1**, when examining the variable of degree of knowledge that teachers and students have about the cross-sectional teaching of the ARC, in the inter-subject tests it is shown that the scores obtained by teachers are higher than

students regardless of the department to which they belong (e.g., $M_{\text{Humanities teachers}} = 4.06$ vs. $M_{\text{Humanities students}} = 3.51$; $p \leq 0.01$; e.g., $M_{\text{Environmental teachers}} = 3.79$ vs. $M_{\text{Environmental students}} = 3.26$; $p \leq 0.01$).

Cognitive Strategies for Teaching and Cross-Evaluating ARC

When the results obtained in the cognitive strategies for teaching and cross-evaluating ARC are compared, by teachers and students from different departments, several trends are observed. On the one hand, as can be seen in **Figure 2**, the trend described above is maintained, that is, teachers show higher scores than students. Likewise, it is observed that, of the cognitive strategies examined: the extraction of main ideas, the establishment of relationships between the ideas of some readings with previous knowledge, the development of conclusions and inferences not explicit in the readings, the application of solutions suggested in the readings to other aspects, the explanation of the content of some reading and retrieval of previously read information use for instructing. Statistically significant differences are only shown in establishing relationships between the ideas of the readings with previous knowledge, perceiving that the use of this is lower among Computer Science students followed by those of Environmental, Energy, Exact, Industrial, Economic, Law and Humanities both in teaching and in evaluation. On the other hand, in the case of professors, the order is modified starting with those of the Industrial department and followed by those of Environmental, Economics, Exact, Informatics, Law, Energy and Humanities.

Textual Genres to Teach and Cross-Evaluate ARC

Regarding the textual genres examined to teach and cross-evaluate ARC, statistically significant differences are only evidenced in the use of argumentative texts (see **Figure 3**) and in trials (see **Figure 4**), not being observed in comparison and contrast, literary analysis, bibliographic reviews, cause-effect or problem-solution, definitions, and reviews on the state of the question.

In this sense, as can be seen in **Figure 3**, the trend shown above is partially maintained, in the case of argumentative texts, that is, teachers achieve higher scores than students except for those assigned to the departments of Environmental, Industrial and Economic. In addition, in the case of students there is a trend of greater use depending on the department [Computer Science, Environmental, Exact and/or Energy (these last two departments are the only ones whose order is altered), Industrial, Economics, Law and Humanities]. In the same way, in the case of teachers,

TABLE 2 | Tests of the inter-subject effects of the professors on the instructional variables in ARC considering the department as a grouping variable.

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Please indicate your level of knowledge ... (min 1–max 5)											
Of the transversal teaching of generic competences of critical reading and textual construction in the subject	4.23 (0.83)	3.48 (0.75)	3.79 (0.78)	4.14 (0.77)	3.82 (0.77)	4.05 (0.50)	3.44 (0.98)	4.06 (1.00)	2.13	0.04	0.10
Advanced reading comprehension—has ever been used in this subject tasks, techniques or teaching strategies (or instructional) of... (mín 1–max 5)											
Strategies and processes											
Relate ideas from some reading with previous knowledge (WORKED)	4.31 (0.63)	4.10 (0.77)	3.71 (1.04)	3.93 (1.07)	4.25 (0.65)	4.33 (0.58)	3.56 (1.20)	4.56 (0.51)	2.78	0.01	0.13
Relate ideas from some reading with previous knowledge (EVALUATE)	4.23 (0.60)	4.05 (0.81)	3.67 (1.05)	4.07 (0.92)	4.29 (0.66)	4.33 (0.58)	3.50 (1.34)	4.56 (0.51)	2.91	0.01	0.14
Relate ideas from some reading with previous knowledge (UTILITY)	4.46 (0.66)	4.24 (0.77)	3.96 (1.00)	4.43 (0.65)	4.32 (0.61)	4.19 (0.68)	3.67 (1.19)	4.75 (0.45)	2.71	0.01	0.13
Textual genres											
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (WORKED)	3.69 (0.63)	3.43 (1.17)	3.13 (1.12)	4.07 (0.83)	4.11 (0.99)	3.57 (1.36)	3.39 (1.38)	4.31 (0.60)	2.63	0.01	0.13
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (EVALUATE)	3.62 (0.65)	3.38 (1.20)	3.17 (1.13)	4.14 (0.77)	4.07 (1.05)	3.48 (1.33)	3.28 (1.45)	4.25 (0.68)	2.55	0.01	0.12
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (UTILITY)	3.77 (0.73)	3.86 (0.85)	3.50 (1.18)	4.36 (0.75)	4.29 (0.81)	3.57 (1.21)	3.22 (1.44)	4.38 (0.72)	3.10	0.00	0.14
Reading of some comparison and contrast text : two theories, concepts, stories, authors, figures, preferences (WORKED)	3.38 (0.77)	3.29 (1.19)	3.42 (0.97)	3.57 (1.02)	4.00 (0.90)	3.86 (0.96)	3.06 (1.16)	3.88 (0.89)	2.05	0.04	0.10
Reading of some comparison and contrast text : two theories, concepts, stories, authors, figures, preferences (UTILITY)	3.62 (0.87)	3.71 (0.96)	3.54 (0.93)	4.14 (0.77)	4.25 (0.65)	3.81 (0.98)	3.22 (1.22)	4.06 (0.93)	2.50	0.01	0.12
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (WORKED)	4.00 (0.58)	3.57 (1.03)	3.33 (0.96)	3.79 (0.80)	3.86 (1.08)	3.62 (1.12)	2.94 (1.26)	3.94 (0.77)	2.01	0.05	0.10
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (EVALUATE)	3.92 (0.64)	3.48 (1.03)	3.38 (0.97)	3.79 (0.80)	3.86 (1.08)	3.62 (1.12)	2.72 (1.18)	4.00 (0.73)	2.82	0.01	0.13
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (UTILITY)	4.08 (0.64)	3.76 (0.83)	3.50 (1.02)	3.93 (0.83)	4.07 (0.94)	3.71 (1.19)	3.11 (1.37)	4.19 (0.83)	2.11	0.04	0.10

(Continued)

TABLE 2 | Continued

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Reading of some analysis text (description, literary analysis, process analysis) (EVALUATE)	4.00 (0.91)	3.29 (0.96)	3.42 (1.18)	3.57 (1.16)	4.04 (0.79)	3.95 (0.92)	3.11 (1.37)	3.56 (1.26)	2.06	0.04	0.10
Reading of some analysis text (description, literary analysis, process analysis) (UTILITY)	4.15 (0.90)	3.71 (0.90)	3.58 (1.18)	4.00 (0.68)	4.14 (0.65)	4.05 (0.92)	3.06 (1.51)	4.13 (1.09)	2.42	0.02	0.12
Medium used											
Reading some text in blogs (internet tool) (WORKED)	3.08 (1.26)	3.29 (1.10)	2.96 (1.12)	2.71 (1.44)	3.04 (1.07)	2.76 (1.14)	1.78 (1.06)	2.88 (1.03)	2.88	0.01	0.14
Reading some text in blogs (internet tool) (EVALUATE)	3.00 (1.29)	3.19 (1.12)	2.87 (1.23)	2.64 (1.45)	2.96 (1.04)	2.71 (1.15)	1.89 (1.18)	2.88 (1.03)	2.37	0.02	0.11
Reading some text in blogs (internet tool) (UTILITY)	3.00 (1.08)	3.67 (0.86)	3.00 (1.22)	3.07 (1.39)	3.25 (0.89)	2.81 (1.12)	2.22 (1.17)	3.00 (1.10)	2.60	0.01	0.12
Reading some text in wikis (internet tool) (WORKED)	2.85 (0.99)	2.48 (0.93)	2.83 (1.01)	2.36 (1.55)	2.75 (1.18)	2.24 (1.18)	1.61 (1.04)	2.50 (1.03)	2.67	0.01	0.13
Reading some text in wikis (internet tool) (EVALUATE)	2.77 (0.83)	2.38 (1.02)	2.83 (1.01)	2.29 (1.54)	2.75 (1.18)	2.24 (1.18)	1.56 (1.04)	2.38 (0.89)	2.91	0.01	0.14
Reading some text in wikis (internet tool) (UTILITY)	3.00 (1.16)	2.86 (1.01)	3.00 (0.98)	2.86 (1.51)	3.00 (1.02)	2.24 (1.14)	2.11 (1.18)	2.69 (1.20)	2.32	0.02	0.11
Readings are preferably done in digital format (Word, PDF, eBook, ePub) (WORKED)	3.46 (0.78)	3.43 (1.03)	3.83 (0.87)	3.57 (0.76)	3.82 (0.95)	4.05 (0.59)	3.44 (1.20)	3.37 (0.89)	2.00	0.05	0.10
Readings are preferably done in digital format (Word, PDF, eBook, ePub) (EVALUATE)	3.46 (0.78)	3.38 (1.12)	3.83 (0.87)	3.50 (0.86)	3.79 (0.969)	4.00 (0.63)	3.33 (1.28)	3.37 (0.89)	2.86	0.01	0.13

Only statistically significant variables are included.

TABLE 3 | Tests of the inter-subject effects of the students in the instructional variables in ARC considering the department as a grouping variable.

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Please indicate your level of knowledge ... (min 1–max 5)											
of the SaberPro tests on critical reading	3.14 (0.74)	3.23 (0.89)	3.14 (0.83)	3.17 (0.74)	3.31 (0.79)	3.10 (0.68)	3.22 (0.82)	3.41 (0.78)	2.60	0.00	0.01
of the evaluation of critical reading in the subject	3.24 (0.70)	3.41 (0.83)	3.21 (0.82)	3.32 (0.74)	3.43 (0.83)	3.36 (0.65)	3.39 (0.79)	3.58 (0.81)	4.27	0.00	0.02
of the evaluation of textual construction in the subject	3.15 (0.74)	3.34 (0.83)	3.24 (0.82)	3.32 (0.78)	3.44 (0.76)	3.23 (0.73)	3.36 (0.80)	3.60 (0.78)	5.49	0.00	0.02
of the teaching of critical reading in the subject	3.25 (0.78)	3.48 (0.80)	3.26 (0.82)	3.34 (0.82)	3.52 (0.80)	3.37 (0.73)	3.46 (0.79)	3.65 (0.84)	5.07	0.00	0.02
of the transversal teaching of generic competences of critical reading and textual construction in the subject	3.15 (0.81)	3.32 (0.83)	3.26 (0.81)	3.28 (0.77)	3.36 (0.83)	3.20 (0.75)	3.30 (0.80)	3.51 (0.78)	3.32	0.00	0.01
Indicate your degree of interest... (min 1–max 5)											
for critical reading to work on key aspects of the subject	3.45 (0.85)	3.77 (0.80)	3.53 (0.83)	3.61 (0.81)	3.80 (0.84)	3.48 (0.78)	3.64 (0.78)	3.97 (0.77)	7.26	0.00	0.03
for critical reading to evaluate on key aspects of the subject	3.48 (0.86)	3.73 (0.83)	3.59 (0.76)	3.58 (0.80)	3.80 (0.78)	3.47 (0.73)	3.60 (0.78)	3.92 (0.79)	5.90	0.00	0.02
Advanced reading comprehension-has ever been used in this subject tasks, techniques or teaching strategies (or instructional) of... (min 1–max 5)											
Strategies and processes											
Extract the main ideas of some reading (WORKED)	3.53 (1.02)	3.90 (0.90)	3.66 (0.95)	3.65 (0.94)	3.89 (0.95)	3.72 (0.97)	3.83 (0.92)	4.16 (0.85)	9.05	0.00	0.04
Extract the main ideas of some reading (EVALUATE)	3.47 (1.07)	3.47 (1.07)	3.86 (0.91)	3.59 (0.95)	3.90 (0.91)	3.70 (0.93)	3.82 (0.91)	4.16 (0.83)	10.97	0.00	0.04
Extract the main ideas of some reading (UTILITY)	3.65 (0.99)	3.90 (0.84)	3.61 (0.99)	3.70 (0.93)	3.99 (0.87)	3.64 (0.96)	3.80 (0.92)	4.23 (0.83)	10.00	0.00	0.04
Relate ideas from some reading with previous knowledge (WORKED)	3.59 (0.96)	3.92 (0.82)	3.62 (0.88)	3.74 (0.88)	3.98 (0.85)	3.64 (0.95)	3.80 (0.87)	4.20 (0.78)	9.51	0.00	0.04
Relate ideas from some reading with previous knowledge (EVALUATE)	3.53 (0.99)	3.91 (0.82)	3.54 (0.92)	3.67 (0.88)	3.97 (0.85)	3.58 (0.95)	3.80 (0.89)	4.13 (0.82)	10.14	0.00	0.04
Relate ideas from some reading with previous knowledge (UTILITY)	3.67 (0.97)	3.91 (0.77)	3.57 (0.96)	3.71 (0.87)	3.94 (0.89)	3.70 (0.88)	3.82 (0.89)	4.18 (0.80)	9.09	0.00	0.04
Draw conclusions and inferences not explicit in the readings (WORKED)	3.47 (1.00)	3.73 (0.88)	3.48 (1.05)	3.59 (0.90)	3.81 (0.86)	3.58 (0.94)	3.70 (0.92)	4.08 (0.82)	8.62	0.00	0.03
Draw conclusions and inferences not explicit in the readings (EVALUATE)	3.44 (1.00)	3.73 (0.86)	3.49 (0.99)	3.59 (0.94)	3.80 (0.87)	3.57 (0.95)	3.73 (0.93)	4.10 (0.80)	9.35	0.00	0.04

(Continued)

TABLE 3 | Continued

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Draw conclusions and inferences not explicit in the readings (UTILITY)	3.56 (1.00)	3.73 (0.85)	3.57 (1.01)	3.62 (0.92)	3.86 (0.87)	3.50 (0.94)	3.72 (0.91)	4.08 (0.83)	7.92	0.00	0.03
Apply solutions to other aspects, suggested by the readings (WORKED)	3.39 (0.99)	3.70 (0.88)	3.50 (0.99)	3.56 (0.90)	3.83 (0.83)	3.55 (0.93)	3.69 (0.90)	4.03 (0.81)	8.54	0.00	0.03
Apply solutions to other aspects, suggested by the readings (EVALUATE)	3.41 (1.02)	3.68 (0.90)	3.42 (0.98)	3.55 (0.92)	3.82 (0.84)	3.54 (0.96)	3.69 (0.89)	4.01 (0.84)	8.45	0.00	0.03
Apply solutions to other aspects, suggested by the readings (UTILITY)	3.50 (0.94)	3.72 (0.86)	3.53 (0.92)	3.58 (0.88)	3.90 (0.80)	3.53 (0.91)	3.65 (0.87)	4.03 (0.77)	8.71	0.00	0.03
Explain the basic content of some reading (WORKED)	3.53 (0.98)	3.86 (0.85)	3.62 (0.96)	3.66 (0.86)	3.97 (0.82)	3.62 (0.94)	3.83 (0.87)	4.22 (0.74)	11.99	0.00	0.05
Explain the basic content of some reading (EVALUATE)	3.55 (1.04)	3.82 (0.87)	3.51 (0.97)	3.63 (0.88)	3.98 (0.79)	3.58 (0.95)	3.75 (0.89)	4.18 (0.78)	10.80	0.00	0.04
Explain the basic content of some reading (UTILITY)	3.65 (0.99)	3.84 (0.80)	3.59 (0.94)	3.65 (0.88)	4.05 (0.81)	3.69 (0.94)	3.81 (0.88)	4.18 (0.88)	11.28	0.00	0.04
Remember without consulting the information previously read (retrieve). To consult again later and to recover without consulting to go completing what is not remembered (WORKED)	3.34 (1.01)	3.58 (0.87)	3.40 (0.95)	3.52 (0.90)	3.73 (0.86)	3.55 (0.98)	3.60 (0.89)	3.85 (0.83)	5.05	0.00	0.02
Remember without consulting the information previously read (retrieve). To consult again later and to recover without consulting to go completing what is not remembered (EVALUATE)	3.34 (1.03)	3.60 (0.87)	3.41 (0.99)	3.51 (0.91)	3.72 (0.89)	3.57 (0.97)	3.59 (0.88)	3.85 (0.85)	5.26	0.00	0.02
Remember without consulting the information previously read (retrieve). To consult again later and to recover without consulting to go completing what is not remembered (UTILITY)	3.44 (0.97)	3.60 (0.85)	3.48 (0.97)	3.55 (0.85)	3.73 (0.87)	3.39 (1.00)	3.59 (0.82)	3.90 (0.79)	5.47	0.00	0.02
Textual genres											
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (WORKED)	3.36 (1.05)	3.71 (0.93)	3.43 (1.04)	3.53 (1.01)	3.93 (0.92)	3.54 (1.07)	3.63 (0.94)	4.12 (0.80)	11.86	0.00	0.05
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (EVALUATE)	3.33 (1.07)	3.69 (0.95)	3.43 (1.01)	3.51 (1.01)	3.92 (0.90)	3.46 (1.08)	3.58 (0.99)	4.08 (0.819)	11.16	0.00	0.04
Reading of some argumentative text (defend ideas, debate, refute, convince, justify) (UTILITY)	3.39 (1.10)	3.74 (0.87)	3.48 (1.01)	3.58 (0.95)	3.94 (0.83)	3.52 (1.06)	3.61 (0.92)	4.14 (0.80)	11.84	0.00	0.05

(Continued)

TABLE 3 | Continued

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Reading of some comparison and contrast text : two theories, concepts, stories, authors, figures, preferences (WORKED)	3.30 (1.04)	3.64 (0.94)	3.37 (0.95)	3.50 (0.90)	3.81 (0.96)	3.51 (1.00)	3.63 (0.90)	4.06 (0.84)	11.47	0.00	0.04
Reading of some comparison and contrast text : two theories, concepts, stories, authors, figures, preferences (EVALUATE)	3.31 (1.02)	3.64 (0.94)	3.33 (0.95)	3.49 (0.97)	3.81 (0.95)	3.53 (1.02)	3.62 (0.93)	4.04 (0.78)	11.47	0.00	0.04
Reading of some comparison and contrast text : two theories, concepts, stories, authors, figures, preferences (UTILITY)	3.43 (1.06)	3.71 (0.87)	3.51 (0.90)	3.54 (0.90)	3.87 (0.86)	3.50 (0.98)	3.61 (0.88)	4.09 (0.80)	11.41	0.00	0.04
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (WORKED)	3.36 (1.07)	3.67 (0.91)	3.40 (0.97)	3.53 (0.94)	3.83 (0.84)	3.48 (0.99)	3.57 (0.91)	3.93 (0.84)	6.76	0.00	0.03
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (EVALUATE)	3.38 (1.04)	3.65 (0.93)	3.39 (0.96)	3.50 (0.95)	3.78 (0.93)	3.44 (1.04)	3.63 (0.89)	3.88 (0.84)	6.66	0.00	0.03
Reading of an essay (question-answer, admission, answer, scientific, test answers, opinion) (UTILITY)	3.45 (1.01)	3.69 (0.87)	3.45 (0.92)	3.59 (0.88)	3.89 (0.81)	3.50 (0.99)	3.61 (0.86)	4.02 (0.81)	8.82	0.00	0.03
Reading of some analysis text (description, literary analysis, process analysis) (WORKED)	3.42 (1.07)	3.68 (0.90)	3.41 (0.89)	3.55 (0.89)	3.83 (0.86)	3.56 (1.00)	3.64 (0.87)	4.00 (0.84)	8.35	0.00	0.03
Reading of some analysis text (description, literary analysis, process analysis) (EVALUATE)	3.38 (1.05)	3.68 (0.93)	3.40 (0.95)	3.50 (0.92)	3.83 (0.87)	3.53 (1.03)	3.64 (0.91)	4.01 (0.82)	9.52	0.00	0.04
Reading of some analysis text (description, literary analysis, process analysis) (UTILITY)	3.44 (1.03)	3.71 (0.84)	3.51 (0.92)	3.54 (0.88)	3.86 (0.89)	3.45 (1.00)	3.61 (0.85)	4.02 (0.80)	9.92	0.00	0.04
Reading a bibliographic review (WORKED)	3.21 (1.08)	3.50 (1.04)	3.36 (0.98)	3.43 (0.99)	3.62 (0.98)	3.43 (1.03)	3.49 (1.01)	3.89 (0.89)	6.31	0.00	0.03
Reading a bibliographic review (EVALUATE)	3.16 (1.14)	3.52 (1.03)	3.30 (1.01)	3.43 (1.01)	3.70 (0.96)	3.47 (1.03)	3.49 (1.04)	3.90 (0.87)	6.91	0.00	0.03
Reading a bibliographic review (UTILITY)	3.32 (1.05)	3.55 (1.00)	3.47 (0.99)	3.46 (0.96)	3.70 (0.90)	3.43 (1.04)	3.56 (0.96)	3.97 (0.80)	7.69	0.00	0.03
Reading some cause-effect/problem-solution text (WORKED)	3.36 (0.87)	3.62 (0.94)	3.40 (1.02)	3.50 (0.94)	3.68 (0.90)	3.58 (0.93)	3.65 (0.93)	3.82 (0.90)	4.67	0.00	0.02

(Continued)

TABLE 3 | Continued

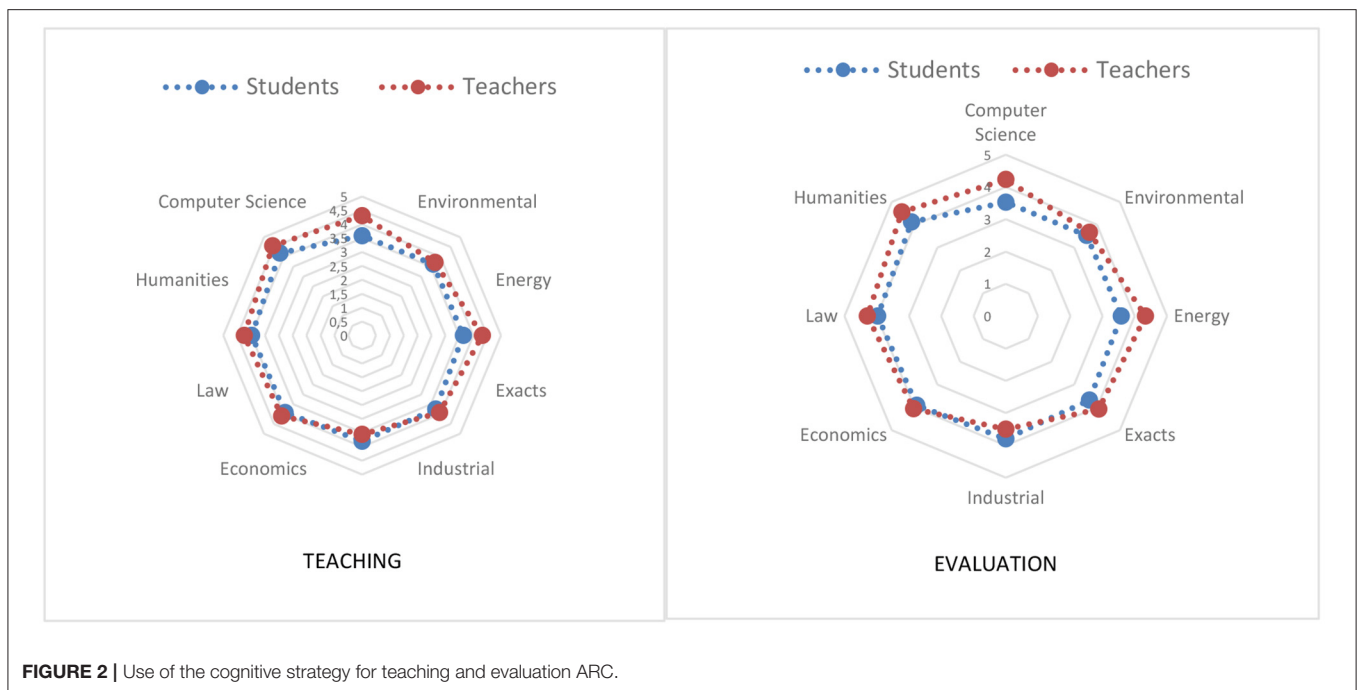
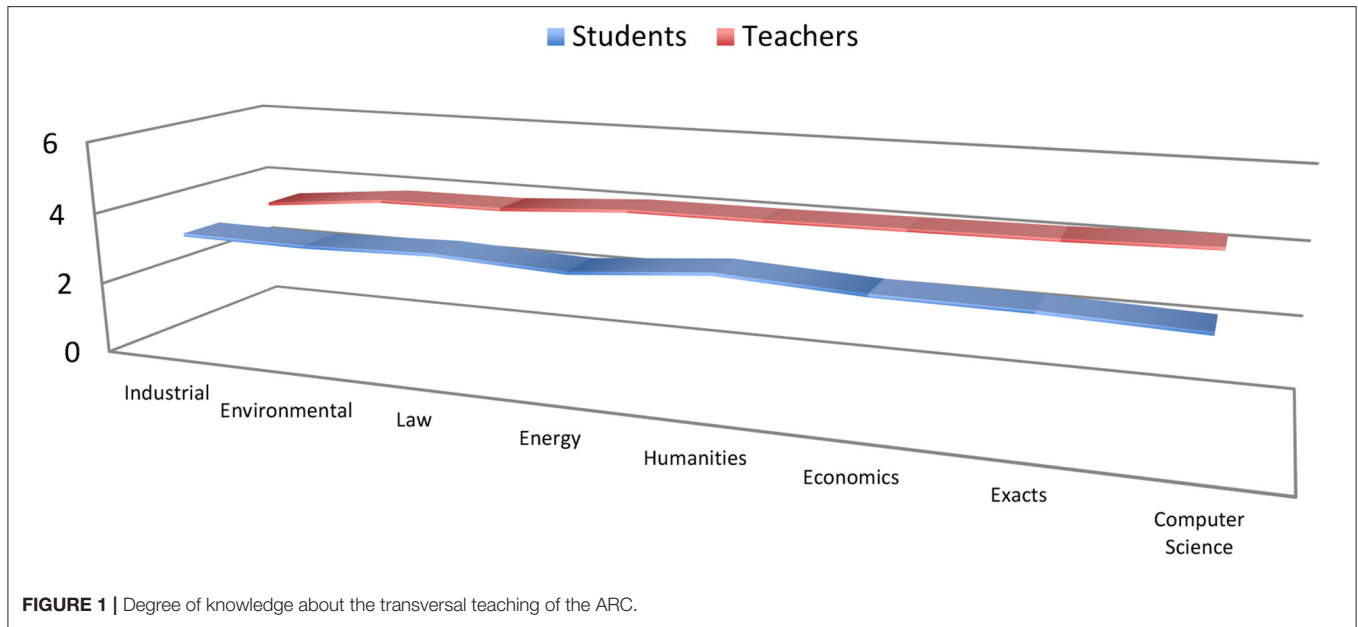
Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Reading some cause-effect/problem-solution text (EVALUATE)	3.38 (1.12)	3.62 (0.94)	3.34 (1.02)	3.47 (0.97)	3.72 (0.91)	3.62 (0.97)	3.66 (0.92)	3.82 (0.90)	5.81	0.00	0.02
Reading some cause-effect/problem-solution text (UTILITY)	3.47 (1.04)	3.64 (0.88)	3.47 (0.88)	3.53 (0.91)	3.80 (0.88)	3.55 (0.98)	3.66 (0.92)	3.91 (0.83)	5.50	0.00	0.02
Reading some definition text (WORKED)	3.46 (1.03)	3.74 (0.85)	3.49 (0.91)	3.60 (0.86)	3.86 (0.89)	3.63 (1.03)	3.73 (0.94)	3.97 (0.94)	6.65	0.00	0.03
Reading some definition text (EVALUATE)	3.44 (1.08)	3.74 (0.87)	3.51 (0.90)	3.59 (0.87)	3.80 (0.93)	3.57 (1.04)	3.70 (0.93)	3.94 (0.84)	6.40	0.00	0.03
Reading some definition text (UTILITY)	3.49 (1.06)	3.74 (0.83)	3.59 (0.84)	3.63 (0.86)	3.82 (0.91)	3.51 (1.03)	3.73 (0.90)	3.97 (0.79)	6.06	0.00	0.02
Reading of a text to review the state of the question (topic, theory, approach, scientific, antecedents, previous experiences, successful solutions) (WORKED)	3.37 (1.01)	3.61 (0.89)	3.42 (0.86)	3.52 (0.91)	3.72 (0.91)	3.58 (1.01)	3.57 (0.98)	3.97 (0.80)	6.89	0.00	0.03
Reading of a text to review the state of the question (topic, theory, approach, scientific, antecedents, previous experiences, successful solutions) (EVALUATE)	3.42 (1.04)	3.61 (0.92)	3.43 (0.88)	3.52 (0.89)	3.72 (0.90)	3.55 (0.98)	3.60 (0.98)	4.00 (0.81)	7.40	0.00	0.03
Reading of a text to review the state of the question (topic, theory, approach, scientific, antecedents, previous experiences, successful solutions) (UTILITY)	3.48 (1.02)	3.66 (0.85)	3.59 (0.80)	3.54 (0.89)	3.83 (0.89)	3.53 (0.99)	3.61 (0.89)	4.02 (0.79)	8.75	0.00	0.03
Medium used											
Reading a document in digital databases (WORKED)	3.64 (0.96)	3.64 (0.96)	3.38 (1.04)	3.50 (0.98)	3.54 (0.97)	3.70 (1.04)	3.65 (0.92)	3.93 (0.90)	4.91	0.00	0.02
Reading a document in digital databases (EVALUATE)	3.53 (1.00)	3.63 (0.99)	3.32 (0.98)	3.52 (0.97)	3.57 (1.00)	3.70 (1.06)	3.62 (0.98)	3.96 (0.86)	5.25	0.00	0.02
Reading a document in digital databases (UTILITY)	3.71 (0.94)	3.69 (0.98)	3.45 (0.94)	3.61 (0.92)	3.69 (0.87)	3.65 (0.97)	3.64 (0.92)	4.01 (0.79)	4.94	0.00	0.02

(Continued)

TABLE 3 | Continued

Department	Computer Science	Economics	Exacts	Environmental	Law	Energy	Industrial	Humanities	F	p	η^2
Readings are preferably done in traditional format (paper, notes, articles, books) (WORKED)	3.29 (1.01)	3.52 (0.98)	3.58 (0.92)	3.52 (0.94)	3.80 (0.89)	3.71 (0.87)	3.65 (0.99)	3.78 (0.96)	4.23	0.00	0.02
Readings are preferably done in traditional format (paper, notes, articles, books) (EVALUATE)	3.30 (1.08)	3.56 (0.99)	3.52 (0.94)	3.52 (0.94)	3.74 (0.95)	3.74 (0.87)	3.60 (0.99)	3.84 (0.92)	4.49	0.00	0.02
Readings are preferably done in traditional format (paper, notes, articles, books) (UTILITY)	3.43 (1.07)	3.59 (0.95)	3.58 (0.93)	3.63 (0.86)	3.79 (0.86)	3.64 (0.86)	3.58 (0.87)	3.89 (0.90)	3.70	0.00	0.02
Readings are preferably done in digital format (Word, PDF, eBook, ePub) (WORKED)	3.67 (0.94)	3.81 (0.90)	3.54 (0.95)	3.68 (0.88)	3.80 (0.94)	3.72 (0.93)	3.84 (0.88)	4.04 (0.86)	4.41	0.00	0.02
Reading some text in blogs (internet tool) (WORKED)	3.64 (0.94)	3.80 (0.88)	3.41 (1.00)	3.63 (0.90)	3.76 (0.92)	3.72 (0.91)	3.77 (0.87)	4.04 (0.78)	5.64	0.00	0.02
Reading some text in blogs (internet tool) (EVALUATE)	3.72 (0.90)	3.80 (0.87)	3.65 (0.80)	3.69 (0.88)	3.84 (0.90)	3.74 (0.89)	3.81 (0.86)	4.06 (0.78)	4.36	0.00	0.02
Audiobooks or video documents are used (WORKED)	3.20 (1.19)	3.19 (1.20)	2.95 (1.19)	3.17 (1.16)	3.12 (1.16)	3.37 (1.12)	3.18 (1.22)	3.46 (1.17)	2.15	0.02	0.01
Audiobooks or video documents are used (EVALUATE)	3.16 (1.24)	3.17 (1.18)	2.96 (1.23)	3.15 (1.15)	3.13 (1.19)	3.37 (1.13)	3.23 (1.21)	3.49 (1.17)	2.53	0.00	0.01

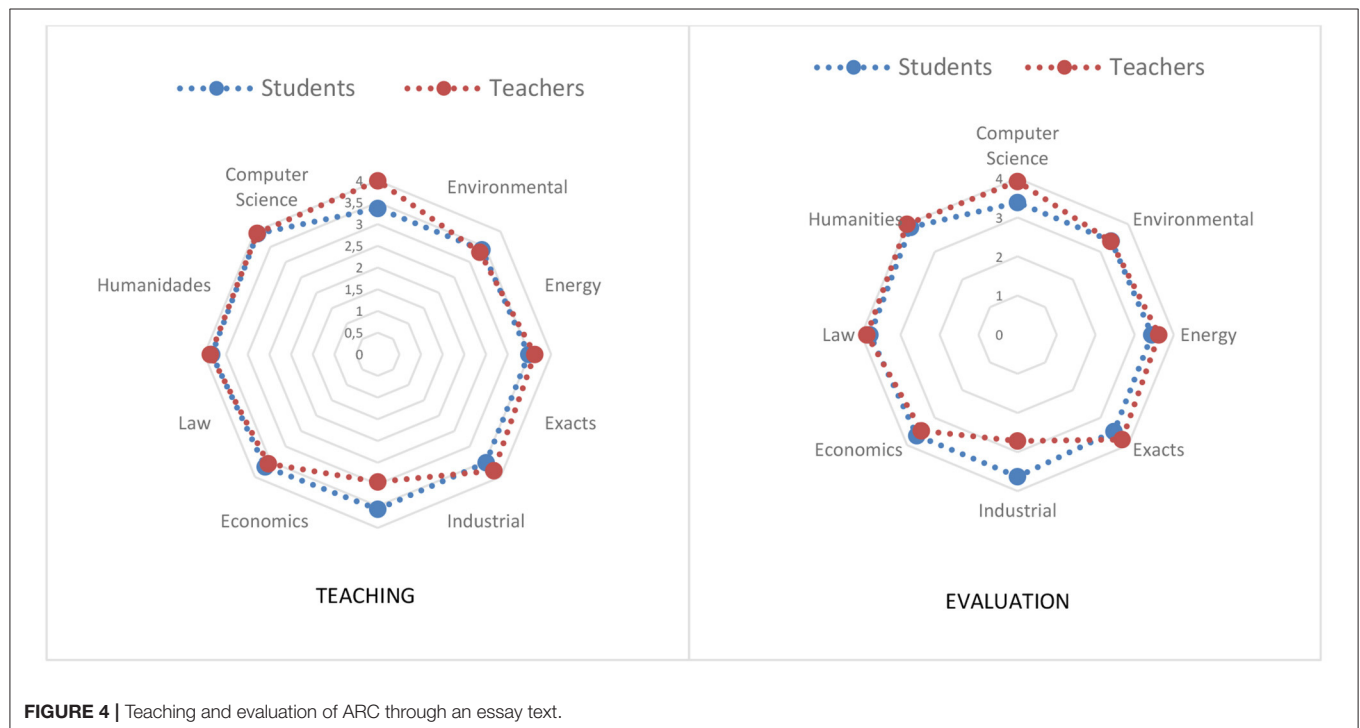
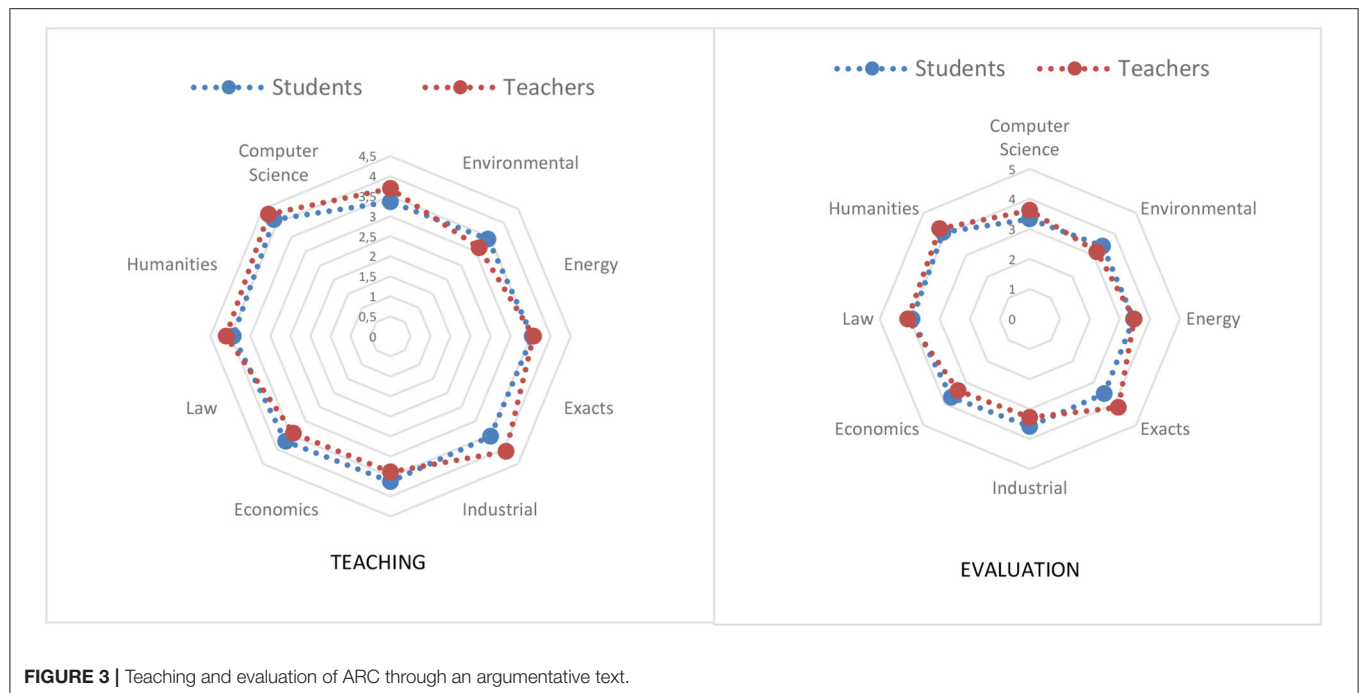
Only statistically significant variables are included.



another clear trend is evident by department [Environmental, Industrial, Economics, Energy, Computer Science, Law and/or Exact (these last two are the only departments whose order is altered) and Humanities].

As shown in **Figure 4**, the trend evidenced in the argumentative texts is maintained in the essays, that is, the teachers show higher scores than the students, except for the Environmental, Industrial and Economic departments. In addition, when examining both their

use when teaching and cross-evaluating the ARC, in the case of students a trend of greater use is observed by department (Computer Science, Environmental, Energy, Exact, Industrial, Economics, Law and Humanities) and of the in the same way, in the case of teaching staff, there is another trend by department [Industrial, Natural Sciences, Economics, Energy, Civil, Law, Humanities and/or Informatics (these last two are the only departments whose order is altered)].



DISCUSSION AND CONCLUSIONS

This reaches the initial purpose of analyzing the cognitive strategies and the textual genres used in teaching and in the transversal evaluation of critical reading or advanced reading comprehension, in a Latin American university, on the Coast, from a double perspective: that of the teaching staff and that of the students. Likewise, the hypotheses are confirmed, on the

one hand, in relation to the first, it is evidenced that teachers show greater knowledge of cognitive strategies and textual genres to teach and evaluate ARC than students, regardless of the department of assignment. In the same way, the second hypothesis is confirmed since differential patterns are observed in the use of cognitive strategies and textual genres to teach and evaluate ARC by teachers and students, depending on the department to which they are assigned. In this sense, it is not

surprising that it is the professors assigned to the Humanities department that show a greater use of cognitive strategies and textual genres for teaching and cross-sectional evaluation of the ARC and that it is those of Computer Science that exhibit the least.

However, it is true that the extrapolation of the results obtained to the entire Colombian higher education territory is questionable, given that the data collected comes from the same higher education institution. That is why the replication of this study in other Colombian universities is recommended in order to verify or refute the trends in the use of cognitive strategies and textual typologies used both to teach and to cross-evaluate ARC. This approach is in line with the most recent advances in the field of literacy across content/discipline/curriculum and the need for its empowerment and deployment in the university (Miller et al., 2018; Scott and Washburn, 2018; Van Ockenburg et al., 2019). The peculiar characteristics of the Latin American university refer to the need to advance in the knowledge of the cognitive strategies and the textual genres used for the teaching and evaluation of the generic competence of ARC, in the different subjects of the different careers and studies. The identification of these, as well as their adaptation and use in the different fields of knowledge, seems mandatory, for the improvement of educational quality and learning results, both in generic competences and in the rest of the subjects. In this sense, there are several emerging research lines that require more evidence, such as the specificity in each professional field, the key personal and institutional variables, as well as other desirable measures that should be implemented in the future, including the deployment of studies with other methodologies different from self-report, such as the observation or contribution of evidence and learning results and experiences validated empirically or scientifically (Valero et al., 2015; Uttl et al., 2017; Zhao and Zhang, 2017; Graham et al., 2018; CEEDAR, 2021; USOE, 2021).

It seems reasonable to include various performances in the future. On the one hand, the implementation of longitudinal studies and the inclusion of digital literacy seem obligatory, which would allow us to know more precisely the nature and role of generic competences, specifically advanced reading comprehension, as well as its impact on learning, rest of psychoeducational competencies and variables (García-Martín and García-Martín, 2021; Robledo-Ramón and García-Gutiérrez, 2021). In addition, the need for exploratory studies on the training that teachers and students, from different fields and careers, must receive and promote to implement the mainstreaming of advanced reading comprehension, such as online workshops and webinars (Hatlevik and Hatlevik, 2018; Daumiller et al., 2019; He et al., 2020). It is also evident that the development of strategies, techniques and instructional procedures for cross-sectional evaluation and teaching, as well as virtual ones, must be contemplated, both for students and teachers (López et al., 2021), the construction of systems of evaluation and instruction, among others (Utama et al., 2020).

Another action could involve the design of forms on advanced reading comprehension strategies, such as the case of checklists or others (Mohamadi, 2018; Yunusa and Umar, 2021). Finally, the implementation of online records and protocols for the instruction and cross-sectional evaluation of advanced reading comprehension and other generic and specific competences, as has been done with the citizen with lateral reading and checking fakes while reading information from various subjects and fields (Brodsky et al., 2021; Cabrera et al., 2021; Fandiño-Parra et al., 2021).

Quality Assessment

In short, it is therefore a study that allows optimizing this transversal competence in students that, in turn, favors the performance of these in the *Saber Pro* tests of the reading module, to which everyone who wants to opt for a degree in Colombian Higher Education, in accordance with the guidelines of the Colombian Institute for the Evaluation of Education (Calderón et al., 2021) and thus reducing the chances of university student desertion (CUC, 2018).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité De Ética De La Universidad De La Costa. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

J-NG-S and JG-M: implemented the study. J-NG-S and JG-M conducted the statistical analyses and drafted the manuscript. Both authors performed substantial contributions to the conception of the study, reviewed the manuscript critically for relevant intellectual content, and approved the submitted version. All authors contributed to the article and approved the submitted version.

FUNDING

Competitive Research Project evidence CONV-ÍNDEx Ref. 13-2018, Code of Project INV. 150-01-007-13. Assessed by the ACAC (MEN Colombia).

ACKNOWLEDGMENTS

Universidad de la Costa (Colombia)—Universidad de León (Spain)—Universidad de Salamanca (Spain).

REFERENCES

- Arias, D. H., Díaz, O. C., Garzón, I., León, A. C., Rodríguez, S. P., and Valbuena, E. O. (2018). *Between the Demands of Quality and the Conditions of Inequality: Initial Teacher Training in Colombia*. Bogotá: UPN. Available online at: <http://repositorio.pedagogica.edu.co/handle/2~0.500.12209/7826>
- Barnet, R. (2001). *The Limits of Competition. Knowledge, Higher Education and Society*. Barcelona: Gedisa.
- Brodsky, J. E., Brooks, P. J., Scimeca, D., Todorova, R., Galati, P., Batson, M., et al. (2021). Improving college students' fact-checking strategies through lateral reading instruction in a general education civics course. *Cogn. Res.* 6:23. doi: 10.1186/s41235-021-00291-4
- Cabeza, L., Lombana, J., and Castrillón, J. (2020). External factors in the performance of generic state tests (Saber Pro) of English in business and related students in Colombia. *Rev. Iberoam. Educ. Super.* 11, 182–200. doi: 10.22201/iisue.20072872e.2020.30.595
- Cabrera, L., Sokolow, J., and Cabrera, D. (2021). "Developing and validating a measurement of systems thinking: the systems thinking and metacognitive inventory (STMI)," in *Routledge Handbook of Systems Thinking*, eds D. Cabrera, L. Cabrera, and G. Midgley (London: Routledge), 1–35.
- Calderón, A., Parra, C. A., and Piñeros, M. A. (2021). *Saber Pro Orientation Guide: Generic Skills Module*. Bogotá: Instituto Colombiano para la Evaluación de la Educación (ICFES). Available online at: <https://www.icfes.gov.co/documents/20143/2297330/Guia+de+orientacion+modulos+genericos+Saber-Pro+2021.pdf>
- Cassany, D. (2003). Approaches to critical reading: theory, examples and reflections. *Rev. Electron. Investig. Innov. Educ.* 1, 113–132.
- Castellar, A., Villadiego, D., Gamero, H., and Gamarra, J. (2021). Academic Support Plan: incidence in the Development of Generic Competences in University Students. *Revista de Ciencias Sociales*. Available online at: <https://produccioncientificaluz.org/index.php/rcs/index>
- CEEDAR (2021). *Evidences Based Practices. CEEDAR Innovation Configurations. Tools for Programs Reforms*. Available online at: <http://ceedar.education.ufl.edu/tools/innovation-configuration-ns/>
- Chen, Y. H., and Lin, Y. J. (2018). Validation of the short self-regulation questionnaire for taiwanese college students (TSSRQ). *Front. Psychol.* 9:259. doi: 10.3389/fpsyg.2018.00259
- Cubides, C., Rojas, M., and Cárdenas, R. (2017). Critical reading. Definitions, experiences and possibilities. *Saber Cienc.* 12, 184–197. doi: 10.18041/2382-3240/saber.2017v12n2.1586
- CUC (2018). *High Quality Accreditation Self Report*. Barranquilla: CUC
- Cuesta, M., Suárez-Álvarez, J., Lozano, L. M., García-Cueto, E., and Muñoz, J. (2018). Assessment of eight entrepreneurial personality dimensions: validity evidence of the bepe battery. *Front. Psychol.* 9:2352. doi: 10.3389/fpsyg.2018.02352
- Daumiller, M., Dickhäuser, O., and Dresel, M. (2019). University instructors' achievement goals for teaching. *J. Educ. Psychol.* 111, 131–148. doi: 10.1037/edu0000271
- Del Moral-Pérez, M. E., Villalustre, L., and Neira, M. R. (2016). Digital stories: activating narrative and digital communication skills in initial teacher training. *Ocnos* 15, 22–41. doi: 10.18239/ocnos_2016.15.1.923
- Díaz, J. P., Bar, A., and Ortiz, M. (2015). Critical reading and its relationship with the disciplinary training of university students. *Rev. Educ. Super.* 44, 139–159. doi: 10.1016/j.resu.2015.12.006
- Fandiño-Parra, Y. J., Muñoz-Barriga, A., López-Díaz, R. A., and Galindo-Cuesta, J. A. (2021). Teacher education and critical thinking: systematizing theoretical perspectives and formative experiences in Latin America. *Rev. de Investig. Educ.* 39, 149–167. doi: 10.6018/rie.416271
- García, M., and Jiménez, M. F. (2017). "Critical reading of scientific articles," in *Investigación en enfermería. Teoría y práctica*, eds M. Lleixà Fortuño, and P. Montesó (Tarragona: Publicacions URV), 159–172.
- García, J. N., Conde, M., Inciarte, A., Sánchez, E., Marín, F., and García-Martín, J. (2019). Review of international studies on evaluation and methodologies university professors. *Int. J. Dev. Educ. Psychol.* 31, 273–282. doi: 10.17060/ijodaep.2019.n1.v3.1489
- García, J. N., and García-Martín, J. (2021a). "From learning to write to writing to learn," in *Aflar la Mente Como herramienta del Pensar*, eds A. Dosil (Vigo: Publicac Universidad), 252–270.
- García, J. N., and García-Martín, J. (2021b). Reading comprehension across the curriculum: instructional and psychoeducational variables. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* 19 (in press).
- García, J. R., Sánchez, P. A., Orozco, M., and Obredor, S. (2019). Knowledge extraction for the prediction and analysis of the results of the quality test of higher education in Colombia. *Form. Univ.* 12, 55–62. doi: 10.4067/S0718-50062019000400055
- García-Martín, J., García, J. N., Inciarte, A. J., Marín, F. V., Sánchez, E. R., and Conde, M. E. (2019a). Teacher evaluation systems and methodologies and their impact on generic competences (EMICOG-PROFESORS). *Int. J. Dev. Educ. Psychol.* 31, 325–336. doi: 10.17060/ijodaep.2019.n1.v3.1498
- García-Martín, J., García, J. N., Inciarte, A. J., Marín, F. V., Sánchez, E. R., and Conde, M. E. (2019b). Teacher evaluation systems and methodologies and their impact on generic competences (EMICOG-STUDENTS). *Int. J. Dev. Educ. Psychol.* 31, 337–352. doi: 10.17060/ijodaep.2019.n1.v3.1499
- García-Martín, J., and García-Martín, S. (2021). Uso de herramientas digitales para la docencia durante la pandemia COVID-19. *Rev. Española Educ. Comparada* 38, 151–173. doi: 10.5944/rec.38.2021.27816
- García-Martín, J., and García-Sánchez, J. N. (2020). Methodologies used by University lecturers in teaching and assessment of reading competence. *Ocnos* 19, 55–71. doi: 10.18239/ocnos_2020.19.3.2378
- Graham, S. (2021). Through the looking glass: reflections of a writing scholar. *Acquired wisdom series. Educ. Rev.* 28, 1–21. doi: 10.14507/er.v2.8.3209
- Graham, S., Liu, X., Ng, C., Bartlett, B., Harris, K. R., and Kavanaug, C. (2018). Effectiveness of literacy programs balancing reading and writing instruction: a meta-analysis. *Read. Res. Q.* 53, 279–304. doi: 10.1002/rq.194
- Habók, A., and Magyar, A. (2018). Validation of a self-regulated foreign language learning strategy questionnaire through multidimensional modelling. *Front. Psychol.* 9:1388. doi: 10.3389/fpsyg.2018.01388
- Hatlevik, I. K. R., and Hatlevik, O. E. (2018). Examining the relationship between teachers' ICT self-efficacy for educational purposes, collegia collaboration, lack of facilitation and the use of ICT in teaching practice. *Front. Psychol.* 9:935. doi: 10.3389/fpsyg.2018.00935
- He, P., Zhou, Q., Zhao, H., Jiang, C., and Wu, Y. J. (2020). Compulsory citizenship behavior and employee creativity: creative self-efficacy as a mediator and negative affect as a moderator. *Front. Psychol.* 11:1640. doi: 10.3389/fpsyg.2020.01640
- Inciarte, A. J., García, J. N., Sánchez, E. R., Marín, F. V., Conde, M. E., Castillo, J. S., et al. (2019). Impact of evaluation and teacher methodologies in the development of generic competences. *Int. J. Dev. Educ. Psychol.* 31, 221–230. doi: 10.17060/ijodaep.2019.n1.v3.1475
- Jiménez, D., González, J. J., and Tornel, M. (2020). Active methodologies in the university and their relationship with teaching approaches. *Profesorado. Revista de Currículum y Formación de Profesorado* 24, 76–94.
- Li, J., Ye, H., Tang, Y., Zhou, Z. y Hu, X. (2018). What are the effects of self-regulation phases and strategies for Chinese students? A meta-analysis of two decades research of the association between self-regulation and academic performance. *Front. Psychol.* 9, 1–13. doi: 10.3389/fpsyg.2018.02434
- López, P., Torrance, M., Rijlaarsdam, G., and Fidalgo, R. (2021). Evaluating effects of different forms of revision instruction in upper-primary students. *Read. Writ.* 34, 1741–1767. doi: 10.1007/s11145-021-10156-3
- Marín, F. V., García, J. N., Inciarte, A. J., Sánchez, E. R., Conde, M. E., and García-Martín, J. (2019). Teacher evaluation and methodologies and their impact on generic competences: theoretical perspectives. *Int. J. Dev. Educ. Psychol.* 31, 187–198. doi: 10.17060/ijodaep.2019.n1.v3.1469
- Martínez-Ferrer, B., Moreno, D., and Musitu, G. (2018). Are adolescents engaged in the problematic use of social networking sites more involved in peer aggression and victimization? *Front. Psychol.* 9:801. doi: 10.3389/fpsyg.2018.00801
- Miller, D. M., Scott, C. E., and McTigue, E. M. (2018). Writing in the secondary-level disciplines: a systematic review of context, cognition and content. *Educ. Psychol. Rev.* 30, 83–120. doi: 10.1007/s10648-016-9393-z
- Mohamadi, Z. (2018). The Effect of online summative and formative teacher assessment on teacher competences. *Asia Pac. Educ. Rev.* 20, 343–359. doi: 10.1007/s12564-018-9566-1

- OCDE (2016). *Education in Colombia*. MEC: OCDE Available online at: <https://www.oecd.org/education/school/Educacion-en-Colombia-Aspectos-Destacados.pdf>
- Orozco, L. (2010). Academic quality and social relevance of higher education in Latin America. *Rev. Iberoam. Educ. Super.* 1, 24–36. Available online at: http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S2007-28722010000100003&lng=es&nrm=iso&tlng=en
- Peng, P., Wang, T., Wang, C., and Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: effects of tasks, age, and social economics status. *Psychol. Bull.* 145, 189–236. doi: 10.1037/bul0000182
- Puente-Martínez, A., Páez, D., Ubillos-Landa, S., and Da Costa-Dutra, S. (2018). Examining the structure of negative affect regulation and its association with hedonic and psychological well-being. *Front. Psychol.* 9:1592. doi: 10.3389/fpsyg.2018.01592
- Reyes, C., Díaz, A., Pérez, R., Marchena, R., and Sosa, F. (2020). Learning assessment: perceptions and practices of University teachers on assessment. *Profesorado, Revista de currículum y formación del profesorado* 24, 136–162. doi: 10.30827/profesorado.v24i1.8449
- Robledo-Ramón, P., and García-Gutiérrez, V. (2021). Evolution of textual planning and revision strategies and their relationship with the written product. *Cult. y Educ.* 12, 9–26. doi: 10.17981/cultedusoc.12.2.021.01
- Romera, E. M., Herrera, M., Casas, J. A., Ortega, R., and del Rey, R. (2018). How much do adolescents cybergossip? Scale development and validation in Spain and Colombia. *Front. Psychol.* 9:126. doi: 10.3389/fpsyg.2018.00126
- Rueda, M., Fernández, N., García, P., Bakieva, M., González, J., Jornet, J. M., et al. (2019). Institutional practices and conditions for the development of teaching in Ibero-American universities. *Publicaciones* 49, 19–37. doi: 10.30827/publicaciones.v49i1.9850
- Samper, J., and Ospino, I. (2018). Citizen competences from the perspective of critical reading in University education. *Enfoque Latinoamericano* 1, 63–77. doi: 10.32012/26195399/rel12201833
- Sánchez, E. R., Conde, M. E., Marín, F. V., Inciarte, A. J., García, J. N., and García-Martín, J. (2019). Own successful antecedents-added value for higher education in Colombia. *Int. J. Dev. Educ. Psychol.* 31, 109–120. doi: 10.17060/ijodaep.2019.n1.v3.1455
- Scott, C. E., McTigue, E. M., Miller, D. M., and Washburn, E. K. (2018). The what, when, and how of preservice teachers and literacy across the disciplines: a systematic literature review of nearly 50 years of research. *Teach. Teach. Educ.* 73, 1–13. doi: 10.1016/j.tate.2018.03.010
- Sinval, J., Marques-Pinto, A., Queirós, C., and Maróco, J. (2018). Work engagement among rescue workers: psychometric properties of the Portuguese UWES. *Front. Psychol.* 8:2229. doi: 10.3389/fpsyg.2017.02229
- USOE (2021). *OSEP Ideas that Work. Office of Special Education Programs, EEUU: Department of Education*. Available online at: <https://www2.ed.gov/about/offices/list/osep/programs.html>
- Utama, C., Sajidan, Nurkamto, J., and Wiranto. (2020). The instrument development to measure higher-order thinking skills for pre-service biology teacher. *Int. J. Instr.* 13, 833–848. doi: 10.29333/iji.2020.13451a
- Uttl, B., White, C. A., and Wong, D. (2017). Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. *Stud. Educ. Eval.* 54, 22–42. doi: 10.1016/j.stueduc.2016.08.007
- Valero, M. J., Vázquez, B., and Cassany, D. (2015). Unraveling the web: the critical reading of foreign language learners in digital environments. *Ocnos* 13, 7–23. doi: 10.18239/ocnos_2015.13.01
- Van Ockenburg, L., Daphne van Weijen, D., and Rijlaarsdam, G. (2019). Learning to write synthesis texts: a review of intervention studies. *J. Writ. Res.* 10, 401–428. doi: 10.17239/jowr-2019.10.03.01
- Yunusa, A. A., and Umar, I. N. (2021). A scoping review of critical predictive factors (CPFs) of satisfaction and perceived learning outcomes in E-learning environments. *Educ. Inf. Technol.* 26, 1223–1270. doi: 10.1007/s10639-020-10286-1
- Zeng, G., Chen, X., Cheung, H. Y., and Peng, K. (2019). Teachers' growth mindset and work engagement in the Chinese educational context: well-being and perseverance of effort as mediators. *Front. Psychol.* 10:839. doi: 10.3389/fpsyg.2019.00839
- Zhao, H., and Zhang, X. (2017). The influence of field teaching practice on pre-service teachers' professional identity: a mixed methods study. *Front. Psychol.* 8:1264. doi: 10.3389/fpsyg.2017.01264

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 García-Sánchez and García-Martín. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



By Toutatis! Trainee Teachers' Motivation When Using Comics to Learn History

Juan Ramón Moreno-Vera^{1*}, Santiago Ponsoda-López de Atalaya² and Rubén Blanes-Mora²

¹Faculty of Education, University of Murcia, Murcia, Spain, ²Faculty of Education, University of Alicante, Alicante, Spain

OPEN ACCESS

Edited by:

Pedro García Guirao,
University of Ostrava, Czechia

Reviewed by:

Delfín Ortega-Sánchez,
University of Burgos, Spain
Alvaro Morote,
University of Valencia, Spain

*Correspondence:

Juan Ramón Moreno-Vera
jr.moreno@um.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 September 2021

Accepted: 01 October 2021

Published: 27 October 2021

Citation:

Moreno-Vera JR, Ponsoda-López de Atalaya S and Blanes-Mora R (2021)
By Toutatis! Trainee Teachers'
Motivation When Using Comics to
Learn History.
Front. Psychol. 12:778792.
doi: 10.3389/fpsyg.2021.778792

The main goal of this research was to analyse the perception of trainee primary education teachers regarding motivation when using comics as a resource to teach and learn history. To achieve this objective, a history education programme was designed based on the use of comics and the outcomes evaluated via a mixed qualitative-quantitative post-test questionnaire (Likert scale 1–5). Two hundred twenty-one trainee primary teachers from the University of Alicante, Spain participated in the study during the 2020/2021 academic year. Data were collected using the IBM SPSS v.24 statistical package and AQUAD 7. The results showed that the majority of future teachers felt highly motivated when using comic resources to learn history instead of textbooks (90.5% of participants); trainee teachers recognise that the use of comics improves their capacity to be more creative and that they feel able to design and use their own comic resources to teach history in the future.

Keywords: social sciences, teaching-learning, historical thinking, motivation, comic

INTRODUCTION

Nowadays, the educational resources at the disposal of teachers are not only more numerous and diverse than a few years ago, but they also require a greater degree of knowledge when they are confined to different subjects or heterogeneous nature, enriching a little more one of the most important professions in our society. In the case of history teaching, the methodological revolution which has taken place over the past two decades has been unstoppable and raises the issue of the total renovation of the teaching profession (Boerman-Cornell, 2015; Souto and Martínez, 2016). This paper addresses the use of comics, and in some cases graphic novels (Montenegro, 2012; Sebastián-Faubel, 2016; Williams, 2008), as an educational resource on the rise which aims to facilitate history learning in a motivating way, encouraging critical thinking and, in turn, stimulating students' creativity.

For many of us, brought up in a society in which the mass media are ever-present, cartoon strips and comics have formed part of what could be referred to as our sentimental upbringing (Gubern, 1977). It is true to state that never before have so many graphic novels been produced and read as is the case today. Thus, this manifestation requires considered reflection and, in our opinion, a suitable educational context in which these resources can be implemented in an effective manner. In this way, the educational potential which comics put at our disposal can be prioritised. If a correct reading is given, by stimulating observation, reading, analysing

and interpreting a series of actions reflected in the cartoons, it is possible to consolidate and acquire a historical approach capable of handling and distinguishing between fiction, truth and reality (Lluch-Prats et al., 2016; Sebastián-Faubel, 2016).

In the words of McCloud (1993, p. 9), comics are juxtaposed illustrations and other images in deliberate sequences which aim to transmit information or obtain an aesthetic response from the reader. Gubern (1977), one of the sharpest analysts of the popular and visual culture of our times, characterised comics as a sequence of consecutive images which articulate a narrative in which at least one stable character and an integration of the text in the image can be found throughout the series. Therefore, the most imperative characteristic is the capacity of the comic or graphic novel to naturally integrate iconic and verbal language (Arango, 2012).

If these classical definitions, used in their time to characterise a society ever more focused on aesthetics (as predicted by Berger, 1990), are taken as a starting point, education cannot overlook one of the most important cultural manifestations of the past decades. In other words, it seems to be the ideal moment to incorporate these forms of mass media, along with others such as cinema and photography, into the classroom, adapting them to educational projects resulting from discussion and intellectual debate.

If, as stated by Gubern (1977), ours is a culture of the printed image for the masses, history teaching should provide itself with valid tools for this culture, not so much to ignore what has been done or taught to date, but to offer a re-reading or an alternative version of history (Arango, 2012; Steiner et al., 2014). In the words of Saitua (2018a), history teachers should provide their students with materials and training resources which foster a clear understanding of complex historical processes, enable them to identify historical causality and to develop the skills associated with argumentation. Consequently, comics, as stated by many different authors (Williams, 2008; Aguilera, 2011; Blay, 2015; Boerman-Cornell, 2015; Souto and Martínez, 2016), could be an ideal resource for history teaching and the development of historical thinking (Ortega-Sánchez and Pagès, 2017) due to their recreational and motivational character, plot clarity and multi-modal nature. For Zagkotas (2019) and Sebastián-Faubel (2016), the narrative formula of comics, an element which should, no doubt, be borne in mind, is another way of narrating the past which is closely linked with the narrative discourse of the historian. As the student is presented with the world *via* an imaginary lens, he/she is encouraged to seek information in an autonomous manner, observing the necessity to confront and contrast the documentary sources, thus fostering discussion and interpretation of the narration, leading to an attitude which is surely more positive concerning the place of history in his/her education.

In close relation to the concepts and ideas described here, Saitua (2018b) points out an element which is fundamental for our interests: reading a comic can enable the student to deploy the necessary mental frameworks for historical thinking. Without doubt, this is an important step in considering this means of expression as a first-order tool in the teaching of history, as is stated in the studies by Aguilera (2011),

Boerman-Cornell (2015) and Souto and Martínez (2016). These authors stress the importance of the identification, contextualisation and corroboration of all of the historical elements which appear in the comic, interpreting them *via* a critical, efficient reading, a characteristic of history teaching itself. In this regard, Arango (2012), Saitua (2018a) and Blay (2015) state that this reading, formulated within an educational project which takes in some of the issues formulated above, not only helps to promote critical thinking among students, but also imagination and creativity. These authors consider that this stimulus, which promotes and encourages imagination, also takes place on what is specifically known as “historical imagination,” defined by Lowenthal (1985, p. 213) as “collective memory is constructed through shared images.” Indeed, the comic strips which present certain historical contents do not restrict themselves exclusively to returning to these shared mental images and reproducing them, as was the case with the well-known graphic novel by Art Spiegelman, *Maus: A Survivor's Tale* (1992). But rather, they transform and insert new images into this ensemble, a fundamental aspect when their use is contemplated in the education of students who are in the process of forming their own historical imagination (Arango, 2012; Lluch-Prats et al., 2016).

Among the main advantages of using comics to teach history, it can be highlighted that not only must the student adopt a role as a researcher-reader, as mentioned by many authors (Aguilera, 2011; Montenegro, 2012; Souto and Martínez, 2016; Zagkotas, 2019), evaluating sources, interpreting information and considering different perspectives, but also different key concepts can be worked on, such as identity-otherness, multiple causality, change and continuity, diversity, inequality and conflict, according to the pattern proposed by Sebastián-Faubel (2016). Thus, comics, used as an educational resource, put into images what textbooks explain in words, fostering students' creativity and increasing their interest in what is taught, as stated by Blay (2015). This is yet another incentive for the use of new methodologies in history teaching as it becomes possible for the teacher to catch his/her students' attention thanks to the complementarity between what is iconic and what is written, as the task of opening the eyes of the reader is facilitated (Arango, 2012; Boerman-Cornell, 2015).

Ultimately, the challenge with which we are confronted when considering the suitability of the use of comics in the history classroom resides in the capacity of teachers to select which comics are to be used in class, what prior information to offer to students and, last but not least, what type of choice is sought from the reading and analysis of this educational resource (Montenegro, 2012; Steiner et al., 2014; Lluch-Prats et al., 2016). In brief, comics are a cultural product generated in certain specific social and productive conditions (Arango, 2012). Therefore, it is the teacher who is responsible for incorporating them into a broader ensemble of strategies with the aim of substantially transforming the teaching of history (Aguilera, 2011; Blay, 2015).

In this regard, the main aspects which make comics a resource of great educational potential lie in their capacity to motivate (Altarriba, 2003) and in the fact that they are considered

to be a transversal and multidisciplinary tool which can facilitate integral learning (Jiménez García et al., 2019). Likewise, Barrero (2002), cited by Jiménez García (2020), points out certain reasons for the benefits of using comics in language teaching, which can, in our opinion, also be applied to other subjects, such as their narrative and visual aspects, the fostering of creativity and the capacity for reflection on cultural and social reality. To all of this can be added the fact that comics enable the introduction of specific contents for different subjects (Pons, 2017).

Along these lines, it can be stated that comics make it possible to work on values and beliefs (Onieva, 2015), reflect on the world around us and favour research (Guzmán, 2011). They are, therefore, an ideal resource for teaching and learning history as their use in the classroom is related, to a great degree, with the objectives and intentions of teachers (Gómez-Trigueros and Ruiz-Bañuls, 2019).

Taking into account the aforementioned statement, comics can be used in history classes in many different ways. First of all, comics can be used as a historical source to address the study and knowledge of the society which produces them (Gual Boronat, 2011, 2013; Sola Morales and Barroso Peña, 2014; Calviño Freire, 2018), as they can show features which reflect the moment of time in which they were created (Jiménez García, 2020). Secondly, another way of implementing the use of comics in the classroom is *via* the design of educational activities based on the contents of a published comic, making it possible to use works which address the specific period being studied or to be studied in class, thus transforming the comic into a tool enabling the study of a specific historical reality *via* argumentation (Saitua, 2018b; Fernández de Arriba, 2019). A third approach to working on historical contents based on comics is the creation of comics by teachers and students. This creation can be by hand or *via* the use of any of the different digital applications available on the Internet. Independently of the tool used to create them, in the case of teachers, the creation and use of comics in the classroom makes it possible to present contents in a different way which can be adapted to both the learning contents and objectives. As far as the creation of comics by students is concerned, this can be the final result of research in the style of a historical narrative.

MATERIALS AND METHODS

Objectives

The main objective of this research was to analyse the motivation of future primary education teachers when using comics as an educational resource in the teaching of history. In order to achieve this objective, three specific objectives (SO) were proposed:

- SO 1: To analyse the elements of motivation that trainee teachers have when using comics as an educational resource in the teaching of history.

- SO 2: To study the positive aspects which future teachers consider when creating their own comics as an educational resource.
- SO 3: To establish the limitations and difficulties of the creation of comics for the teaching of history according to the trainee teachers.

This research is based on the assumption that, in general, the future primary education teachers have not worked with comics as a resource for learning history before and that, therefore, they have not considered comics as a future educational resource which can help them to improve their social sciences classes. Therefore, the following research questions have been proposed as a result of this work: Do the future teachers consider that comics are a motivating resource for history classes? What are their perceptions regarding motivation for learning social sciences?

Participants and Context

This research was carried out on an intentional sample as its participants were, in all cases, trainee teachers from the third year of a degree in primary education at the University of Alicante, Spain.

The sample consisted of 221 future primary education teachers ($n=221$) taking the subject entitled “Didactics of the Social Sciences: History” in which the training programme took place.

As far as the characteristics of the participants are concerned, there was a total of 63 men (28.5%) and 158 women (71.4%) with an average age of 22.4 years.

An Educational Programme Based on Comics to Learn History

The ultimate objective of the educational programme developed was for students to learn to develop their own educational resources using comics as a tool for the teaching and learning of historical contents. For this purpose, first of all, the educational possibilities of comics were introduced.

Thus, in the first phase, an activity was proposed based on a series of cartoons from the comic *Asterix and the Chariot Race* (Gosciny et al., 2017a,b). Several questions were posed with the ultimate aim of working on a specific issue: the importance of roads for the creation and consolidation of the Roman Empire. This activity served as an example for the next educational proposal to be developed, which was based on the creation of a comic by the students. Each student was required to create a comic using a digital application (*Pixton* or *Storyboard*) focusing on a historical episode, character or event.

Once the comic had been created, the next step was to design a series of related activities to work on the contents addressed following the example of the first activity.

Instrument Design, Validation and Research Process

The research tool was designed *ad hoc* for this study, although it was based on similar tools which have already been validated

and published. Like the tool designed by Gómez-Carrasco et al. (2020), its aim was to reveal students' motivation concerning the teaching and learning of the social sciences.

In order to carry out the research, a mixed quantitative and qualitative questionnaire was designed (**Supplementary Material**) with 10 closed numerical statements on a Likert scale (1–5) in ascending level of agreement and two items of an open qualitative nature in which the trainee teachers expressed the main positive and negative aspects of the educational programme.

The items analysed were:

- Item 1. The way the activity was presented and worked on has motivated me to learn more about history.
- Item 2. The way the activity was presented and worked on has motivated me to learn more about the use of comics as a resource for teaching and learning about history.
- Item 3. The activity has improved my motivation to make more of an effort in this subject.
- Item 4. The activity has improved my motivation to achieve better grades.
- Item 5. The activity has motivated me because it has enabled me to contribute my knowledge.
- Item 6. The activity has motivated me because it has enabled me to contribute my creative skills.
- Item 7. The activity has motivated me because we have used resources which are different to those normally used.
- Item 8. The activity has motivated me because I have been able to be responsible for my own learning.
- Item 9. The activity has motivated me because it has taught me to design my own contents and activities via comics.
- Item 10. The activity has motivated me because I consider it to be useful for my future career.
- Item 11 (open). State and explain the positive aspects of the creation of comics for the teaching and learning of history in primary education.
- Item 12 (open). State and explain the negative aspects of the creation of comics for the teaching and learning of history in primary education.

In order to validate the reliability of the construct, an internal consistency analysis (**Supplementary Material**) was carried out on the total of the numerical scale ($n=10$ items) which revealed a Cronbach's alpha value of 0.902 ($0.9>1$, representing an extremely high degree of reliability according to Oviedo and Campo-Arias, 2005) and a value of 0.873 with the Guttman split-half coefficient (**Supplementary Material**), thereby demonstrating a high degree of internal reliability of the items on the scale according to other research in the field of social science education, such as that carried out by Gestsdóttir et al. (2018), Gómez-Carrasco et al. (2019) and Moreno-Vera et al. (2020).

Furthermore, as it is a mixed quantitative-qualitative questionnaire, the Friedman test was employed to establish whether there is dependence between the quantitative items and if these affect the qualitative items (**Supplementary Material**). In this case, the values of the Chi-square test vary between

a low of 69.2 in item 4 to 264.6 for item 7, all of which are far removed from 0 (>0.05), which, according to Satorra and Bentler (2010) and Sharpe (2015), indicates that there is discrepancy between the variables and that, therefore, they are not dependent. This bestows a positive reliability value on the qualitative research, as occurs in other studies on the teaching of the social sciences, such as Moreno-Vera et al. (2020).

Lastly, as far as the research procedure is concerned, it must be pointed out that participation in the research was voluntary and subject to ethical guidelines, and the trainee teachers gave informed consent for their participation. Secondly, both the educational programme and the later test were applied in class time by professionals trained for that purpose. Thirdly, the analysis of the research results was carried out using IBM SPSS v. 24 software for the analysis of the descriptive statistics of the questionnaire, whereas the qualitative analysis was carried out with the AQUAD 7 program (Huber, 2013), establishing the variables based on the narratives of the trainee teachers.

RESULTS

First of all, as far as the analysis of the descriptive statistics is concerned, it must be stated that the analysis of the means, medians and SD (**Supplementary Material**) of the 10 quantitative items of the questionnaire has been extremely positive. The item with the lowest mean score is item 4 with 4.0 out of 5, which, in itself, represents a positive result. Thus, the mean values oscillate between 4.0 for item 4 (The activity has improved my motivation to achieve better grades) and 4.6 for item 7 (The activity has motivated me because we have used resources which are different to those normally used).

In relation to the median, all of the items are situated between 4 and 5. However, in terms of the SD, all of the items analysed present a value of <1 , thus representing a positive detail as no item has a great degree of variability. Item 7 (The activity has motivated me because we have used resources which are different to those normally used) presents the greatest homogeneity of responses with the SD below 0.6, whereas item 5 (The activity has motivated me because it has enabled me to contribute my knowledge) is that which presents the greatest variability with 0.9, which, in spite of everything, is a very positive piece of information.

Motivation of Future Teachers When Using Comics as a Resource to Teach and Learn History

The statistical and descriptive analysis (**Supplementary Material**) of the items reveals extremely positive results concerning the perception of trainee primary education teachers regarding the use of comics as an educational resource for the learning of history.

Item 1 (The way the activity was presented and worked on has motivated me to learn more about history; **Table 1**) has a mean score of 4.3 out of 5, this score being the most represented with 52.5% of the students feeling high, motivated

TABLE 1 | The way the activity was presented and worked on has motivated me to know more about history.

VAR00001				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	2	0.9	0.9	0.9
2.00	1	0.5	0.5	1.4
3.00	29	13.1	13.1	14.5
4.00	73	33.0	33.0	47.5
5.00	116	52.5	52.5	100.0
Total	221	100.0	100.0	

TABLE 2 | The way the activity was presented and worked on has motivated me to know more about the use of comics as a resource for teaching and learning about history.

VAR00002				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	1	0.5	0.5	0.5
2.00	4	1.8	1.8	2.3
3.00	19	8.6	8.6	10.9
4.00	76	34.4	34.4	45.2
5.00	121	54.8	54.8	100.0
Total	221	100.0	100.0	

by comics as they enable them to obtain more conceptual contents on the subject of history. If the accumulated percentages of values 1, 2 and 3 are added up, it can be noted that only 14.5% of trainee teachers consider that the use of comics is not motivating in terms of learning about history.

Item 2 (The way the activity was presented and worked on has motivated me to learn more about the use of comics as a resource for teaching and learning about history; **Table 2**) has a positive mean score of 4.4. In this case, 89.2% of the trainee teachers responded that the activity motivated them to discover more about comics as a resource if the percentages of those who are “in agreement” and “totally in agreement” are taken into account. Only 7.3% of the students did not agree that the activity had motivated them to know more about the use of comics in the teaching of history.

As for the statistical and descriptive analysis of item 3 (The activity has improved my motivation to make more of an effort in this subject; **Table 3**), the majority of the future teachers stated that they were in agreement. The mean for this item was 4.1 out of 5 and it is worthy of note that none of the participants stated that they were “totally in disagreement” with this statement and only 2.3% were in disagreement. On the other hand, the most commonly selected option was “in agreement” with 44.3% of the sample, whereas 38% of the participants were “totally in agreement.”

Item 4 (The activity has improved my motivation to achieve better grades; **Table 4**), as explained above, is the item with the lowest mean value (4.0) which, even so, is an extremely positive result. As far as the percentage is concerned, it is

worthy of note that there were no students who were “totally in disagreement” and only 2.7% were “in disagreement.” It is interesting to note how, in this item, 26.7% of the participants were neither in agreement nor disagreement, which demonstrates certain doubts when responding. This leads us to consider that improving grades and marks is not such an important issue for future teachers. On the other hand, those participants who do indeed feel motivated to obtain better grades represent similar values with 39.4% of the sample being “in agreement” and 33.9% “totally in agreement.”

As for item 5 (The activity has motivated me because it has enabled me to contribute my knowledge; **Table 5**), the mean score is 4.2 with 79.2% of the future teachers declaring that they were “in agreement” or “totally in agreement” that using comics as a teaching resource motivates them to obtain greater conceptual contents of history. In this case, only 5.9% of the participants were “in disagreement” with this statement.

Item 6 (The activity has motivated me because it has enabled me to contribute my creative skills; **Table 6**) has a very positive mean score of 4.4. Only 2.8% of the trainee teachers are “in disagreement,” while the percentage of those in doubt is relatively low at 10.4%. As regards the motivation which they have felt due to being able to develop aspects of their own creativity *via* the activity with comics, the vast majority (89.6%) are “in agreement” that this creative aspect motivates them. Indeed, 58.8% of the participants claim to be “totally in agreement” with this point.

As stated above, the statistical analysis of item 7 (The activity has motivated me because we have used resources which are different to those normally used; **Table 7**) has the highest mean score with 4.6 out of 5. It is noteworthy in this case that 71% of the future teachers are “totally in agreement” that using teaching resources which are different to those normally used proves to be a motivating factor. Furthermore, including the 19.5% who were “in agreement,” the positive percentage for this item increases to 90.5%, the highest of all the items. This indicates that the trainee teachers value positively the use of resources such as comics as it enables them to introduce resources other than the textbook (still the most commonly used resource) into their history classes, thereby offering the possibility of working on history *via* active learning methodologies and preventing students from playing a passive role in the classroom.

When examining the analysis of item 8 (The activity has motivated me because I have been able to be responsible for my own learning; **Table 8**), a high mean score of 4.3 out of 5 can be observed. 84.1% of the trainee primary teachers are “in agreement” or “totally in agreement” with the idea that playing an active role within the teaching and learning process is, in their opinion, a motivating factor. This result is related with the previous item and confirms that students feel more motivated in history classes when they are the ones to carry out activities relating to the application of knowledge rather than merely receiving the lesson in a passive way.

Item 9 (The activity has motivated me because it has taught me to design my own contents and activities *via* comics; **Table 9**) also presents a positive mean score of 4.3. In this

TABLE 3 | Activity has improved my motivation to make more effort in the subject.

VAR00003				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
2.00	5	2.3	2.3	2.3
3.00	34	15.4	15.4	17.6
4.00	98	44.3	44.3	62.0
5.00	84	38.0	38.0	100.0
Total	221	100.0	100.0	

TABLE 4 | Activity has improved my motivation to achieve better grades.

VAR00004				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
2.00	6	2.7	2.7	2.7
3.00	53	24.0	24.0	26.7
4.00	87	39.4	39.4	66.1
5.00	75	33.9	33.9	100.0
Total	221	100.0	100.0	

case, the trainee teachers also consider that the training programme working on history *via* comics motivated them as it enabled them to design their own comics and materials to use in the future with primary schoolchildren. 52.5% of the participants were “totally in agreement” and 36.7% were “in agreement.” Only 2.7% of the participants stated that they did not feel motivated designing and creating their own teaching materials in the format of a comic.

Last of all, item 10 (The activity has motivated me because I consider it to be useful for my future career; **Table 10**) obtained an extremely high mean score of 4.5 out of 5. What is more, 90.1% of the sample (the second highest percentage of the questionnaire) were “in agreement” with the idea that learning about history *via* comics will be useful in their future teaching careers in primary education. Indeed, the lowest value of those participants who were “totally in disagreement” of the whole questionnaire can be observed in this item (only 1.4% of the participants).

Positive Perceptions on Creativity When Using Comics as a Resource for Teaching History

Item 11 of the questionnaire was an open question aimed at revealing the positive aspects which future teachers perceive regarding the use of comics as a resource for teaching and learning history in primary education (**Table 11**).

The analysis from the AQUAD 7 software offers up a series of results (**Supplementary Material**) which, *a priori*, are of interest as there is a total of 580 different open responses from the participants, from which 23 different variables can be extracted which are considered positive, such as: being an accessible resource which is, at the same time, entertaining

TABLE 5 | Activity has motivated me because it has enabled me to contribute my knowledge.

VAR00005				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	1	0.5	0.5	0.5
2.00	11	5.0	5.0	5.4
3.00	34	15.4	15.4	20.8
4.00	69	31.2	31.2	52.0
5.00	106	48.0	48.0	100.0
Total	221	100.0	100.0	

TABLE 6 | Activity has motivated me because it has enabled me to contribute my creative skills.

VAR00006				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	1	0.5	0.5	0.5
2.00	4	1.8	1.8	2.3
3.00	18	8.1	8.1	10.4
4.00	68	30.8	30.8	41.2
5.00	130	58.8	58.8	100.0
Total	221	100.0	100.0	

and fun and fosters research and autonomous work on the part of the students (this is directly related with the fact that the educational action was, in this case, individual).

Three variables particularly stand out which the future teachers perceive to be positive concerning the use of comics in the teaching of history. On the one hand, the most repeated aspect was that it “encourages the understanding of conceptual contents” with 73 responses and 12.6% of the total answers. Secondly, it was considered to be very positive that comics are a “motivating” resource for students. This response was given 72 times with 12.4% of the answers. Thirdly, 69 responses (11.9%) highlighted the fact that comics are a resource which makes it possible to “encourage creativity.”

It is coherent that comics are considered to be a “motivating” resource given that it coincides with and reaffirms the good results of the quantitative items along the same lines. However, it is noteworthy that the most highly valued item regarding motivation was item 7 (The activity has motivated me because we have used resources which are different to those normally used), whereas, in the open responses, it was highlighted that “it favours the understanding of historical contents,” with responses related with using “different resources in the classroom” being a long way behind, only appearing in 4.5% of the responses.

A good example can be found in the response of participant 108, who commented “It increases motivation, encourages creativity, helps children to understand the subject matter better as they use illustrations in the form of cartoon strips, without too much text and putting relevant information.” Furthermore, participant 178 considered that “on the other hand, comics encourage creativity, imagination, initiative and the go-getting spirit sought in so many other subjects.”

TABLE 7 | Activity has motivated me because we have used resources which are different to those normally used.

VAR00007				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
2.00	3	1.4	1.4	1.4
3.00	18	8.1	8.1	9.5
4.00	43	19.5	19.5	29.0
5.00	157	71.0	71.0	100.0
Total	221	100.0	100.0	

TABLE 8 | Activity has motivated me because I have been able to be responsible for my own learning.

VAR00008				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	1	0.5	0.5	0.5
2.00	5	2.3	2.3	2.7
3.00	29	13.1	13.1	15.8
4.00	77	34.8	34.8	50.7
5.00	109	49.3	49.3	100.0
Total	221	100.0	100.0	

Limitations and Difficulties When Using Comics in History Education

As far as item 12 of the questionnaire is concerned, the participants were required to indicate any negative aspects which they had perceived in the research (Table 12).

In this case, the qualitative analysis of the open question (Supplementary Material) yielded a total of 298 different responses, representing almost half the number of negative responses compared to positive ones. Furthermore, 8.4% of them (a total of 25 responses) were merely to state that no great negative aspects had been found when applying the use of comics and creativity for teaching and learning history.

The trainee history teachers mainly pointed out as a negative aspect (22.5%) the “preparation time” required to create their own comic to represent a historical event or character. For example, participant 47 stated “One negative aspect could be the time it took to create. If a comic had to be made for each topic throughout the year, I do not think a teacher would have the time it would take.” The preparation for a comic also involves searching for historical information in order to make the data offered more rigorous. This coincides with the opinion of participant 176: “One of the negative aspects in creating comics is the time employed as it is necessary to plan the story and what you want to tell really well.”

Furthermore, 15.1% of the participants commented on the “difficulty in presenting and adapting historical contents to the comic format.” In this regard, participant 20 remarked: “The only negative aspect I can find is that the year group the comic is aimed at must be taken into consideration due to the fact that, as it is not a constant narration (which is what students are accustomed to), it can be difficult to follow

and understand the story.” This aspect can even be reinforced with perceptions such as that of participant 55, who recalled that “perhaps the only drawback is that when representing certain facts, they can be interpreted as being closed and not continuous.”

Finally, the third most common negative aspect among the participants is that “it is a very time-consuming task,” with 33 responses (11.1% of the total). This is closely related with the lack of time to research, use sources, prepare the topic and, finally, to create their own comic. Thus, participant 123 commented: “It involves a lot of work because, first of all, you have to learn to use the application and then put the contents, search for pictures, etc. You have to be very careful about what you want to explain through the comic.” Even more than this, it was perceived that this negative aspect can be aggravated if the comic has to be devised and created by trainee primary teachers: “It is very time-consuming to create a comic, and even more so if you do so with primary schoolchildren” (participant 40).

DISCUSSION

The research presented in this paper is based on the prior assumption that surmised that the trainee primary teachers had never worked on history *via* a resource such as comics. Thus, the researchers wondered whether the future teachers would feel motivated teaching and learning history using comics as an educational resource.

In this regard, the prime objective posed in this research was to analyse the perceptions of trainee teachers regarding their degree of motivation when learning history through comics.

In order to achieve this aim, a specific educational programme on the use of comics in the teaching of the social sciences was designed, along with a mixed qualitative and quantitative questionnaire created *ad hoc* to evaluate the degree of motivation of the 221 participating trainee primary teachers from the University of Alicante. The results of the construct validity confirm that, methodologically speaking, the tool is highly reliable, as has been the case in other similar research projects on motivation in the teaching of the social sciences, such as that of Gómez-Carrasco et al. (2020).

When analysing the results, three different specific objectives were proposed. First of all, the different aspects of motivation perceived by the participants when using comics in social science classes have been researched (SO1). Here, it should be highlighted that the results concerning motivation have been extremely positive, with mean scores of above 4 out of a possible 5. Item 7 (The activity has motivated me because we have used resources which are different to those normally used) stands out with the highest mean score of 4.6 and, furthermore, a positive result indicating that 90.5% of the participants were “in agreement” that working with educational resources which were different to those normally employed proves motivating in history classes.

This can basically be explained by the fact that the teaching of the social sciences, and history in particular, still maintains

TABLE 9 | Activity has motivated me because it has taught me to design my own contents and activities *via* comics.

VAR00009				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
2.00	6	2.7	2.7	2.7
3.00	18	8.1	8.1	10.9
4.00	81	36.7	36.7	47.5
5.00	116	52.5	52.5	100.0
Total	221	100.0	100.0	

TABLE 10 | Activity has motivated me because I consider it to be useful for my future career.

VAR00010				
Valid	Frequency	Percentage	Valid percentage	Cumulative percentage
1.00	1	0.5	0.5	0.5
2.00	1	0.5	0.5	0.9
3.00	20	9.0	9.0	10.0
4.00	53	24.0	24.0	33.9
5.00	146	66.1	66.1	100.0
Total	221	100.0	100.0	

an extremely rigid and traditional structure (Estepa, 2017; Moreno-Vera and Alvé, 2020), in which the explanatory method (the masterclass) continues to be the most common approach when constructing new knowledge. The textbook is still the most commonly used educational resource in history classes, as pointed out by Vera et al. (2014), who, following an investigation of history teaching in Spain, concluded that more than 85% of teachers continued to employ textbooks as the only resource in the classroom. Thus, using resources other than the textbook in itself represents an element of motivation according to the future primary teachers.

This result is confirmed if the fact is taken into account that 84.1% of the trainee teachers feel motivated when they are responsible for their own learning. The participants in the study consider it to be a factor of motivation that it is the students who shoulder the responsibility of constructing knowledge (Powell and Kalina, 2009), putting into practice educational strategies which facilitate creativity (Cooper, 2018), historical research employing sources (Prats and Santacana, 2001; Ortega-Sánchez and Pagès, 2020), active learning methods (Gómez-Carrasco et al., 2019) and the use of narratives to learn history (Moreno-Vera, 2015).

Some of the results prove to be of particular interest, such as those of item 4, which have the lowest mean score of the study, thereby indicating that the trainee teachers do not attribute so much importance to motivation depending on the final mark that will be awarded.

SO2 was concerned with the analysis of the positive perceptions of the participants when using comics as a resource for the teaching of history. In this case, a wide variety of favourable opinions can be found among a total of 580 responses,

TABLE 11 | Positive aspects regarding the use of comics as a resource for learning history.

Variable	Frequency	%
Favours understanding of contents	73	12.6
Motivation	72	12.4
Encourages creativity	69	11.9
Visual resource facilitates understanding	46	7.9
Entertaining resource	45	7.7
Innovative resource/innovation	37	6.4
Makes it possible to adapt contents	28	4.8
Favours autonomous and active learning	28	4.8
Makes it possible to work on basic competences	27	4.6
Makes it possible to learn in a different way	26	4.5
Favours working with competences associated with the development of historical thinking	23	4
Arouses students' curiosity/interest/attention	19	3.3
Encourages reading	17	3
An attractive resource	15	2.6
A dynamic resource	12	2.1
Favours significant learning	11	1.9
Encourages research	8	1.4
Favours the capacity of synthesis	7	1.2
Transversality	7	1.2
Simple to carry out	4	0.7
Favours group work	3	0.5
Possibility of developing critical thinking	2	0.3
An accessible resource	1	0.2
Total	580	100

among which those referring to the capacity of comics to “encourage understanding of historical contents” stand out, thus coinciding with prior research which has treated narrative resources as a fundamental element when learning history in primary and early-years education (Moreno-Vera, 2015), in which narration has been a significant element for researchers such as Egan (1994).

Indeed, as stated by González Gaxiola et al. (2020), one of the main reasons why teachers resort to graphic narratives (comics, graphic novels, etc.) is to facilitate understanding of abstract concepts among students. Without doubt, the narrative nature of comics with the combination of text and images enables a better understanding of the ideas and contents presented. Therefore, the visual aspect of this resource must be highlighted, as images have, in the words of Becerra Romero and Jorge Godoy (2014, p. 17), the “added value that they avoid pages and pages of descriptions and help to set the scene and contextualise the characters.” In this way, comics facilitate the understanding and visualisation of historical information (Saitua, 2018b; Ortega-Sánchez et al., 2019).

TABLE 12 | Negative aspects regarding the use of comics as a resource for learning history.

Variable	Frequency	%
Preparation time	67	22.5
Difficulty in presenting/ adapting contents	45	15.1
Time-consuming work	33	11.1
None	25	8.4
Lack of ICT training	19	6.4
Payment for applications	18	6
Perception of entertainment	17	5.7
Difficulty for students	12	4
Insufficient for learning historical contents	10	3.3
Lack of necessary (ICT) resources	9	3
Lack of training for teachers	9	3
Effort	8	2.7
Lack of creative and artistic skills	5	1.7
The elaboration of subsequent activities	5	1.7
Lack of motivation among students	4	1.3
Students do not like it	3	1
Perception of lack of learning	3	1
May cause disruptive behaviour in class	2	0.7
Necessary to know the specific language of comics	2	0.7
Intimidating for students due to its great amount of possibilities	1	0.3
The starting point	1	0.3
Total	298	100

Furthermore, the results obtained *via* the open questions endorse the idea that comics are a tool with great potential in terms of motivating students (Altarriba, 2003; Gutiérrez, 2006; Saitua, 2018a). This perception is, doubtless, connected with the fact that they are entertaining, which, in the light of the participants' responses, has passed from being considered prejudicial, as has traditionally been the case (Saitua, 2018a) and as can be observed in some of the responses regarding negative aspects, to being considered by the vast majority as a positive element which can attract the students' interest and facilitate learning.

These positive results regarding the "motivation" which comics produce among students reinforce the encouraging data shown by the research on motivation *via* the analysis of the quantitative data. Last of all, the participants highlighted the capacity of comics as a resource to "foster students' creativity," thus coinciding with Orlich et al. (1994) and Baur (1978), who stated that, in order to achieve creative thinking, practices which foster it must be worked on in the classroom.

SO3 focused on the qualitative analysis of the possible difficulties, limitations and negative aspects of the use of comics

in history classes. In this regard, of a total of 298 responses, the most noteworthy were focused on formal aspects.

Without doubt, introducing comics into the classroom, be it *via* pre-existing editions or original creations, has a series of limitations or negative aspects which should be taken into account, such as the preparation time required (mentioned by 22.5% of the participants) and, in many cases, the work required, implying the preparation of materials, the use of sources and historical evidence (Seixas and Morton, 2013), research (Prats and Santacana, 2001) and, last of all, the creation of the comic itself.

Likewise, another of the drawbacks or objections mentioned by the participants relates to the difficulty of presenting or adapting historical contents to the format of a comic (15.1% of the total). In this regard, it must be highlighted that this requires a significant capacity for synthesis and organisation of the structure of the comic. However, as stated by Becerra Romero and Jorge Godoy (2014), it should be taken into account that comics are not history books, but rather a resource or complementary material to assist students in approaching a specific topic or subject.

In conclusion, therefore, it can be affirmed that the trainee primary education teachers consider the use of comics to be a motivating resource for the teaching and learning of history, particularly because it makes it possible to work with resources other than the textbook, something that also happen in geography, according to authors like Morote et al. (2021). Furthermore, they consider the use of comics to be extremely useful for their future careers. In addition, they highlight as a positive aspect the fact that comics are capable of facilitating the understanding of historical contents, fostering creativity and motivating their future students. As far as negative aspects are concerned, the participants highlight the excessive time and difficulty implied by creating an original comic in class and the difficulty of correctly adapting historical contents to the format of a comic.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de ética de la Universidad de Murcia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JM-V: materials and methods, quantitative analysis and results, and discussion. SP-L: instrument design, data collection,

educational programme design, qualitative results, and discussion. RB-M: introduction, references, data collection, and discussion. All authors contributed to the article and approved the submitted version.

FUNDING

This research is part of the project “Accept the challenge! Gamification in online higher education” (2020-1-PL01-KA226-HE-096034) funded by the European Union ERASMUS+ KA2 programme Cooperation and innovation for good practices.

ACKNOWLEDGMENTS

The researchers would like to express their gratitude for the help and collaboration offered in the application of this

study by our colleagues from the DICSO (Didáctica de las Ciencias Sociales de la Universidad de Murcia) research group, through which they participate in the research project which has funded this study. We would particularly like to thank the principal investigators Pedro Miralles and Cosme J. Gómez. We would also like to thank the department of Didactics of the Social Sciences of the University of Alicante and, in particular, the colleagues who participate in the innovation network in the Didactics of the Social Sciences, in which this educational experiment has been carried out.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.778792/full#supplementary-material>

REFERENCES

- Aguilera, I. B. (2011). Las TIC's y las viñetas: una propuesta didáctica sobre los totalitarismos a través del cómic Maus. *Espiral. Cuadernos del profesorado* 4, 42–57. doi: 10.25115/ecp.v4i8.924
- Altarriba, A. (2003). Enseñar con viñetas. Peculiaridades y aplicaciones didácticas de la historieta. Mosaico: revista para la Promoción y Apoyo a la enseñanza del Español, 23, 4–9.
- Arango, N. M. (2012). De lo distractivo a lo instructivo: algunas aproximaciones al uso de la historieta histórica en la enseñanza de la historia. *Praxis Pedagógica* 12, 166–193. doi: 10.26620/uniminuto.praxis.12.13.2012.166-193
- Barrero, M. (2002). Los cómics como herramientas pedagógicas en el aula. Jerez de la Frontera (Cádiz): Jornadas sobre Narrativa Gráfica.
- Baur, E.K. (1978.) *La historieta como experiencia didáctica*. Mexico: Ed. Nueva Imagen.
- Becerra Romero, D., and Jorge Godoy, S. (2014). Un acercamiento didáctico a la primera mitad del S. XX a través de los cómics. *Espacio, Tiempo y Forma. Serie V Historia Contemporánea* 26, 15–40. doi: 10.5944/etfv.26.2014.14501
- Berger, J. (1990). *El sentido de la vista*. Madrid: Alianza Editorial.
- Blay, J. M. (2015). Dibujando la Historia. El cómic como recurso didáctico en la clase de Historia. *Revista Supervisión* 21, 1–14.
- Boerman-Cornell, W. (2015). Using historical graphic novels in high school history classes: potential for contextualization, sourcing, and corroborating. *Hist. Teach.* 48, 209–224.
- Calviño Freire, D. (2018). “La historia de los cuatro inmigrantes: el cómic como fuente histórica,” in *Nuevas visiones sobre el cómic. Un enfoque interdisciplinar*. eds. J. A. G. Lana and A. A. Suñer (Zaragoza: Universidad de Zaragoza), 113–118.
- Cooper, H. (2018). What is creativity in history? *Education* 3-13, 636–647. doi: 10.1080/03004279.2018.1483799
- Egan, K. (1994). *Fantasia e imaginación: su poder en la enseñanza*. Madrid: Morata.
- Estepa, J. (2017). *Otra didáctica de la historia para otra escuela*. Huelva: Servicio de Publicaciones de la Universidad de Huelva.
- Fernández de Arriba, D. (2019). *Memoria y viñetas. La memoria histórica en el aula a través del cómic*. Paterna: Desfiladero Ediciones.
- Gestsdóttir, S. M., Van Bostel, C., and Van Drie, J. (2018). Teaching historical thinking and reasoning: construction of an observation instrument. *Br. Educ. Res. J.* 44, 960–981. doi: 10.1002/berj.3471
- Gómez-Carrasco, C. J., Monteagudo, J., Moreno-Vera, J. R., and Sáinz, M. (2019). Effects of a gamification and flipped-classroom program for teacher in training on motivation and learning perception. *Educ. Sci.* 9:299. doi: 10.3390/educsci9040299
- Gómez-Carrasco, C. J., Monteagudo, J., Moreno-Vera, J. R., and Sáinz, M. (2020). Evaluation of a gamification and flipped-classroom program used in teacher training: perception of learning and outcome. *PLoS One* 15:e0241892. doi: 10.1371/journal.pone.0241892
- Gómez-Trigueros, I., and Ruiz-Bañuls, M. (2019). El cómic como recurso didáctico interdisciplinar. Tebeosfera, 10. Available at: https://www.tebeosfera.com/documentos/el_comic_como_recurso_didactico_interdisciplinar.html (Accessed July 29, 2021).
- González Gaxiola, F., Galindo Ruiz de Chávez, M. A., and Gutierrez Rohan, D. C. (2020). Actitudes y disposiciones de docentes de Bachillerato para usar la narrativa gráfica. *Didáctica. Lengua y Literatura* 32, 171–182. doi: 10.5209/dida.71795
- Gosciny, R., and Ferri, J-Y, Uderzo, A., and Conrad, D. (2017a). *Astérix en Italia*. Madrid: Bruño.
- Gosciny, R., Ferri, J-Y, Uderzo, A., and Conrad, D. (2017b). *Asterix and the Chariot Race*. London: Ed. Hachette Children.
- Gual Boronat, O. (2011). El cómic como fuente histórica: el falso testimonio de Tintín en el Congo Belga. *Espacio, Tiempo y Forma* 23, 141–158. doi: 10.5944/etfv.23.2011.1578
- Gual Boronat, O. (2013). *Viñetas de posguerra: los cómics como fuente para el estudio de la historia*. València: Universitat de València.
- Gubern, R. (1977). *Comunicación y cultura de masas*. Barcelona: Península.
- Gutiérrez, T. (2006). El cómic en los adolescentes. Estudio y práctica en el aula. Una propuesta de evaluación. *Arte, Individuo y Sociedad* 18, 29–56.
- Guzmán, M. (2011). El cómic como recurso didáctico. *Pedagogía Magna* 10, 122–131.
- Huber, G.L. (2013). *AQUAD7-Analysis of Qualitative Data*. Tübingen: University of Tübingen.
- Jiménez García, E. (2020). Comunicación verbal y no verbal en el cómic como recurso didáctico en la enseñanza de español como lengua extranjera. *Rev. incl.* 7, 29–37.
- Jiménez García, E., Gómez Redondo, S., and Francisco Carrera, F. J. (2019). “Del cómic al teatro: propuestas didácticas interdisciplinares,” in *Patrimonio, creatividad y teatro*, eds. R. de la Fuente Ballesteros, C. Munilla Garrido and J. García-Medal Villanueva (Valladolid: Verdelis), 305–314.
- Lluch-Prats, J., Martínez Rubio, J., and Souto, L. C. (eds.) (2016). “Propuestas entre la imagen y la palabra: el cómic y sus contornos,” in *Las batallas del cómic: Perspectivas sobre la narrativa gráfica contemporánea*. (Valencia: Diablotexto), 6–14.
- Lowenthal, D. (1985). *The Past Is a Foreign Country*. Cambridge, New York: Cambridge University Press.
- McCloud, S. (1993). *Understanding Comics: The Invisible Art*. New York: Harper Collins.

- Montenegro, C. (2012). "Novelas gráficas: reconstrucciones de la memoria," in *Investigar con jóvenes: ¿Qué sabemos de los jóvenes como productores de cultura visual?* ed. G. I. Edarte (Pamplona: Pamiela-Edarte), 174–184.
- Moreno-Vera, J. R. (2015). La narración literaria, una experiencia didáctica en la enseñanza de historia. *Boletín REDIFE* 4, 42–48.
- Moreno-Vera, J. R., and Alvén, F. (2020). Concepts for historical and geographical thinking in Sweden's and Spain's primary education curricula. *Hum. Soc. Sci. Commun.* 7:107. doi: 10.1057/s41599-020-00601-z
- Moreno-Vera, J. R., Ponsoda-López de Atalaya, S., López-Fernández, J. A., and Blanes-Mora, R. (2020). Holistic or traditional conceptions of heritage among early-childhood and primary trainee teachers. *Sustainability* 12:8921. doi: 10.3390/su12218921
- Morote, A. F., Hernández, M., and Olcina, J. (2021). Are future school teachers qualified to teach flood risk? An approach from the geography discipline in the context of climate change. *Sustain. For.* 13:8560. doi: 10.3390/su13158560
- Onieva, J. (2015). El cómic online como recurso didáctico en el aula. Web y aplicaciones para móviles. *Filología y Didáctica de la Lengua* 15, 105–127.
- Orlich, D. C., Harder, R. J., Callahan, R.C., and Gibson, H. W. (1994). *Técnicas de enseñanza. Modernización en el aprendizaje*. Mexico: Limusa Wiley.
- Ortega-Sánchez, D., and Pagès, J. (2017). Las representaciones sociales de los problemas contemporáneos en estudiantes de magisterio de Educación Primaria. *Revista Investigación en la Escuela* 93, 1–16. doi: 10.12795/IE.2017.i93.01
- Ortega-Sánchez, D., and Pagès, J. (2020). The end-purpose of teaching history and the curricular inclusion of social problems from the perspective of primary education trainee teachers. *Sociol. Sci.* 9:9. doi: 10.3390/socsci9020009
- Ortega-Sánchez, D., Sanz de la Cal, E., and Ibáñez, J. (2019). Literacies and the development of social, critical, and creative thought in textbook activities for primary education in social sciences and the spanish language. *Front. Psychol.* 10:2572. doi: 10.3389/fpsyg.2019.02572
- Oviedo, H. C., and Campo-Arias, A. (2005). An approach to the use of cronbach's alfa. *Rev. Colomb. Psiquiatr.* 34, 572–580.
- Pons, A. (2017). *La cárcel de papel. Diario de un lector de tebeos (2002–2016)*. Madrid: Confluencias.
- Powell, K., and Kalina, C. (2009). Cognitive and social constructivism: developing tools for an effective classroom. *Education* 130, 241–250.
- Prats, J., and Santacana, J. (2001). *Enseñar historia: notas para una didáctica renovadora*. Mérida: Consejería de Educación, Ciencia y Tecnología.
- Saitua, I. (2018a). La enseñanza de la Historia a través de la novela gráfica: una estrategia de aprendizaje emergente. *Didácticas Específicas* 18, 65–87.
- Saitua, I. (2018b). Tristísima ceniza: una propuesta de intervención para enseñar la Guerra Civil en el País Vasco con la novela gráfica. *Clio* 44, 182–209.
- Satorra, A., and Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika* 75, 243–248. doi: 10.1007/s11336-009-9135-y
- Sebastián-Faubel, V. (2016). "El uso del cómic y la narrativa gráfica como estrategia didáctica en el aula de Historia y Ciencias Sociales," in *Las batallas del cómic: Perspectivas sobre la narrativa gráfica contemporánea*. eds. J. Lluch-Prats, J. Martínez-Rubio and L. C. Souto (Valencia: Diablotexto Digital), 314–333.
- Seixas, P., and Morton, T. (2013). *The Big Six Historical Thinking Concepts*. Toronto: Nelson Ed.
- Sharpe, D. (2015). Chi-square test is statistically significant: now what? *Pract. Assess. Res. Eval.* 20, 1–10. doi: 10.7275/tbfa-x148
- Sola Morales, S., and Barroso Peña, G. (2014). El cómic de no-ficción como fuente para el estudio de los conflictos bélicos: Crónicas de Jerusalén. *Hist. Comun. Soc.* 19, 231–248. doi: 10.5209/rev_HICS.2014.v19.47294
- Souto, L. C., and Martínez, J. (2016). Perspectivas, toma de conciencia y consolidación de la historieta en el mundo académico. *Diablotexto digital* 1, 1–5. doi: 10.7203/diablotexto.1.9033
- Steiner, S. F., Peralta, C., and Boothe, D. (2014). "Graphic novels: a bridge between print and media for ELL teaching," in *International Conference ICT for Language Learning. Conference presentation*; November 13, 2014 (Florence: Pixel - International Education and Training Institution).
- Vera, M. I., Moreno-Vera, J. R., and Torres Soto, A. (2014). "La enseñanza de la historia en España, la percepción de los estudiantes en torno a las características de los docentes y su metodología," in *La historia de España en los recuerdos escolares: análisis, interpretación y poder de cambio de los testimonios de los profesores y alumnos*. ed. N. M. Valcárcel (Valencia: Nau Llibres), 179–210.
- Williams, R. M. C. (2008). Image, text, and story: comics and graphic novels in the classroom. *Art Educ.* 61, 13–19. doi: 10.1080/00043125.2008.11652072
- Zagkotas, V. (2019). Are comic books appropriate for teaching history? Three suggestions for greek primary education. *Education* 47, 358–365. doi: 10.1080/03004279.2018.1452955

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Moreno-Vera, Ponsoda-López de Atalaya and Blanes-Mora. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Implicit Teacher Theories Regarding the Argumentative Commentary of Multimodal Texts in the Teaching of Spanish as a Native and Foreign Language

OPEN ACCESS

Edited by:

Pedro García Guirao,
University of Ostrava, Czechia

Reviewed by:

Cristina Cañamares Torrijos,
University of Castilla La
Mancha, Spain
Eva María Iñesta Mena,
University of Oviedo, Spain

*Correspondence:

María Teresa Caro Valverde
maytecar@um.es
José Manuel de Amo Sánchez-Fortún
jmdeamo@ual.es
Juana Celia Domínguez-Oller
jdo357@ual.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 29 July 2021

Accepted: 13 September 2021

Published: 27 October 2021

Citation:

Caro Valverde MT, Amo
Sánchez-Fortún JMd and
Domínguez-Oller JC (2021) Implicit
Teacher Theories Regarding the
Argumentative Commentary of
Multimodal Texts in the Teaching of
Spanish as a Native and Foreign
Language. *Front. Psychol.* 12:749426.
doi: 10.3389/fpsyg.2021.749426

María Teresa Caro Valverde^{1*}, José Manuel de Amo Sánchez-Fortún^{2*} and Juana Celia Domínguez-Oller^{2*}

¹ Teaching Department of Language and Literature (Spanish, English and French), University of Murcia, Murcia, Spain,

² Department of Education, University of Almería, Almería, Spain

The research works linked to the thinking of the teaching staff influence the relevant influence that implicit theories exert on decision-making about classroom practice and on the academic performance of students. In this sense, the present study focuses on the teaching belief system about the development of argumentation in the commentary of multimodal texts. For this, a quantitative methodology based on non-experimental or ex post facto design with semi-structured and closed survey-questionnaire-type instruments has been selected. From a target population made up of Spanish teachers, 502 respondents selected using the non-probabilistic sampling technique applied the accessibility criterion. An *ad hoc* questionnaire has been drawn up consisting of 28 items digitised electronically using the survey platform of the University of Murcia. It has been structured in two blocks: the first aimed at establishing the sociodemographic and professional profile of the participants and the second at collecting data related to the teachers' beliefs regarding the work of the text commentary in class. The results show five professional profiles defined based on the implicit theories and the pedagogical model to which they are associated. It is also found that the majority declare that they align themselves with non-conservative didactic trends or approaches, centred on the student body and oriented toward the construction of critical knowledge. In this regard, manifest contradictions are detected between his implicit and explicit epistemological convictions. The findings of this study offer guidelines for the design of an effective and efficient argumentative text commentary formative proposal.

Keywords: argumentation, text commentary, teachers' beliefs, multimodal texts, teaching Spanish, Spanish foreign language

INTRODUCTION

The Paradigm of Teacher Thinking as a Key Element in Improving Education

The study of education improvement actions based on the paradigm of teacher thinking calls for an examination of their thoughts on learning and teaching, as such thinking has an influence on their day-to-day teaching decisions, both of a normal or innovative nature (Clark and Peterson, 1990; Schön, 1998). Its analysis is just as enriching as it is complicated, portraying the systematic structure of the professional background insofar as the explicit and implicit beliefs, values, ideas, and concepts involved in their teaching action. Given such cognitive polyhedron, its extensive investigative spectrum is based on key terms that can be grouped together in two major blocks of study aimed at discovering “the unique characteristics of the practical thinking of every teacher” (Pérez and Gimeno, 1988, p. 61). One block entails the corresponding constructs and conceptions, while the other the implicit theories and beliefs.

The construct (Kelly, 1955) is shaped by individuals with hypotheses and interpretations that afford them a predictive kind of reflective discernment with which to organise and contrast their progressive experience, in an adaptive manner, with the actions and situations entailed in the process itself, all of which is of particular interest in studying education agents through grid techniques (Pope and Ken, 1981). Today they are usually recognised under the term “concepts” (Pratt, 1992), which, in the education field, affect teaching and learning with positions that oscillate between the mere transfer of knowledge, up to the negotiation of meanings, and the fostering of knowledge generation (Hernández-Pina and Maquilón-Sánchez, 2011, p. 169).

Implicit beliefs or theories refer to the personal frameworks that serve as a starting point for making prioritised decisions. Belief may be understood in many different ways: as a conviction or subjective truth that is different from knowledge as an objective truth (Bunge, 2009); as a socially assumed reflective judgment (Ortega Gasset, 1976); and, as a psychological disposition to act in a certain way (Díez, 2017, p. 136).

As regards beliefs in education, the studies of Pajares (1992) are pioneering, pointing to the early acceptance of beliefs and to a great reluctance to change in adulthood, and, above all, they serve to assess the influence that they have on the decisions made by every teacher on the knowledge they impart, which range from the simple memorising of authoritative sources to the sophistication of critical and creating thinking.

Implicit theories, given their idiosyncratic nature, extend the framework of studies on beliefs, affecting principles, beliefs, goals, expectations, values, and practise models (Mitchell, 1995), and, due to their interdisciplinary study, there are connexions with social representations (Castorina et al., 2005). It is worth highlighting that, in the education context, the implicit theories are reconstructed based on educational knowledge gained through training and work actions (Marrero, 1993) in such a way that, compared with the knowledge representation of the explicit educational theories, the implicit theories entail knowledge

attribution, beliefs that the individual pragmatically assumes (Ros-Garrido and Chisvert-Tarazona, 2018, p. 99; Maldonado et al., 2019). As their acceptance reveals a semiological relationship between theory and action, where the intentional perspective is prioritised over linear consistency, it is important to consider the holistic role of teachers, who work not under instruction, but rather as creators of a sense of unity (Zabalza, 1987, p. 115):

teachers do not usually act with rigid and standard systems, but rather they interpret situations, create overall views on the indicators or clues observed in the classroom and act accordingly. It is not, therefore, the case of teachers in a space of certainty and automatic connexions between thought and actions, but rather one of teachers in a context of permanent hypotheticality (“given the situation, I think the best thing I can do is...” [sic].

As such, the scientific consideration on the thinking of teachers in their intellectual intent and complexity allows us to delve into the reason for being of educational processes and in the “re-conceptualisation of educational research with and based on teachers” (Jiménez-Llanos and Feliciano-García, 2006, p. 113). Therefore, it is important to consider the psychological determinants (implicit theories, values, beliefs) and environmental determinants (resources, external situations, administrative limitations) (Fandiño, 2007), as well as the difficulties involved in this study, particularly that relating to the verbalisation of their professional thinking due to two reasons: firstly, teachers tend not to communicate it in a reflective manner, but rather through intuition; and, secondly, their ideas and actions often respond to pragmatic and sentimental occurrences, which are difficult to compare with logical-causal categories. Questionnaires are useful instruments for measuring implicit beliefs and theories (Schommer-Aikins, 2004). Specifically, Likert scale questionnaires with multiple choice answers allow the key areas on the subject in question to be covered (Serrano, 2010, p. 274), provided the constructs and dimensions for their matrix are suitable for the assessment purpose sought (Vizcaíno et al., 2015).

The Implicit Theories of Teachers on the Argumentative Commentary of Multimodal Texts

Addressing the analysis of the implicit theories of teachers on the argumentative commentary of multimodal texts focuses research into what Grossman (1990) understands as “content knowledge” (substantively specialised in a subject matter) and “pedagogical content knowledge” (pedagogical knowledge on the teaching of the specialised subject matter). In this regard, considering their background and current status of this scientific issue is of interest.

Inspired by pioneering studies on the analysis of teacher thinking centred on the teaching of language (Pearson and Stephens, 1994; Woods, 1996; Borg, 2003; White and Bruning, 2005), Spanish studies have been conducted over the last two decades that analyse the interaction of beliefs, assumptions and knowledge as to such regard (Cambra et al., 2000; Ballesteros et al., 2001; Munita, 2013). As for teacher thinking centred on the teaching of Spanish, many more studies have been conducted

on Spanish as a native language than on Spanish as a foreign language. There is a thriving flow of research in Spain and Latin America on the thinking of teachers with regard to academic writing (Carlino, 2008; Ortiz et al., 2009; Martins, 2012, 2014; Capomagi, 2013; Castell and Mateos, 2015; Giraldo, 2015; Flores, 2018; Bigi et al., 2019; Cordero and Carlino, 2019; Gordillo, 2019) and on reading literacy and comprehension of academic texts (Makuc, 2008; Suárez, 2015; Mojarro-Delgadillo and Alvarado-Nando, 2021), where diverse implicit theories are usually used: the linear or decoding theory (Laberge and Samuels, 1994), the generative cognition theory (Chomsky, 1974), the interactive or procedural theory (van Dijk and Kintsch, 1983; Rumelhart, 1997); the transactional theory (Goodman, 1994; Rosenblatt, 1996; Mendoza, 2001).

On the other hand, with regard to analytical studies on teacher thinking relating to Spanish as a foreign language, there are only a few publications in different countries (Minervini, 2004; Usó, 2012; Almeda, 2014; Santos and Alexopoulou, 2014; Zhiying, 2018; Domínguez et al., 2019). More specifically, there is an extreme lack of studies that regard beliefs relating to the teaching of argumentation (Martínez, 2018) and there is no known record of studies on the exploration of teacher thinking related to the argumentative commentary of multimodal texts, except those conducted by the research team currently co-led by the authors of this article.

Therefore, the scientific initiative on which this article provides knowledge, opens up a new line of research on an international level. To date, in terms of the paradigm of studying teacher thinking, only one analysis has been published on the teaching needs of professional training in the teaching of argumentation in text commentaries (de Vicente-Yagüe et al., 2019), as well as an exploration of academic teaching customs on the suitable methodology for undertaking informal argumentation in text commentaries that affect the textual typologies, negotiations with students, the procedural sequence of oral and written tasks, and the revision and assessment strategies of said commentaries (Caro et al., 2018).

The exploratory research set out in this document is consistent with a dialogic model of text commentary in its rhetoric architecture sustained over the logical course of informal argumentation (Caro and González, 2018), given that the latter is conceived within a multimodal aspect, both generic and digital, that is inseparable from the functioning of the discourse and that is evident in its linguistic modalisation and is illustrative of any kind (Amossy, 2008, p. 12). Such model starts by looking into the epistemic culture of two inherent activities of the commentator, the interpretation of texts and intertexts, and the formulation of ideas and the writing of the commentary with the hypothesis-arguments-conclusion sequence. In this regard, it is important to be aware of the problem that may reveal the thinking of teachers based on resolute beliefs and customs that, responding to authoritative and mimetic pedagogical models, are reluctant to grant readers inventive power to generate knowledge. Therefore, using exploratory questionnaires on academic customs, beliefs and demands of Spanish language teachers for such purpose, prior validation and analysis of their internal consistency, is a priority objective (Caro et al., 2021).

The professional thinking profile is explained with regard to the classification of teacher responses according to the criteria stipulated relating to the kind of implicit theory and the pedagogical model associated with it. A recurring classification is that of Hargreaves and Goodson (1996, p. 4–19), entailing the following profiles: classical (shared technical culture as organisational self-regulation to provide good customer service), flexible (collaborative culture in communities of professional practise improvement), practical (culture that dignifies practise as a source of knowledge), expanded (connective culture of theory and practise and wide-ranging collective planning), and complex (culture committed to solving problems and uncertainties). However, in view of the innovative expectations with respect to traditional pedagogies on text commentary and argumentation, it is preferable to start from a classificatory model that differentiates with evolutionary clarity the main implicit theories of teaching, such as the one referred to by Marrero (1993, p. 251–255), in these terms schematised by Beltrán (2019, p. 205). See **Table 1**. We have chosen Marrero's (1993) classificatory model in order to transpose into it the didactic characteristics of text commentaries and argumentation which, guided by our previous research on the subject (Caro and González, 2012, 2015, 2018; Caro, 2015), we specify below in a systematic way in correlation with the type of theory and its corresponding pedagogy. We intend this new classificatory model to serve as a preliminary organiser for the discernment of professional profiles through the teaching responses on the subject (**Table 2**).

Alongside the implicit theories that are clarified regarding the teaching of argumentative textual commentary, it would be appropriate to critically consider the trends that currently mark its teaching ethos, as teachers are in a transition between the parameters in which they were trained and the need to train for the challenges that the current situation poses. As such, just as a few decades ago, the idea of autonomous professionals who chose the most appropriate methods for their students evolved toward that of associated professionals who carry out their professional development collaboratively to face challenges and uncertainties with daily work in learning communities; in the Knowledge Society, we are moving toward a “post-professional age” where inter-institutional permeation is growing through digital communication and the commercialisation of education is accentuated with client-like relationships that, unfortunately, devalue teachers by denying them their autonomy (Montero and Gewerc, 2018) and that objectify their expectations of competence innovation in consumer banners (Caro, 2017). For this reason, the need to focus teacher thinking on authentic and multicultural communicative expectations (Dorfsman, 2018) that differ from the neo-conductive market impostures, and that advance with emancipatory theories of a critical pedagogical model, is currently gaining scientific momentum.

In the current context, the relevance of online educators and the need to promote the development of their digital competence has increased, as recognised by the TALIS 2018 report (OECD, 2019, p. 12). Educational institutions should approach this challenge with a community commitment that does not reduce it to mere instrumentalist executive work according to neo-behaviourist taxonomic competencies. In this sense, we consider

TABLE 1 | Generic classification model of implicit teacher theories (Marrero, 1993).

Type of theory	Characteristics	Pedagogical model with which it is associated
Dependent	Teacher-guided and teacher-directed teaching, so that the same pace of learning is maintained for all students; it is thought that, if the teacher does not teach, students are not capable of learning on their own; a distant attitude toward students and a conception of the school outside social and political conflicts are postulated.	Traditional
Productive	Teaching is the pursuit of results and the enhancement of effectiveness in teaching and learning. Teaching by objectives becomes relevant.	Technical
Expressive	It recognises student activity as the core element of the teaching and learning process. Permanent indicators include experimentation, education for life, the number of activities to be carried out and the permanent occupation of students.	Active
Interpretative	Pedagogy centred on students (their needs, resources, and learning processes) and an interpretative attitude (search for more or less formalised explanations of teaching practises) coincide. It stresses the importance of processes over outcomes and emphasises the communicative aspects of teaching.	Constructivist
Emancipatory	It has a strong moral and political character in a broad sense. The concern for the contextual legitimisation of certain objectives and contents of teaching, the link between teaching practises and the political-social framework of the actions of students and teachers accentuate the critical character and the corresponding emancipatory intentionality.	Crítico

the opportunity to work on heuristic argumentation in authentic situations of shared learning to activate the hypothetical thinking of teacher action-research with the strategic use of digital media (Caro, 2018). Furthermore, we undertake the line of research in the argumentative commentary of multimodal texts where all students, in singular or shared leadership, can expand their critical interpretation and their emancipatory hypotheses. This line has a bearing on a key issue for such pedagogical renewal, as text commentary as an academic discursive genre has been one of the bastions on which teaching based on the transmission of knowledge has survived (Bordieu, 1989, p. 28) by reducing it to dissertation obedient to the principle of authority and the ideological control it entails (Foucault, 1973). In fact, traditional pedagogical models with this theoretical profile, replicated in textbooks, are still applied in schools today (López, 2008; Lluch and Serrano, 2016; Rodríguez-Martínez, 2016).

Therefore, the study of implicit teacher theories on the teaching of argumentative commentary of multimodal texts has to start from an enquiry into their epistemological beliefs about argumentation and text commentary, the assumptions of which are not usually made explicit in teaching practise nor in the teaching models that support them, although they are fundamental to discern two psychological models of understanding the production of knowledge: the mimetic model of “stating” knowledge and the creative model of “transforming” knowledge (Scardamalia and Bereiter, 1992).

Likewise, studying their professional expectations about the use of ICTs in the teaching of argumentative commentary on multimodal texts will allow us to gather valuable information regarding the teachers’ vision of technology, allowing several aspects to be contrasted: one is whether they replicate the utopian theses disseminated by educational institutions (neutral and controllable tools that procure prosperity) or, conversely, whether they denounce them in their dystopia (corrupt force that will destroy humanity); the other is whether they possess implicit emancipatory theories in this respect, according to the characteristics pointed out by Castañeda et al. (2018, p.

13) on the emergent network pedagogy as “increased reflective practise,” exercising social engagement with personal learning environments specific to the current technological context.

The outcome of the teacher thinking analysis on this subject will result in proposals on their professionalism (Englund, 1996) or the diagnosis of the quality of their work, taking into account their method and style, and the scientific-technical standards that serve as a framework, all with the aim of reflecting objectively on their improvement through innovative processes.

In line with the aforementioned scientific bases, the objectives of this research are established below:

General Objective

To interpret the pedagogical profile of Spanish language teachers in their implicit theories on the development of argumentation in the commentary of multimodal texts based on the analysis of their corresponding teaching beliefs.

Specific Objectives

1. To describe teachers’ epistemological convictions on the definition of text commentary (SO1).
2. To identify teachers’ preferences on the didactic modalities of text commentary (SO2).
3. To determine teachers’ judgments on the generic effect of argumentation (SO3).
4. To explore pragmatic teaching models on the verbal communication of argumentation (SO4).
 - 4.1. To identify argumentative models in expressive activity (SO4.1).
 - 4.2. To recognise argumentative models in the comprehensive activity of text commentary (SO4.2).
5. To analyse teaching assumptions regarding the value that the commentator should place on the wording of the text (SO5).
6. To discover teachers’ attributions regarding the argumentative key of commentary (focal point of enquiry/matter of controversy) (SO6).

TABLE 2 | Classification of implicit teacher theories (Caro et al., 2021).

Type of theory	Characteristics	Pedagogical model with which it is associated
Dependent	<p>Logocentric teaching based on the transmission of knowledge from a neoclassical positivist paradigm:</p> <ol style="list-style-type: none"> (1) It conceives text commentary as an individual representation of the meaning of the text according to the author's ideas. (2) It attributes personal critical argumentation to few genres, but not to commentary. (3) It gives authoritarianism to teachers: their monologue instructs the contents without classroom discussion and without attending to diversity; they make a pyramidal selection of texts (canon); they evaluate the performance of students based on a controlled pattern of concepts and behaviours. (4) It affords passivity to students: reproducing information; practises entail the application of theory. (5) Conservatism (presumed school neutrality). (6) It rejects Information and Communication Technologies (hereinafter, "ICTs"). 	Traditional
Productive	<p>Functional teaching for training effectiveness from a positivist paradigm of technical rationality:</p> <ol style="list-style-type: none"> (1) It conceives text commentary as an individual reconstruction of the meaning of a specialised text and of the author's ideas. (2) It attributes personal critical argumentation to few genres, but not to commentary. (3) Teachers as facilitators of the contents to achieve the objectives: they make a pyramidal selection of texts (canon) and evaluates the students' results from their programmed expert disciplinary control of concepts and skills. (4) Students as consumers of knowledge models; their practises are of a behaviourist nature in terms of theory application. (5) Conservatism (presumed school neutrality). (6) The use of ICTs enhances information. 	Technical
Expressive	<p>Spontaneous teaching that promotes learning for life from a humanistic experiential paradigm:</p> <ol style="list-style-type: none"> (1) It conceives text commentary as an individual reconstruction of the meaning of the text and of the author's ideas, based on which they give a personal opinion. (2) It attributes personal critical argumentation to few genres, including the commentary. (3) Teachers as facilitators of content learning through activities: they take into account the students' tastes in the textual selection (educational canon) and evaluates their performance according to the corresponding experiences. (4) Students as the centre of the teaching-learning process based on the meaningful motivation of the actions. (5) Spontaneous dialogue (no promotion of critical ideology). (6) The use of ICTs enhances feedback. 	Active
Interpretative	<p>Cognitive teaching that meets the needs, resources and learning processes from a humanistic interpretative-symbolic paradigm:</p> <ol style="list-style-type: none"> (1) It conceives text commentary as an individual reconstruction of the meaning of the text (hypertext) from the author's perspective or thesis and as a construction of the critical sense of the commentator through argued hypotheses. (2) The personal critical argumentation is multimodal (including the commentary). (3) Reflective teachers: critical action-research in the teaching-learning processes: it takes into account students' tastes in textual selection (educational canon) and evaluates their procedural performance according to the competences demonstrated in the tasks; interprets teaching practises. (4) Reflective students: the centre of the competence-based teaching-learning process, constructing and self-assessing in a procedural way their tasks. (5) Critical dialogue (ideological perspective). (6) The multimodal use of ICTs enhances feedback. 	Constructivist
Emancipatory	<p>Democratic teaching that meets the needs, resources and learning processes based on and organised under a critical emancipatory paradigm:</p> <ol style="list-style-type: none"> (1) It conceives text commentary as an individual or collective reconstruction of the contextual meaning of the text (hypertext) from the author's perspective or thesis and as a construction of the critical sense of the commentator through argued hypotheses. (2) The personal critical argumentation is multimodal (including the commentary). (3) Teachers as intellectual transformers committed to the sustainable development of the community: critical and meta-reflexive research-action in the teaching-learning processes: they take into account the tastes and ingenuity of the students in the textual selection (educational canon) and in the didactic proposals for social improvement; they evaluate the students' procedural performance according to the competences demonstrated in the contextualised performance of the individual and collaborative tasks. (4) Students as autonomous, responsible and transformative people committed to the sustainable development of the community: the centre of the teaching-learning process based on competences; they learn theory for practical problem solving; construct personal theories from their own research and reading reflection; generate knowledge and experience social empowerment; undertake a self-evaluation in a procedural manner on their performance in the key and global competence framework. (5) Innovative critical dialogue (transformative ideological perspective on equity and diversity). (6) Multimodal communication with ICTs enhances social emancipation. 	Crítico

7. To enquire into teachers' beliefs regarding teaching resources for argumentative text commentary (SO7).
 - 7.1. To explore teachers' judgment on the methodological suitability of textbooks for teaching argumentation in text commentary (SO7.1).
 - 7.2. To explore teachers' expectations of ICTs in argumentative commentary (SO7.2).
 - 7.3. To explore teachers' pedagogical preferences on the design of didactic guides for argumentative text commentary (SO7.3).
8. To discern teachers' implicit theories regarding the argumentative commentary of texts by means of a contrastive analysis of the answers to the exploratory questionnaire according to typological parameters (SO8).

MATERIALS AND METHODS

A quantitative methodology based on a non-experimental or *ex post facto* design with semi-structured, closed-ended and survey-questionnaire instruments has been selected. Implicit teacher beliefs or theories about the development of argumentation in text commentary were assessed by means of a Likert-style questionnaire, the design of which was previously inspired by the literature on quantitative research on teaching beliefs (Inganzo, 2010; Castañeda and Ortiz, 2017; Vizcaíno et al., 2018). For the analysis of the internal consistency of this instrument, the percentage of reliability was measured in the three thematic blocks (teacher customs, beliefs and academic demands) in which the questionnaire was structured, obtaining fairly acceptable results. The analysis of construct validity was also carried out by checking the correlation matrix to determine the degree of variable correlation.

Sample Population

With a target population made up of teachers of Spanish as a Native Language (hereafter referred to as SNL) and Spanish as a Foreign Language (hereafter referred to as SFL), the teaching staff in the area of Spanish Language and Literature in Spanish-speaking countries at different educational stages and the university teaching staff of SFL in different countries have been established as the sample framework (or study population). The subjects were selected using the non-probabilistic sampling technique, applying the criterion of accessibility. A total of 502 teachers took part in the questionnaire, 390 of which were SNL teachers and 112 SFL teachers.

In order to define the characteristics of the selected population, a series of items related to the socio-demographic and academic data of the teachers surveyed were integrated (Table 3).

Data Collection Instrument

We have developed an *ad hoc* questionnaire consisting of 28 items (see Annex 1), geared toward to SNL and SFL teachers. This is an instrument of open and closed questions digitised

TABLE 3 | Socio-demographic and professional profile of participants.

Age	Under the age of 20	6	Mean:
	21–30	88	43.17
	31–40	138	
	41–50	114	
	51–60	120	
	61 or more	28	
Gender	Male	142	
	Female	354	
	Don't know/not answered	6	
Academic training	Diploma	40	
	Licentiate degree/degree	200	
	Master's degree	92	
	Phd	162	
	Don't know/not answered	8	
Specialised studies undertaken	Philology	144	
	Language and literature teaching	62	
	Education or teaching	70	
	Geography and history, and philosophy	26	
	Sociology	4	
	Communication	20	
	Special education	4	
	Don't know/not answered	172	
Teacher experience	1–5 years	128	Mean:
	6–10 years	78	15.62
	11–15 years	54	
	16–20 years	68	
	21 or more	168	
	Don't know/not answered	6	
Teaching given	Pre-primary education	2	
	Primary education	60	
	Compulsory secondary education (ESO)	20	
	Further education (<i>Bachillerato</i>)	98	
	University degree	92	
	Postgraduate degree	40	
	Does not work on text commentary in class	150	
	In different educational levels	18	
	Don't know/not answered	22	

electronically using the University of Murcia's survey platform¹. In terms of its structure, it is made up of two blocks:

Block 1: Academic and Socio-Demographic Data

This allows the research to be contextualised and the characteristics of the selected population to be defined. This is made up of a series of items related to the academic data of the teachers surveyed, such as for example, the teaching given, studies undertaken, age, among others. In total, 5 items.

Block 2: Teachers' Academic Beliefs

This block collects data related to the academic beliefs of teachers regarding text commentary work in class with variables such as type of analysis and interpretation procedure, dialogue and argumentation or scientific cognition, individual/collective and oral/written modality, use and function of argumentation, beliefs about commentators, teaching manuals, use of ICTs, appropriate

¹<https://encuestas.um.es>.

procedures for elaborating teaching guides, among others. In total, 23 items.

We have undertaken the reliability analysis and the beginning of the validation process of the data collection instrument, i.e., the analysis of the internal consistency of the questionnaire scale applied in a pilot test (28 items).

RESULTS

For data processing and analysis, the statistical analysis programme IBM SPSS (version 27) was used. The Cronbach's Alpha coefficient obtained in this block corresponds to 0.8327. Consequently, the reliability of the measurement instrument comprising the set of items, Likert scale 4, with the following levels: "Strongly disagree," "Disagree," "Agree," "Strongly agree," is high.

The analysis of grouped relative frequencies has allowed us to explore the distribution and extent of beliefs about argumentative text commentary extracted from the answers given by respondents (**Figure 1**).

The 23 items corresponding to block 2 of the survey related to the teachers' academic beliefs on argumentative text commentary were as follows:

- Item 6. Text commentary should be first and foremost a procedure of "analysis, interpretation and evaluation of textual data"
- Item 7. Text commentary should be first and foremost a procedure of "dialogue with texts in which a personal position is argued"
- Item 8. Text commentary should be first and foremost a procedure of "scientific cognition that uses language in an interdisciplinary way"

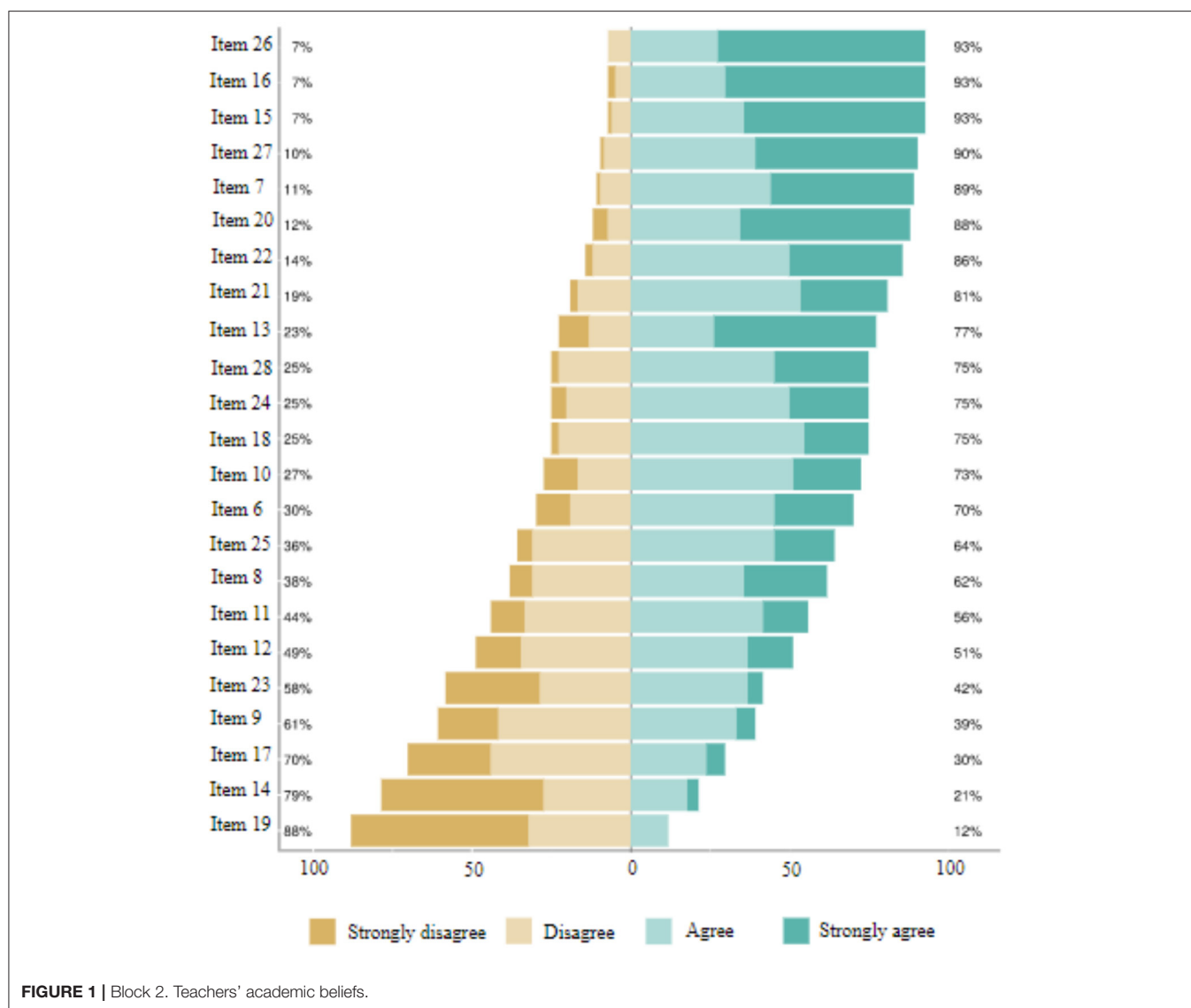
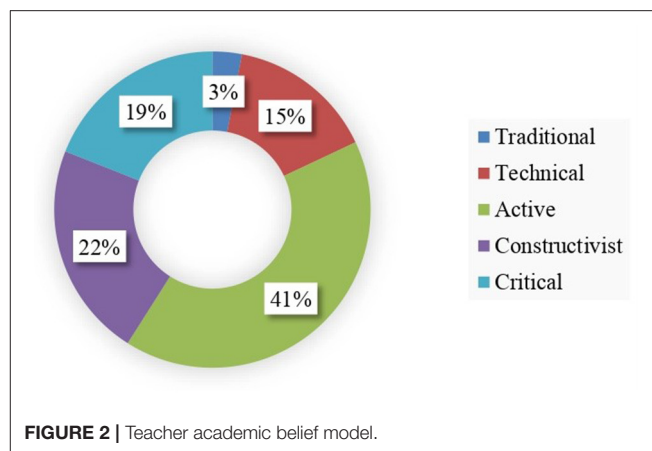


FIGURE 1 | Block 2. Teachers' academic beliefs.

- Item 9. The preferred mode of text commentary should be “individual and oral”
- Item 10. The preferred mode of text commentary should be “individual and written”
- Item 11. The preferred mode of text commentary should be “collective and oral”
- Item 12. The preferred mode of text commentary should be “collective and written”
- Item 13. Argumentation can appear in any type of text
- Item 14. Argumentation can appear only in academic and opinion texts
- Item 15. Argumentation serves to provide a space for discussion between two or more perspectives, ideologies, etc.
- Item 16. Argumentation serves to express a personal or collective position on an issue.
- Item 17. The argumentation of the text commentary focuses primarily on the recognition of explicit aspects
- Item 18. The argumentation of the text commentary focuses above all on the interpretation of the implicit aspects
- Item 19. In text commentary, “what is stated” in the text is presupposed as something unquestionable
- Item 20. Text commentary presupposes “what is stated” in the text as subjective and subject to critical review
- Item 21. The commentator makes incognito enquiries in order to propose a solution and to argue their defence
- Item 22. The commentator chooses controversial issues in order to argue, dispute, deliberate and engage in dialogue with arguments
- Item 23. The textbooks provide teaching material in line with the teaching methodology I consider suitable for argumentation in text commentary
- Item 24. The use of ICTs could improve argumentative skills in text commentary
- Item 25. Focusing the analysis of the text on the understanding of the literal and implicit contents of the author’s intention
- Item 26. Contrasting the author’s intention with the commentator’s perspective in order to promote critical thinking
- Item 27. Providing guidance to the commentator on how to organise the sections and the writing of the argumentative commentary
- Item 28. Giving commentators the freedom to use their critical sense with their own contextualised logic and style.

In order to identify teachers’ profiles according to their implicit theories, new categorical variables were established by grouping the contiguous values of the variables studied in block 2 of the questionnaire.

From the resulting Gaussian bell, five categories were obtained which described the generic model of classification of teachers’ implicit theories on the teaching of text commentary and argumentation described above (see **Figure 2**). The descriptive statistical analysis showed the distribution of these categories according to the percentages obtained. It became clear that a minority of respondents (2.8% declared themselves as traditional and 15.1% technical) positioned themselves within the parameters that define a conservative, teacher-centred



teaching model, focused on the transmission and reproduction of the knowledge given. In contrast, the highest percentage of participants (41.4%) in the study was concentrated in the expressive-active model, based on spontaneous teaching that promotes learning for life from a humanistic experiential paradigm (**Figure 2**). This may be related to their usual practises of teaching programming and designing classroom projects, where the possibility of specifying the active methodologies for competence training required by the Bologna process is often resolved with simple expressions such as “active and participatory methodology” without specifying the specific systems of task- and project-based learning. This also shows the scarce scientific-academic impact that institutional proposals for initial and ongoing teacher training focused on educational methodologies have on teachers’ beliefs, as they do not use them unless they change their habit of relying on the use of manuals for teaching to design projects that favour constructivist and emancipatory learning. It may also be related to the fact that their licentiate degree or degree training in Hispanic Philology and related subjects does not tend to focus on issues of linguistics applied to education.

As for the comparative analysis of the epistemological and didactic convictions linked to the commentary of argumentative texts between teachers of Spanish as a native language (390 respondents) and those of Spanish as a foreign language (112 participants), the only relevant difference was found in the categories related to the constructivist and critical pedagogical model. It was the SFL teachers who registered a greater awareness of democratic teaching based on the critical emancipatory paradigm. This may be related to the methodological parameters established by the Council of Europe in the Common European Framework of Reference for Languages (CEFR) and, specifically, in the Curricular Plan of the Instituto Cervantes, a transcript of the CEFR, for the teaching of Spanish in the world, whose methodological models of study are used with institutional imperative in the teaching of foreign languages. It also introduces SFL teachers to the usual didactic practise that promotes task-based learning and innovative teaching strategies linked to innovative educational materials hosted on institutional training promotion websites.

In this sense, the crossover between the variables linked to years of work experience and the model of the academic beliefs of teachers has allowed us to discover that it was the participants with <6 years of teaching experience (in both teaching modalities) who presented a more innovative professional profile, as shown by the exponential trend line in **Figures 3, 4**. These results may be related to the context of initial training which, for more than a decade, has been offered by the Master's Degree in Teacher Training for Secondary Education in Spain, as, in comparison with the previous professional training system, the Teacher Training Course (*Curso de Aptitud Pedagógica*, CAP), it has improved knowledge of innovative methodologies of a constructivist and emancipatory type by having in its academic programme subjects exclusively for this purpose of pedagogical renovation. This upward innovative dynamic was also observed when looking at the university education of the teachers surveyed: the higher the academic level (Master's and PhD), the greater their commitment was to educational proposals that sought to move away from traditional teaching paradigms. For the same reason, the university teaching staff who teach and guide Master's theses have improved in this respect in their ongoing training and by tutoring educational innovation and research projects, the backbone of which are these methodologies. This aspect was particularly relevant in postgraduate studies linked to the area of Language and Literature Teaching.

Having established the general framework of the models of teachers' academic and pedagogical beliefs, the results related to each of the specific objectives formulated were specified:

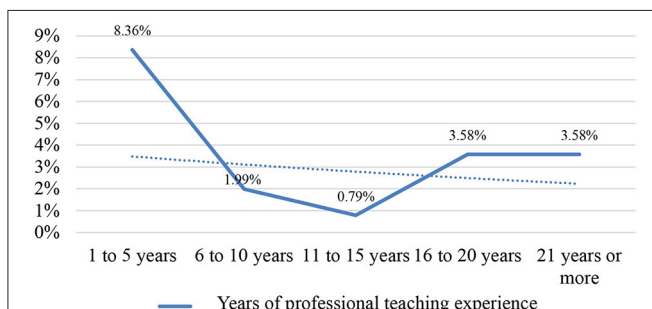


FIGURE 3 | Critical pedagogical model.

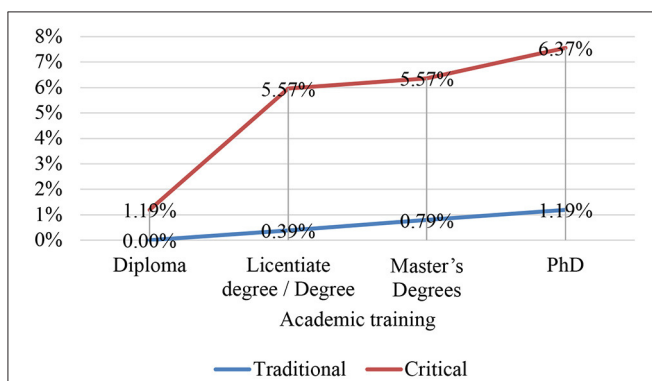


FIGURE 4 | Trends in pedagogical models.

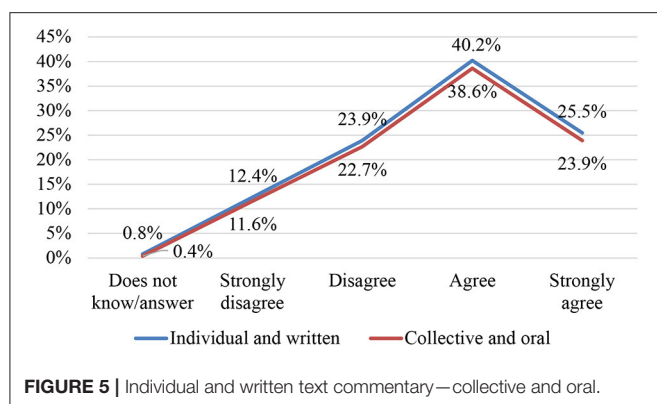
SO1. To Explore Teachers' Epistemological Convictions on the Definition of Text Commentary: Items 6, 7, and 8

In relation to the definition of text commentary, the answers are mainly concentrated on the interpretative aspect of the students (77.29%). Respondents understand that the individual representation of textual meaning according to the author's ideas is not enough, but rather that emphasis should be placed on the student's personal position vis-à-vis the text and, in particular, on the construction of critical meaning through the hypotheses generated, discarded or corroborated during the process of reception. From this perspective, text commentary has been understood as a complex, dialogical, creative, constructive and interactive cognitive process through which the individual participates with full autonomy in different socio-cultural contexts. In this sense, commentators use their cognitive-linguistic skills, activate their analytical thinking to read between the lines, identify and question the underlying ideologies, unravel what is implicit in the statements, argue in a well-founded and contrasted way, etc., through a dialogue between the appellative structure of the text, the author's intentions and the receiver's knowledge of the world and life experiences. This advanced conviction that corresponds to the interpretative theory of reading and its responsive writing in the commentary is consistent with a constructivist pedagogical model that has been consolidated as a belief based on the dialogical conception of communicative competence in the construction of knowledge that teachers of different educational stages have been progressively assimilating in their teaching work and particularly since the LOE (Spanish organic law of education) came into force 15 years ago.

SO2. To Explore Teachers' Preferences on the Didactic Modalities of Text Commentary: Items 9, 10, 11, and 12

The answers given by the participants to the questions related to the oral or written aspect and to the individual or collective nature of the text commentary, made it possible to identify their beliefs in this regard. The identical behaviour of the percentages represented showed two very clear preferences according to the sequential process of the argumentative text and its teaching purpose (**Figure 5**):

1. 65.7% of those surveyed opted for written text commentary when it came to individually developing the cognitive and linguistic-textual skills and abilities linked to argumentative competence. From this perspective, they have revealed, on the one hand, to possess a graph-centric image by understanding that the written discourse genre favoured a greater degree of formality, planning and critical distancing; and, on the other, didactically perceiving the text commentary as the final product produced and not as a complex constructive process in which interpretative mechanisms come into play, contrasting opinions, as well as cognitive and metacognitive skills associated with analytical thinking and, in general, argumentative competence. Although in their theoretical



prolegomena they declare the action of commenting on texts from a dialogical and constructivist approach in accordance with the legal expectations of the twenty-first century curriculum, specific teaching decisions emerge in their teaching practise that contradict this ideal by preferring the written mode of commentary over the oral mode for reasons inherent to implicit theories dependent on the traditional pedagogical model that promotes writing as a product, not as a process. Such a dominant response thus proves the stubborn persistence of the so-called “hidden curriculum” in teachers’ implicit theories that disables the proper management of competence-based education.

2. A significant proportion of respondents recognised that textual interpretation by each student should be accompanied by communicative exchange processes in which different points of view are contrasted in order to construct and share a more complex and plural reading (or *archi-reading*) within the interpretative community of the classroom. The emphasis was therefore on transactional exchanges to expose, discuss and reformulate perspectives, ideological positions or simply hermeneutical tasks.

As such, their explicit theories once again clash with their implicit theories, while recognising the benefit of working on the interpretation of texts in communicative acts with group interaction, the responsibilities they grant to text commentary remain individualistic and silent.

SO3. To Explore Teachers’ Judgments on the Generic Effect of Argumentation: Items 13 and 14

Although it is true that a third of the participants admitted that argumentation belongs exclusively to academic and opinion texts, a judgment that reduces the intellectual scope to the genres traditionally recognised as argumentative due to their verbal explanations (stating and defending author’s theses), it has been observed that the majority response corresponds to the conviction that it could be found in any discursive genre or textual modality. Such a belief opens the way to the possibility of the emancipatory exercise of argumentation in class, given that it directly affects both the processes of interpretative

reading attentive to the implications of the commented text and the criteria for the selection of texts for argumentative commentary, and its predisposition to offer an environment rich in written, oral and multimodal material. Getting students used to dealing with the greatest variety of texts associated with the different contexts of social and cultural life (personal, public, educational, and professional) facilitates specific and different ways of cooperating with the text and of constructing their personal critical point of view in their reading reception writing through commentary, which is no longer an explanatory gloss on the text but an expansion of its meaning in the horizon of each reader’s expectations.

With regard to the analysis of the variable of argumentative expression, the teaching judgment prefers to grant argumentation a pragma-dialectical communicative purpose, as it understands that its discursive utility gives rise to spaces for discussion and contrast between different points of view, ideological convictions and beliefs. Therefore, teachers should promote the transformation of the classroom into a space for social interaction between peers, where dialogue after reading is encouraged as a way of:

- Expressing and respecting personal and collective opinion on a given topic.
- Exchanging views, so that no single interpretation is forced on others.
- Consensual negotiating the meaning of the text within the class-community.
- Reaching deeper levels of understanding and interpretation based on individual contributions.

However, the prevailing belief of limiting argumentation to a dialectical model implies that it can only exist in productions or verbal receptions where there is confrontation of criteria and decision-making in favour of or against them, thus neglecting those argumentative initiatives of a focal model that enable the shared construction of knowledge and, therefore, curtailing the treatment of textual multimodality. Moreover, it limits critical commentary to polemical verbal acts subject to verdict and neglects its heuristic possibilities for classroom research.

OE4. To Explore Pragmatic Teaching Models on the Verbal Communication of Argumentation: Items 15, 16, 17, and 18

The majority of the teachers consulted all believe that argumentation is used to express both singular and plural subjective perspectives and, therefore, they accept without reservation its critical purpose, an aspect due to which a personal critical understanding of the exercise of text commentary can also be accepted, which is interesting for both individual and cooperative oral or written expression actions (Tables 4, 5). It can also be noted that, by referring to the term “position,” in their acceptance of the critical sense, they continue to maintain a dialectical approach to their discursive practise aimed at defending ideas in the face of different or adverse positions.

In relation to the answers given to questions 17 and 18 regarding the variable of argumentative comprehension, it should

TABLE 4 | Item 15. Argumentation as a space for discussion.

		Frequency	Percentage	Valid percentage	Cumulative percentage
Valid	Does not know/answer	4	0.8	0.8	0.8
	Strongly disagree	6	1.2	1.2	2.0
	Disagree	42	8.4	8.4	10.4
	Agree	178	35.5	35.5	45.8
	Strongly agree	272	54.2	54.2	100.0
	Total	502	100.0	100.0	

TABLE 5 | Item 16. Argumentation serves to express one's position on an issue.

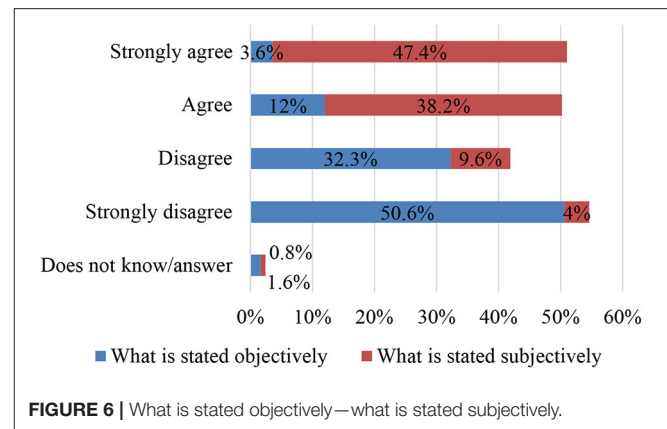
		Frequency	Percentage	Valid percentage	Cumulative percentage
Valid	Does not know/answer	4	0.8	0.8	0.8
	Strongly disagree	6	1.2	1.2	2.0
	Disagree	36	7.2	7.2	9.2
	Agree	178	35.5	35.5	44.6
	Strongly agree	278	55.4	55.4	100.0
	Total	502	100.0	100.0	

be noted that, based on pragmatic assumptions, the concept of implicature was developed due to its essential value in the argumentative processes of text commentary. The recognition of information presented by the sender in a non-explicit way becomes the focus of argumentation in text commentary. This is the perception of the teachers surveyed at least: almost 75% of respondents stated that they agreed or strongly agreed with this statement. This teaching conviction justifies the design of contextualised argumentative practises where students are trained in strategies to identify the implicit and select relevant information in oral and written texts.

However, it should be noted that a significant percentage of participants (33.9%) indicated that recognising what is explicit should be the focus of argumentation. Most are teachers with professional experience of between 1 and 5 years. It becomes clear, in any case, that argumentative discourse consists of an inferential process based on investigating the relationship between what is implicit and explicit in the text.

SO5. To Explore Teaching Assumptions Regarding the Value That the Commentator Should Place on the Wording of the Text: Items 19 and 20

With regard to the analysis of the value of the wording of the text, the teachers surveyed assume that, in text commentary,

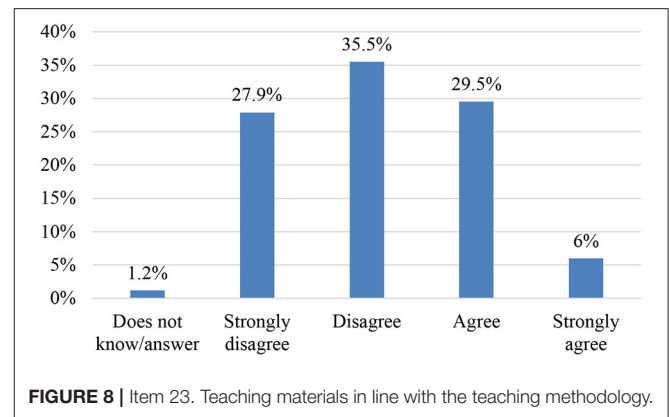
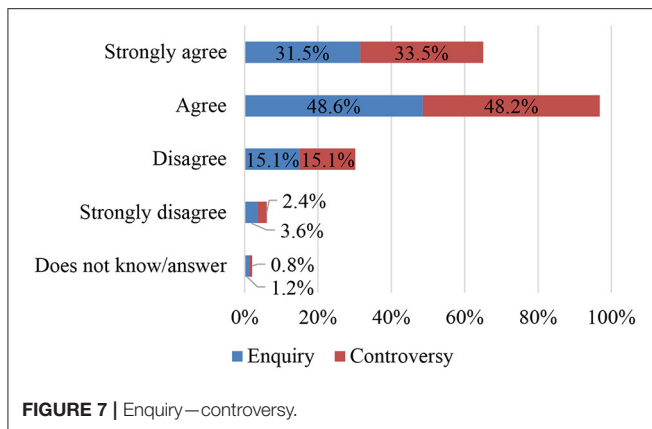
**FIGURE 6 |** What is stated objectively—what is stated subjectively.

students should not accept the postulates, premises, opinions or ideological positioning of the author without questioning such. What is explicitly stated in the text needs to be checked and corroborated with other opinions, theories or reasoning. Thus, over 82% of the participants indicate that they disagree with the fact that what is stated in the text should be considered as something indisputable, rather it should be considered as something subjective that is subject to constant critical review (Figure 6). Such a consideration serves to empower the interpreting reader's horizon of expectations in the classroom when establishing a dialogue *between peers*—between two subjective perspectives—with the text, which shows a major advance with respect to the traditional profile of the commentator whose mission is to extract only the authorial sense of the text, as fostering the interpreting reader's power of judgment is very beneficial for leading the way to the implicit emancipatory theory incited by the genuinely democratic practise of commentary in the classroom.

SO6. To Explore Teachers' Attributes Regarding the Argumentative Key of Commentary (Focal Point of Enquiry/Matter of Controversy): Items 21 and 22

In line with the answers given in question 15, teachers' implicit theories tend to attribute to argumentative commentary the defence of a dialectical position which justifies the text and the commentary as a matter of controversy: the purpose of argumentation is to convince interlocutors to adopt a particular point of view or specific opinions. To do so, the starting point must be a communicative situation in which opposing positions are discussed. Based on this belief in controversy as a principle of dialogue, the study participants (over 81%) emphasise the students' selection of controversial issues in relation to which they can debate and deliberate by means of arguments to defend their own theories and refute those of others (Figure 7).

It is incongruous that, with this being their belief regarding the argumentative key of commentary, they attribute preferences of

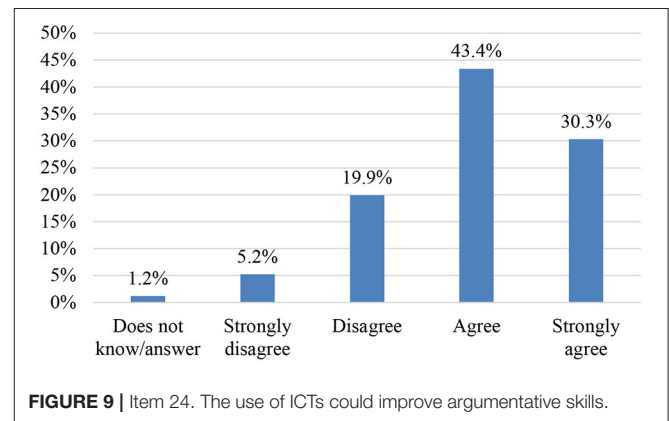


a similar percentage to the key corresponding to the focal point of enquiry, since their fundamental communicative strategy is different to the dialectic one, as they are reasoned commentaries that are contributed in a collective heuristic paper for the shared construction of valid knowledge for the community, thus their nature is predominantly synergic and includes dialectic arguments with fewer resources to reinforce the conviction of their reasoning. It may be the case that teachers have not grasped this distinction of argumentative keys for commenting, as dialectics prevails in legal situations closed by an excluding veridiction and the enquiry key in research situations open to the substantiated postulation of hypotheses as theories that enable problems to be solved by generating knowledge through investigation into unknowns. It is also presumed that in their educational teaching beliefs, there are still no clichés related to research that would allow them to perceive this discernment, nor is there a consolidated line of research on this subject so as to be able to establish them in their initial and ongoing training experience, as research methodology is recent knowledge in their professional training.

It is possible that in the implicit theories through which the teachers have given this answer about the “focal point of enquiry,” it is plausible to argue that in order for the persuasive intention to be effective in convincing the other person of one’s own position, it is necessary to activate strategies that encourage solutions to the problem posed through building arguments centred on the cause, through the pursuit of univocity or semantic precision in the definition of the concepts presented, through the use of authoritative quotations, analogy, etc. In this sense, 80% of the sample attributed an essential value to the process of enquiry that the commentator must carry out in order to propose solutions and argue their theory.

SO7. To Explore Teachers’ Beliefs Regarding Teaching Resources for Argumentative Text Commentary: Items 23 and 24

63.4% of teachers consider that textbooks do not provide teaching material in line with the teaching methodology suitable for



argumentation in text commentary. One of the major challenges, according to the results obtained, lies mainly in the development of a rich, varied and available repository for the development of argumentative competence and critical reading.

In relation to the use of ICT, the majority of respondents assign an important role to ICT regarding improving argumentative skills in text commentaries (Figures 8, 9).

The increasing use of hypertext—the electronic arrangement of networked structures made up of blocks of content interconnected by links, should be taken into account. In this connexion, the links take on the leading role, as they are responsible for allowing the user to move freely through the text, planning or choosing their own route. With a simple click, the reader enters a new textual space and becomes an active part of the argumentative creation process. These possibilities, in many cases, enable the receiver to not only make decisions about their reading project and the construction of their point of view, but also to modify the nature of what is written, manipulate content, contribute documents, transform the discourse, etc. From this perspective, hypertext leads to a multidimensional form of writing and, essentially, a more open and interactive type of hyperlinear reading.

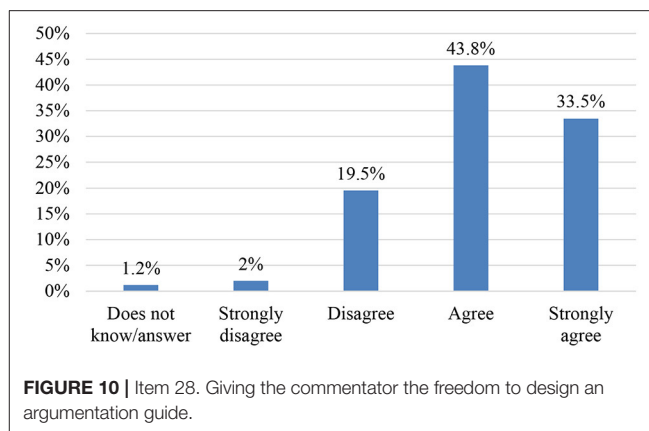
SO8. To Discern Teachers' Implicit Theories Regarding the Argumentative Commentary of Texts by Means of a Contrastive Analysis of the Answers to the Exploratory Questionnaire According to Typological Parameters: Items 25, 26, 27, and 28

With regard to effective procedures for the teaching of argumentative text commentary, it is noted that the participants tend to strongly agree (90.9%) with the fact that, when designing a teaching guide for effective text commentary, priority should be given to the cognitive processes involved in the contrast between the author's intention and the commentator's perspective in order to promote critical thinking. It is also the case that a high percentage of participants indicate (67.7%) that text analysis should focus on understanding the author's literal and implicit content.

The contrastive analysis of the responses given to questions 25 and 26 reveals incongruence in the options assessed, as it shows that two very different implicit teaching theories are positively regarded with very high and similar percentages: question no. 25 (percentage of responses: 50.60% agree, 17.13% strongly agree) focuses the text commentary guide on the comprehensive analysis of authorial intent and, therefore, corresponds to the type of productive implicit theory associated with a technical pedagogical model for teachers that promotes the consumption of knowledge without giving the commentator the opportunity for critical expression. On the other hand, question 26 (percentage of responses: 31.8% agree and 58.96% strongly agree) focuses this guide on the cognitive processes involved in the contrast between the author's intentional sense and the reading sense that the commentator believes it has in order to promote critical thinking, which corresponds to the types of implicit interpretive and emancipatory theories associated with the constructivist and critical pedagogical models, respectively. Therefore, as there are quite a number of teachers who equally value both issues of very different logic, their attributes reveal a deficit of epistemological discernment which may be due to confused or ambiguous beliefs about the practise of text commentary because they want to balance the functional customs of the profession (which tend to focus only on commenting on the author's arguments) with the innovative expectations of the current pedagogical renewal (which tend to focus on the dialogue between the author's and the reader's arguments).

Moreover, regarding the effective procedure for the discursive organisation of the argumentative commentary, 83.27% of teachers prefer the argumentation guide to contain guidelines so that students learn to develop it in their sections and writing appropriately (question 27), and 77.29% of teachers choose to give freedom to the logical, stylistic and contextualised expression of their personal critical sense (question 28). See **Figure 10**.

Thus, once again, an overlap in the responses emerges in the also positive assessment, with very high percentages for two very different issues: giving guidelines or giving freedom when commenting by giving arguments. It would be appropriate for



those who agree with question 27 to disagree with question 28 and vice versa. Nevertheless, the answers given, ranging from “agree” to “strongly agree,” show that teachers' implicit theories on these issues are ambiguous. The cause possibly lies in the fact that they usually practise their teaching methods without having reflected epistemologically on the matter in order to take defined positions, or perhaps it is due to the fact that the training and teaching materials they use have not facilitated such a reflective process, as there are certainly no scientific publications to date that have studied which of the two procedures indicated (with guidelines or free) is the more effective.

DISCUSSION AND CONCLUSION

The belief system related to argumentative text commentary influences not only the way teachers make decisions and approach their teaching practises in the classroom, but also students' academic performance (Trigwell and Prosser, 1991; Estévez-Nenninger et al., 2014). Our work contributes relevant information to this debate by understanding that the identification of this framework of thinking encourages awareness of actions aimed at transforming educational practise in order to improve the quality of teaching and learning.

In this sense, one of the most significant results of this research helps to classify and frame teacher profiles according to their implicit theories. These pedagogical models are distributed according to the role given during learning to the different education agents, the value and understanding of argumentative text commentary in the development of critical competence, the way of building knowledge and the implication of teaching methodologies.

The diversified results of the eight areas of the exploratory study conducted on teachers' implicit theories regarding the development of argumentation in the commentary of multimodal texts reveal important teacher training deficiencies on the epistemology that should underpin their teaching processes in this respect in order to firmly and clearly move toward the constructivist and critical pedagogical models recommended by current education laws. This finding is in line

with the results of previous research (de Vicente-Yagüe et al., 2019), which point to teachers' need for theoretical models on the didactics of argumentation that support and guide the design and implementation of educational proposals leading to the improvement of students' critical competence.

The exploration of teachers' epistemological convictions regarding the definition of text commentary (SO1) provides advanced assumptions corresponding to the interpretive theory of reading and its receptive writing in commentary. This theoretical explanation is consistent with a constructivist pedagogical model that has been consolidated as a manifest belief based on the dialogical conception of communicative competence in the construction of knowledge that teachers of different educational stages have been progressively assimilating in their teaching work since the LOE (Spanish organic law of education) came into force 15 years ago.

However, in the subsequent exploratory aspects (SO2–SO8), the profile of this initial assumption shown as an explicit theory with legal support and professionally accepted by the current institutional context comes into conflict with teachers' other deeper beliefs which reveal the tenacious persistence in their teaching work of the *hidden curriculum* of traditional and technical teaching, as Torres (2005) notes, based on implicit logocentric and functional theories:

- The exploration of teachers' preferences regarding the teaching modes of text commentary (SO2) shows that, although in their theoretical prolegomena they declare the action of commenting on texts from a dialogical and constructivist approach in accordance with the legal expectations of the twenty-first century curriculum, specific teaching decisions emerge in their teaching practise that contradict this ideal by preferring the written mode of commentary over the oral mode for reasons proper to implicit theories dependent on the traditional pedagogical model that promotes writing as a product, not as a process, which prevents the adequate management of competency-based education. Their explicit theories of innovative will again clash with their implicit theories when, while recognising the benefit of working on the interpretation of texts in communicative acts with group interaction, the attributes they concede to text commentary remain individualistic and silent. We agree with previous studies (Aubert et al., 2009; Giménez and Subtil, 2015) which show that teachers continue to identify text commentary as a solitary student activity consisting of unravelling the author's ultimate intention.
- The exploration of teachers' judgments on the generic effect of argumentation (SO3) shows a change in mindset that overcomes the traditional academic confinement of argumentation in the few genres where it is voluntary and explicit, since today's teacher acknowledges that argumentation can appear in any type of text, which implies a predisposition toward the interpretive study of its multimodal implicatures in the different socio-cultural contexts, processes and expressive formats that require the reader's cooperation to establish the meaning of the text.

Thanks to the recognition of the multimodal and implicit condition of argumentation—possibly motivated by the open vision of textuality that the intense use of ICTs in the Knowledge Society provides, text commentary can cease to be understood as an explanatory note on the text and become an expansion of its meaning through the hypertextual horizon established by the interpretive and expressive expectations of each reader. This conclusion is reinforced by the analysis of the results of SO7, where the majority of teachers highlight the insufficiency of the teaching materials available to teach argumentation in text commentary, and rely on the competent use of ICT to improve students' argumentative skills due to the real training opportunities relating to strategic, creative and critical maturity provided by multidimensional and open communication in hypertextual dynamics.

- The exploration of pragmatic teaching models on the verbal communication of argumentation (SO4) visualises, in the analysis of argumentative expression, the dialectical approach prevailing in such models with deep cultural and authoritative roots that continue to give prevalence to the legal rhetorical model of argumentation that translates dialogue as the discussion of opposing positions, whose democratic education emphasises respectful speaking times and consensual negotiation of outcomes. Indeed, this vision of the dialogic learning of argumentative expression as a dispute between antithetical positions stems from the strong roots of legal argumentation in academic discursive practises, maintained from Greek classicism to the present day (Plantin, 2005), without any loss, in a communicational facticity approach (van Eemeren and Grootendorst, 1984) which has been included in language teaching studies for decades (Camps, 1995; Cros, 2003).

We regret that, due to this belief where dialectics monopolises the notion of dialogue, the construction of knowledge does not progress beyond dissension or consensus, toward other heuristic and emancipatory spaces of coexistence and science, as these would be possible if it were understood that dialogic communication can take place by cooperating in focus groups with shared synergies where debate is a minor factor and not the centre of the argumentative activity essential to undertake, develop and finalise common projects with a plausible sense. Previous papers insist on the need to deconstruct teaching habits that are opposed to more democratic and dialogical channels of participation in classroom dynamics (Caro et al., 2018). We believe that it is necessary to generate educational research knowledge to tackle this problem, providing teachers with the necessary training to relativise dialectical argumentation in text commentary as one of its possible teaching approaches, as, in addition to commenting on controversial texts whether for or against, personal critical commentary could be used in the classroom to launch heuristic hypotheses, for example, to explain an unknown or to formulate a challenge with convincing arguments. We believe that, in this way, the deliberative argumentation of commentary would leave behind the only area of confrontations where only opinion serves to

advance toward the open field of research that opens up and substantiates knowledge.

Moreover, the dialectical approach is consistent with the professional belief in the subjectivity of argumentation for both individual and collective expression and this supports the widespread conviction of the critical practise of text commentary.

With regard to argumentative comprehension, the predominant professional belief regarding the essential value of reading in depth the implicatures of the texts commented on represents another indispensable step forward in their mature convictions to design relevant and well-contextualised argumentative practises, making it clear that quality argumentative discourse requires an inferential process based on enquiries that clarify the implicit meaning and sense elicited from textual explicitness.

- The exploration regarding the value that the commentator should place on the wording of the text (SO5), given that the majority of responses show a tendency to subjectively and critically question it, provides a rationale for pedagogical initiatives in line with the current principles of competency-based education, as such a consideration serves to empower the interpreting reader's horizon of expectations in the classroom when establishing a dialogue *between peers*—between two subjective perspectives—with the text and its author. This shows a major advance with respect to the traditional profile of the commentator whose mission is to extract only the authorial sense of the text, as fostering the interpreting reader's power of judgment is very beneficial for leading the way to the implicit emancipatory theory incited by the genuinely democratic practise of commentary in the classroom.
- The exploration conducted on the attributes regarding the argumentative key of commentary (SO6) focused on two different possibilities: the dialectic key, which prevails in legal situations closed by an excluding veridiction (van Eemeren and Houtlosser, 2002), and the enquiry key, which prevails in investigative situations open to the substantiated postulation of hypotheses enabling problem-solving by generating knowledge from unknowns. The results analysed are consistent with the results of SO4, as once again dialectics predominates as the key to justifying commentary as a discursive space for taking a stance on controversies with arguments defending a certain theory. Nevertheless, as the teachers also rated the key to the focal point of enquiry with a similar percentage, which unfolds when the commentary runs heuristically with synergies for the shared construction of knowledge, one of the following two possibilities can be interpreted: teachers accept both argumentative keys, or they lack epistemological discernment between the two, as there is still no consolidated line of research in the area of Spanish language teaching regarding the key to the focal point of enquiry applied to text commentary that may have been disseminated in initial and ongoing teacher training and,

furthermore, research methodology is recent knowledge in their professional training.

- Finally, the discernment of teachers' implicit theories on the argumentative commentary of texts regarding the comprehension and expression strategies that an *ad hoc* teaching guide should offer (SO8) shows their preference for dialogue between the author's theory and the commentator's critical hypothesis. However, this presumed critical emancipation of teachers is not clear due to the epistemological deficit revealed in the analysis of results, as they preferred in high and similar proportions two questions concerning disparate pedagogical models, which could be related to the dissociation that exists between their habits in the classroom (comments made adhering to guidelines and comments on text content and author's arguments) and the prologues of the laws governing the curriculum (dialogical and critical approach; freedom of expression), as well as the shortage of training opportunities and professional publications to reflect on such educational issues in focus groups.

The findings of this study raise additional questions that could open up new lines of work. The identification of teachers' beliefs involves the construction of a training proposal which, based on constructivist and critical pedagogical models, is the independent variable in quantitative experimental research with two treatment and comparison groups.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research has been derived from the R + D + I Project entitled Epistemic innovation of a model of argumentative comment of multimodal texts in the teaching of Spanish as a mother and foreign language (ref. PGC2018-101457-B-I00), funded by FEDER/Ministry of Science and Innovation - State Research Agency of the Government of Spain.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.749426/full#supplementary-material>

REFERENCES

- Almeda, R. (2014). Las creencias de profesores brasileños sobre el currículo de español como lengua extranjera en la enseñanza media. *Phonica* 9, 156–163. doi: 10.1344/phonica.2014.0.156-163
- Amossy, R. (2008). Argumentation et Analyse du discours: perspectives théoriques et découpages disciplinaires. *Argumentation Analyse Discours* 1, 1–18. doi: 10.4000/aad.200
- Aubert, A., García, C., and Racionero, S. (2009). El aprendizaje dialógico. *Cultura Educ.* 21, 129–139. doi: 10.1174/113564009788345826
- Ballesteros, C., Llobera, M., Cambra, M., Palou, J., Riera, M., Civera, I., et al. (2001). “El pensamiento del profesor. Enseñanza de lengua y Reforma,” in *El aula como espacio de investigación y reflexión*, coord. A. Camps (Barcelona: Graó), 195–207.
- Beltrán, D. (2019). La profesión docente y las teorías implícitas sobre la enseñanza. Una reflexión crítica. *Rev. Currículum* 32, 195–212. doi: 10.25145/j.qurricul.2019.32.11
- Bigi, E., García, M., and Chacón, E. (2019). ¿Qué textos académicos escriben los estudiantes universitarios de Educación? *Zona Próxima* 31, 26–55. doi: 10.14482/zp.31.378.2
- Bordieu, P. (1989). “Sistemas de enseñanza y sistemas de pensamiento,” in *La enseñanza: su teoría y su práctica*, eds J. Gimeno and A. Pérez (Madrid: Akal), 20–36.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Lang. Teach.* 36, 81–109. doi: 10.1017/S0261444803001903
- Bunge, M. (2009). *Tratado de Filosofía. Semántica II. Interpretación y verdad*. Barcelona: Gedisa.
- Cambra, M., Civera, I., Palou, J., Ballesteros, C., and Riera, M. (2000). Creencias y saberes de los profesores en torno a la enseñanza de la lengua oral. *Cultura Educ. Rev. Teoría Investig. Prác.* 17–18, 25–40. doi: 10.1174/113564000753837179
- Camps, A. (1995). Aprender a escribir textos argumentativos: características dialógicas de la argumentación escrita. *Comun. Lenguaje Educ.* 25, 51–63. doi: 10.1174/021470395321340439
- Capomagi, D. (2013). La escritura académica en el aula universitaria. *Rev. Educ. Desarrollo* 25, 29–40. Available online at: <https://www.imbiomed.com.mx/articulo.php?id=94387>
- Carlino, P. (2008). Concepciones y formas de enseñar la escritura académica: un estudio contrastivo. *Signo Seña* 16, 71–117. doi: 10.34096/sys.n16.5707
- Caro, M. T. (2015). El desarrollo dialógico de la argumentación informal en el comentario de texto. *Profesorado Rev. Currículum Formación Profesorado* 19, 436–451. Available online at: <https://recyt.fecyt.es/index.php/profesorado/article/view/41050>
- Caro, M. T. (2017). “La renovación sistémica en la sociedad del siglo XXI,” in *La universidad en cambio. Gobierno, renovación pedagógica, ética y condiciones laborales del profesorado*, coords. J. M. Escudero Muñoz and M. Vallejo Ruiz (Murcia: Editum), 13–135.
- Caro, M. T. (2018). La comunicación argumentativa en la Sociedad del Conocimiento, clave del liderazgo distribuido para un cambio educativo desde el desarrollo profesional. *Rev. Educ. Distancia* 56, 1–30. doi: 10.6018/red/56/8
- Caro, M. T., de Amo, J. M., and Domínguez-Oller, J. C. (2021). “Innovación epistémica de un modelo de comentario argumentativo de textos multimodales en la enseñanza del español como lengua materna y extranjera. Consistencia interna de una exploración,” in *Nuevos retos y perspectivas de la investigación en Literatura, Lingüística y Traducción*, coords. A. Flores and R. Pérez (Madrid: Dykinson), 1738–1755.
- Caro, M. T., de Vicente-Yagüe, M. I., and Valverde, M. T. (2018). Percepción docente sobre costumbres metodológicas de argumentación informal en el comentario de texto. *Rev. Española Pedagogía* 76, 273–293. doi: 10.22550/REP76-2-2018-04
- Caro, M. T., and González, M. (2012). *Comentario de texto fácil para Bachillerato y acceso a la Universidad*. Madrid: Espasa.
- Caro, M. T., and González, M. (2015). “El comentario de textos,” in *Didáctica de la Lengua y Educación Literaria*, coords. P. Guerrero and M. T. Caro (Madrid: Ediciones Pirámide), 349–364.
- Caro, M. T., and González, M. (2018). *Didáctica de la argumentación en el comentario de textos*. Madrid: Síntesis.
- Castañeda, L., Esteve, F., and Adell, J. (2018). ¿Por qué es necesario repensar la competencia docente para el mundo digital? *Rev. Educ. Distancia* 56:2. doi: 10.6018/red/56/6
- Castañeda, S., and Ortiz, B. D. (2017). “El rol de las creencias epistemológicas en el aprendizaje académico en psicología,” in *Obstáculos epistemológicos en la enseñanza y el aprendizaje de la filosofía y de la ciencia*, coords. Z. Monroy, R. León and G. Álvarez (México: UNAM, Facultad de Psicología), 263–283.
- Castell, M., and Mateos, M. (2015). Las representaciones de profesores y estudiantes sobre escritura académica en las universidades españolas. *Cultura Educ. Rev. Teoría Investig. Prác.* 27, 489–500. doi: 10.1080/11356405.2015.1072357
- Castorina, J. A., Barreiro, A., and Toscano, A. G. (2005). “Dos versiones del sentido común: las teorías implícitas y las representaciones sociales,” in *Construcción conceptual y representaciones sociales. El conocimiento de la sociedad*, comp J. A. Castorina (Buenos Aires: Miño y Dávila), 205–238.
- Chomsky, N. (1974). *Estructuras sintácticas*. Mexico City: Siglo XXI.
- Clark, C., and Peterson, P. (1990). “Procesos de pensamiento de los docentes,” in *La investigación de la enseñanza, III. Profesores y alumnos*, ed M. C. Wittrock (Barcelona: Paidós), 443–543.
- Cordero, G., and Carlino, P. (2019). “El análisis de la actividad docente: un medio para reflexionar sobre el uso epistémico de la escritura en una asignatura de ingeniería,” in *Conocer la escritura: investigación más allá de las fronteras / Knowing Writing: Writing Research*, eds Ch. Bazerman, B. González, D. Russell, P. Rogers, L. Peña, E. Narváez et al. (Bogotá, Colombia: Universidad Javeriana - The WAC Clearinghouse), 342–370.
- Cros, A. (2003). *Convencer en clase. Argumentación y discurso docente*. Barcelona: Ariel.
- de Vicente-Yagüe, M. I., Valverde, M. T., and González, M. (2019). Necesidades de formación del profesorado de Lengua y Literatura para el desarrollo de la argumentación informal en el comentario de texto. *Educ. Siglo XXI* 37, 213–234. doi: 10.6018/educatio.363471
- Díez, A. (2017). Más sobre la interpretación (II). Ideas y creencias. *Rev. Asociación Española Neuropsiquiatría* 37, 127–143. doi: 10.4321/S0211-57352017000100008
- Domínguez, J., López, S., and Andión, M. A. (2019). Las creencias de los profesores de español como lengua extranjera sobre los estilos de aprendizaje. *Rev. Española Lingüística Aplicada* 32, 419–454. doi: 10.1075/resla.17041.dom
- Dorfman, M. (2018). El educador en línea: más allá de la digitalidad. *Rev. Educ. Distancia* 56, 1–25. doi: 10.6018/red/56/9
- Englund, T. (1996). “Are professional teachers a good thing?,” in *Teachers’ Professional Lives*, eds I. F. Goodson and A. Hargreaves (London: Falmer Press), 75–87.
- Estévez-Nenninger, H., Valdés-Cuervo, A., Arreola-Olivarria, G., and ZavalaEscalante, G. (2014). Creencias sobre enseñanza y aprendizaje en docentes universitarios. *Magis Rev. Int. Investig. Educ.* 6, 49–64. doi: 10.11144/Javeriana.M6-13.csea
- Fandiño, C. G. (2007). *El pensamiento del profesor sobre la planificación en el trabajo por proyectos*. Grado de transición. Bogotá: Fondo Editorial Universidad Pedagógica Nacional.
- Flores, M. D. (2018). La escritura académica en estudios de ingeniería: valoraciones de estudiantes y profesores. *Rev. Educ. Superior* 47:186. doi: 10.36857/resu.2018.186.344
- Foucault, M. (1973). *El orden del discurso*. Barcelona: Tusquets.
- Giménez, G., and Subtil, C. (2015). Enseñar a argumentar en la escuela. Un estudio a través de los manuales escolares. *Cuaderno Educ.* 13, 1–13. Available online at: <https://revistas.unc.edu.ar/index.php/Cuadernos/article/view/11492>
- Giraldo, C. (2015). La escritura en el aula como instrumento de aprendizaje. Estudio en Universidades. *Ánfora* 22, 39–58. doi: 10.30854/anf.v22.n38.2015.25
- Goodman, K. (1994). “Reading, writing, and written texts: a transactional sociopsycholinguistic view,” in *Theoretical Models and Processes of Reading*, eds H. Singer and B. Rudell (Newark: IRA; Erlbaum), 1093–1130.
- Gordillo, A. (2019). Factores exógenos, teorías implícitas y procedimientos lingüístico-discursivos en la escritura científica. *Enunciación* 24, 184–198. doi: 10.14483/22486798.15199
- Grossman, P. (1990). *The Making of a Teacher. Teacher Knowledge and Teacher Education*. Chicago, IL: Teacher College Press.

- Hargreaves, A., and Goodson, I. F. (1996). "Teachers' professional lives: Aspirations and actualities," in *Teachers Professional Lives*, eds I. F. Goodson and A. Hargreaves (London: Farmer Press), 1–27.
- Hernández-Pina, F., and Maquilón-Sánchez, J. J. (2011). Las creencias y las concepciones. Perspectivas complementarias. *REIFOD* 14, 165–175. Available online at: <https://redined.educacion.gob.es/xmlui/handle/11162/81393>
- Inguanzo, G. (2010). *Creencias de los profesores de nivel de licenciatura sobre la naturaleza del conocimiento y los procesos de enseñanza y aprendizaje* (Tesis Doctoral). Universidad Iberoamericana de Puebla, Puebla, México.
- Jiménez-Llanos, A. B., and Feliciano-García, L. (2006). Pensar el pensamiento del profesorado. *Rev. Española Pedagogía* 64, 105–122. Available online at: <https://revistadepedagogia.org/lxiv/no-233/pensar-el-pensamiento-del-profesorado/101400002461/>
- Kelly, G. (1955). *The Psychology of Personal Constructs*. New York, NY: Norton.
- Laberge, D., and Samuels, S. J. (1994). "Toward a theory of automatic information processing of reading," in *Theoretical Models and Processes of Reading*, eds H. Singer and B. Rudell (Newark: IRA; Erlbaum), 689–718.
- Lluch, G., and Serrano, M. A. (2016). La terminología en el comentario de textos de Bachillerato y las PAU. *Textos. Didáctica Lengua Literatura* 73, 64–69.
- López, J. (2008). *El comentari de text: un instrument didàctic*. Per una pedagogia del discurs. València: Universitat de València.
- Makuc, M. (2008). Teorías implícitas de los profesores acerca de la comprensión de textos. *Rev. Signos* 41, 403–422. doi: 10.4067/S0718-09342008000300003
- Maldonado, A. L., González, E. J., and Cajigal, E. (2019). Representaciones sociales y creencias epistemológicas. Conceptos convergentes en la investigación social. *Cultura Representaciones Soc.* 13, 412–432. doi: 10.28965/2019-26-15
- Marrero, J. (1993). "Las Teorías Implícitas del profesorado: vínculo entre la cultura y la práctica de la enseñanza," in *Las teorías implícitas. Una aproximación al conocimiento cotidiano*, comp M. J. Rodrigo, A. Rodríguez, and J. Marrero (Madrid: Visor), 243–276.
- Martínez, I. (2018). La argumentación: creencias y opiniones de docentes de español como lengua extranjera (ELE). *Signos ELE Rev. Español como Lengua Extranjera* 12, 1–10. Available online at: <https://p3.usal.edu.ar/index.php/ele/article/view/4985>
- Martins, I. (2014). Transformaciones en el pensamiento de unos profesores universitarios de áreas no lingüísticas sobre la escritura académica. *Bellaterra J. Teach. Learn. Lang. Literature* 7, 21–33. doi: 10.5565/rev/jtl3.573
- Martins, I. (2012). El pensamiento del profesor sobre la escritura académica y profesional: un estudio de caso de la escuela de Derecho de la Universidad Central de Venezuela. *Legenda* 16, 71–93. Available online at: <http://erevistas.saber.ula.ve/index.php/legenda/article/view/4263>
- Mendoza, A. (2001). *El intertexto lector. El espacio de encuentro de las aportaciones del texto con las del lector*. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.
- Minervini, R. (2004). *Lingüística aplicada y enseñanza de lenguas próximas en facultades de traducción e interpretación: las teorías implícitas de los profesores* (Tesis Doctoral). Universidad Complutense de Madrid, Madrid, Spain.
- Mitchell, J. (1995). Teachers' implicit theories concerning questioning. *Br. Educ. Res. J.* 20, 69–84. doi: 10.1080/0141192940200107
- Mojarro-Delgadillo, A., and Alvarado-Nando, M. (2021). Práctica docente de la alfabetización inicial desde las concepciones pedagógicas y teorías implícitas del profesorado. *Rev. Educac. Desarrollo* 57, 79–88. Available online at: https://www.cucs.udg.mx/revistas/edu_desarrollo/antiores/57/57_Mojarro.pdf
- Montero, L., and Gewerc, A. (2018). La profesión docente en la sociedad del conocimiento. Una mirada a través de la revisión de investigaciones de los últimos 10 años. *RED. Rev. Educ. Distancia* 56:3. doi: 10.6018/red/56/3
- Munita, F. (2013). Creencias y saberes de futuros maestros (lectores y no lectores) en torno a la educación literaria. *Ocnos. Rev. De Estudios Sobre Lectura* 9, 69–88. doi: 10.18239/ocnos_2013.09.04
- OECD (2019). *TALIS 2018. Marco Conceptual*. Madrid: Ministerio de Educación y Formación Profesional. Available online at: <https://sede.educacion.gob.es/publiventa/talis-2018-marco-conceptual/ensenanza/23447>
- Ortega Gasset, J. (1976). *Ideas y creencias*. Madrid: Espasa-Calpe.
- Ortiz, M., López, T., García, M., Molina, G., Pernía, E., Busto, I., et al. (2009). Representaciones y prácticas de docentes de educación básica en el campo de la lectura y la escritura. *Letras* 51, 89–125. Available online at: http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0459-12832009000200004
- Pajares, M. F. (1992). Teacher's beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62, 307–332. doi: 10.3102/00346543062003307
- Pearson, D., and Stephens, D. (1994). "Learning about literacy: a 30-year journey," in *Theoretical Models and Processes of Reading*, eds H. Singer and B. Rudell (Newark: IRA; Erlbaum), 22–103.
- Pérez, A., and Gimeno, J. (1988). Pensamiento y acción en el profesor: de los estudios sobre la planificación al pensamiento práctico. *Infancia Aprendizaje* 42, 37–63. doi: 10.1080/02103702.1988.10822201
- Plantin, C. (2005). *L'argumentation. Histoire, Theories et Perspectives*. Paris: PUF.
- Pope, M., and Ken, T. (1981). *Personal Construct Psychology in Education*. London: Academic.
- Pratt, D. (1992). Conceptions of teaching. *Adult Educ. Q.* 42, 203–220. doi: 10.1177/074171369204200401
- Rodríguez-Martínez, F. (2016). El comentario de textos literarios en Educación Secundaria: evolución metodológica y papel de la literatura comparada. *Ocnos* 15, 119–135. doi: 10.18239/ocnos_2016.15.2.953
- Rosenblatt, L. (1996). "La teoría transaccional de la lectura y la escritura," in *Textos en contexto. Los procesos de lectura y escritura*, ed S. Kenneth Goodman (Buenos Aires: Proyecto Editorial Lectura y Vida), 13–70.
- Ros-Garrido, A., and Chisvert-Tarazona, M. J. (2018). Las investigaciones sobre las teorías implícitas del profesorado de formación profesional en el estado español. *Profesorado Rev. Currículum Formación Profesorado* 22, 97–115. Available online at: <https://recyt.fecyt.es/index.php/profesorado/article/view/63623>
- Rumelhart, D. (1997). "Hacia una comprensión de la comprensión," in *La lectura*, comp E. Rodríguez and E. Laguer (Santiago de Cali: Editorial Universidad del Valle), 25–51.
- Santos, I., and Alexopoulou, A. (2014). Creencias y actitudes de profesores y alumnos griegos de español ante las técnicas de corrección en la interacción oral: estudio comparativo intragrupos. *Didáctica. Lengua Literatura* 26, 429–446. doi: 10.5209/rev_DIDA.2014.v26.46845
- Scardamalia, M., and Bereiter, C. (1992). Dos modelos explicativos de los procesos de composición escrita. *Infancia Aprendizaje* 58, 43–64. doi: 10.1080/02103702.1992.10822332
- Schommer-Aikins, M. (2004). Explaining the epistemological belief system: introducing the embedded systemic model and coordinated research approach. *Educ. Psychol.* 39, 19–29. doi: 10.1207/s15326985ep3901_3
- Schön, D. (1998). *El profesional reflexivo. Cómo piensan los profesionales cuando actúan*. Barcelona: Paidós.
- Serrano, R. (2010). Pensamientos del profesor: un acercamiento a las creencias y concepciones sobre el proceso de enseñanza-aprendizaje en la Educación Superior. *Rev. Educ.* 352, 267–287. Available online at: http://www.revistaeducacion.educacion.es/re352/re352_12.pdf
- Suárez, N. M. (2015). *Teorías implícitas y prácticas de enseñanza de la lectura* (Tesis Doctoral). Universidad de la Laguna, Tenerife. Available online at: <https://riull.ull.es/xmlui/handle/915/7151>
- Torres, J. (2005). *El currículum oculto*. Madrid: Ediciones Morata.
- Trigwell, K., and Prosser, M. (1991). Improving the quality of student learning: the influence of learning context and student learning outcomes. *Higher Educ.* 22, 251–266. doi: 10.1007/BF00132290
- Usó, L. (2012). Creencias de los profesores de E/LE sobre la enseñanza/aprendizaje de la pronunciación. *Phonica* 16, 1–25. doi: 10.1344/phonica.2012.8.%25p
- van Dijk, T., and Kintsch, W. (1983). *Strategic of Discourse Comprehension*. New York, NY: Academic Press.
- van Eemeren, F., and Grootendorst, R. (1984). *Speech Acts in Argumentative Discussions. A Theoretical Model for the Analysis of Discussions Directed towards Solving Conflicts of Opinion*. Dordrechts: Foris.
- van Eemeren, F., and Houtlosser, P. (eds.). (2002). *Dialectic and Rhetoric. The Warp and Woof of Argumentation Analysis*. Amsterdam: Kluwer Academic Publishers.

- Vizcaíno, A., Cadalso, A., and Manzano, M. (2015). Adaptación de un cuestionario para evaluar las creencias epistemológicas sobre la matemática de profesores de secundaria básica. *Rev. Complutense Educ.* 26, 255–273. doi: 10.5209/rev_RCED.2015.v26.n2.42981
- Vizcaíno, A. E., López, K., and Klimenko, O. (2018). Creencias de autoeficacia y desempeño docente de profesores universitarios. *Rev. Katharsis* 25, 83–101. Available online at: <https://dialnet.unirioja.es/servlet/articulo?codigo=6259836>
- White, M. J., and Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educ. Psychol.* 30, 166–189. doi: 10.1016/j.cedpsych.2004.07.002
- Woods, D. (1996). *Teacher Cognition in Language Teaching*. Cambridge: Cambridge University Press.
- Zabalza, M. A. (1987). Pensamiento del profesorado y desarrollo didáctico. *Enseñanza Teach. Rev. Interuniversitaria Didáctica* 4, 109–138.
- Zhiying, L. (2018). *Creencias de profesores chinos respecto a la enseñanza del español como lengua extranjera en China* (Doctoral Thesis). Universitat Pompeu Fabra, Barcelona, Spain.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Caro Valverde, Amo Sánchez-Fortún and Domínguez-Oller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Importance of Phonological Awareness in Learning Disabilities' Prevention: Perspectives of Pre-School and Primary Teachers

Lurdes Veríssimo ^{*1,2}, Marisa Costa ^{1,2}, Francisca Miranda ¹, Catarina Pontes ¹ and Isabel Castro ¹

¹Faculty of Education and Psychology, Universidade Católica Portuguesa, Porto, Portugal, ²CEDH—Research Centre for Human Development, Porto, Portugal

OPEN ACCESS

Edited by:

María Isabel de Vicente-Yagüe
Jara, University of Murcia, Spain

Reviewed by:

Gustaf Bernhard Uno Skar,
Norwegian University of Science
and Technology, Norway
Juana Celia Domínguez-Oller,
University of Almería, Spain

*Correspondence:

Lurdes Veríssimo
lverissimo@ucp.pt

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

Received: 30 July 2021

Accepted: 28 October 2021

Published: 15 November 2021

Citation:

Veríssimo L, Costa M, Miranda F,
Pontes C and Castro I (2021) The
Importance of Phonological
Awareness in Learning Disabilities'
Prevention: Perspectives of Pre-
School and Primary Teachers.
Front. Educ. 6:750328.
doi: 10.3389/feduc.2021.750328

There is robust empirical evidence regarding predictors of success in reading and writing. However, pre-school and primary teachers are not always aware of this evidence and often do not know how to apply it in practice. Considering the importance of the role of these teachers and the importance of early school years in preventing learning difficulties, it is essential to study teachers' attitudes, knowledge, and strategies for improving educational processes toward successful development of literacy skills. Particularly, recognizing the impact of phonological awareness on the development of reading skills, this study has two main research questions: What is the importance pre-school and primary teachers attribute to the promotion of phonological awareness? and Which strategies for promoting phonological awareness do they apply in their professional practice? This qualitative study includes individual interviews with 10 pre-school and primary teachers, seeking to understand the perspective of each participant about phonological awareness, its importance, and implementation in practice. Data were collected in a context of great adversity (e.g., child maltreatment, early life stress). The results reveal that both pre-school and primary teachers recognized the relevance of phonological awareness among other pre-academic skills. Regarding the activities teachers implement, in their perspective, the playfulness seems to better result to stimulate this specific skill. Preschool teachers tend to focus on the dynamics that promote syllabic, intrasyllabic and phonemic awareness. Despite recognizing the gaps in phonological awareness in the first graders, primary teachers seem not to prioritize activities that stimulate and consolidate this skill, privileging the fluency and comprehension processes. The collaborative work between professionals was highlighted as very important to optimize the stimulation phonological awareness as well as facing other learning difficulties teachers find in this context. The results are discussed considering the role of pre-school teachers and primary teachers in reading difficulties' prevention in contexts facing multiple vulnerabilities.

Keywords: phonological awareness, learning disabilities, pre-school, primary school, prevention

INTRODUCTION

Reading is considered essential for academic success and school progression (e.g., Lacal et al., 2018; Porta and Ramirez, 2019). For this reason, there are several authors (e.g., Alonzo et al., 2020; Ferraz et al., 2020; Lacal et al., 2018) who, over time, have been dedicated to studying the variables involved in the process of learning to read. Phonological awareness has been pointed out as one of the most significant predictors of success in this domain (e.g., Melby-Lervåg et al., 2012). In fact, there are several studies that prove the importance of phonological awareness (e.g., Newbury et al., 2020), particularly in word recognition and reading fluency and accuracy (e.g., Ferraz et al., 2020), as well as reading speed and comprehension (e.g., Memisevic et al., 2020), stressing the need for early stimulation of this meta-skill (Wanzek et al., 2019). The stimulation of phonological awareness, by facilitating the process of learning to read, prevents the development of learning disabilities (e.g., Amorim et al., 2020). Such prevention is particularly important, since learning disabilities lead to consequences not only in academic domain (e.g., dropping out of school), but also in socioemotional (e.g., low self-esteem), and behavioral dimensions (e.g., maladaptive behaviors) (e.g., Cruz, et al., 2014; Eloranta et al., 2018; Emam, 2018; Razak et al., 2018; Rodrigues et al., 2018; Skues et al., 2016; Watson et al., 2016). Thus, considering the importance attributed to the stimulation of phonological awareness, it is important to emphasize the role of pre-school and primary school teachers. This role is even more relevant with students who experience difficulties in their learning process (Didion et al., 2020; Washburn et al., 2017). According to the *Curricular Guidelines for Pre-School Education* in Portugal, one of the content areas is oral language, and it is up to pre-school teachers to stimulate phonological awareness through rhymes, nursery rhymes, tongue twisters, and riddles (Direção Geral da Educação, 2016). However, a study by Rodrigues and others (2018) identified that, despite these guidelines and the fact that phonological awareness contributes to the reading process, teachers tend to use activities that promote vocabulary and reading comprehension and speed, to the detriment of phonological awareness. In fact, educators and teachers may need some support in understanding what the requirements for learning to read are, in identifying children who have difficulty mastering these requirements, and in implementing strategies that promote and facilitate the learning process of these students (Bratsch-Hines et al., 2017). Such a need is evidenced by Alshaboul, (2018) study, which sought, in general, to explore the beliefs and knowledge of 158 undergraduate students pre-service English language teachers regarding the importance of phonological awareness. This study found that participants had low literacy in phonological awareness but had more knowledge and were more aware of the importance of reading fluency and accuracy, at the expense of phonological awareness. In fact, the future teachers do not perceive themselves as being competent in the teaching of reading, showing low scores regarding the use of activities that relate directly to the development of skills necessary for reading, such as manipulating the sounds of each word and segmenting words

into phonemes. Thus, the participants in this study generally had low literacy regarding the use of pertinent methods and strategies that enable the teaching and assessment of phonological awareness. This type of result highlights the need attributed to the training of these professionals, which, in turn, influences what is each teacher's valuation of phonological awareness (Jordan et al., 2019). Thus, this exploratory study aims to explore the perceptions of valuation of pre-school and primary school teachers regarding phonological awareness in the prevention of learning disabilities.

MATERIALS AND METHODS

The following research questions were defined:

- 1) What importance do pre-school and primary school teachers attach to the promotion of phonological awareness?
- 2) What are the main strategies adopted by pre-school and primary school teachers to promote phonological awareness?

Participants

Five pre-school teachers and five primary school teachers participated in this study (N = 10), whose characterization is presented in **Table 1**.

The participants integrate one Portuguese school included in the Priority Intervention Educational Territories Program (TEIP), warded by the Ministry of Education. The schools that integrate the TEIP Program evidence high levels of school failure and early school dropout.

Instruments

Script for semi-structured interview “The importance of phonological awareness in the prevention of learning disabilities: Perspectives of pre-school teachers” (Pontes et al., 2020).

This script covers four dimensions: 1) socio-professional data; 2) valuation of the stimulation of phonological awareness; 3) strategies used to stimulate phonological awareness; and, 4) receptivity, importance, and perception of early childhood educators regarding phonological awareness stimulation programs.

Script for semi-structured interview “From learning disabilities in reading to intervention: primary school teachers' perspectives” (Miranda et al., 2020).

This script covers five dimensions: 1) socio-professional data; 2) teachers' perception of students' learning disabilities in reading in the early years of schooling; 3) teachers' sense of self-efficacy in addressing students' learning disabilities in reading in the early years of schooling; 4) strategies used at the level of learning disabilities in reading; and finally, 5) teacher training.

Data Collection Procedures

The study obtained approval in the Scientific Council of the Faculty considering all ethical issues related to research with human subjects. Subsequently, a request was made to the

TABLE 1 | Characterization of the participants.

Participant	Sex	Educator type	Grade of teaching	Years of teaching experience
T1	Female	Teacher	2nd grade	3 years
T2	Female	Teacher	1st grade	1 year
T3	Female	Teacher	1st grade	1 year
T4	Female	Teacher	3rd grade	3 years
T5	Male	Teacher	4th grade	1 year
P6	Female	Pre-school Teacher	3, 4 and 5 years	24 years
P7	Female	Pre-school Teacher	3, 4 and 5 years	27 years
P8	Female	Pre-school Teacher	3, 4 and 5 years	21 years
P9	Female	Pre-school Teacher	3, 4 and 5 years	1 month
P10	Female	Pre-school Teacher	3, 4 and 5 years	13 years

T = teacher; P = pre-school teacher.

Direction of the school context in which the participants carried out their professional activity, to obtain authorization to carry out the study. The research team contacted pre-school teachers and teachers, trying to identify potential participants. After authorization, each participant was contacted via email, with the objective of scheduling the interviews at a time that was most convenient. To ensure ethical and deontological issues, informed consents were obtained from the participants. The participants were previously informed about the objectives of the study, and data confidentiality was ensured, delimiting the use of the information collected only for research purposes. Considering the pandemic situation per COVID-19, only three of the ten interviews were conducted face-to-face, while the rest were conducted online through digital platforms (e.g., Zoom, Meet, Skype). Data collection took place on an individual basis, between June and July 2020 and, with due consent from the participants, the interviews were recorded in audio format.

Data Analysis Procedure

The data analysis went through different stages: pre-analysis, material exploration and data processing, and interpretation (Bazeley and Jackson, 2013). In a first moment, all the interviews were transcribed, and then we proceeded to a fluctuating reading of the interviews to obtain general information about the phenomenon under study. Afterwards, a process of creation of common tree node, and subsequent coding and content analysis was initiated using NVIVO software (version 12). This data analysis process followed a semi-inductive logic (Saldaña, 2013), in which the system of categories evolved according to the dimensions previously established in the interview guide, and other categories emerged from the data.

In addition to the definition of categories aligned with the interview guides, other methodological procedures were followed to guarantee the internal reliability of the coding and analysis processes (Gonçalves et al., 2021). Firstly, the 10 interviews were coded and to achieve consistency in the analysis, a second coding was performed 3 months later and compared to the first coding version. Also, the researcher responsible for the coding of systematically comparing the transcripts with the recordings as well as to verify the content analysis and conclusions drawn from the transcripts.

RESULTS AND DISCUSSION

What importance do pre-school and primary school teachers attach to the promotion of phonological awareness?

Among the various pre-academic skills, phonological awareness is the one that educators highlight as crucial for learning to read and write, considering it a factor that “it will contribute in some way to the success of those children. And when I say success, I mean success at school, success as a person, because if they succeed at school, they will be happier” (P6). The valorization of the stimulation of phonological awareness is also associated with “professional experience and to listen to other professionals” (P8) and to the “Pre-School Curricular Guidelines” received by the authority, which “end up pointing a little in this direction as well” (P7). Teachers corroborate this perception, stating that the most frequent manifestations of students’ reading difficulties are related to deficits in phonological awareness, highlighting that “[the students] cannot associate the letter with the sound, the phonetic part, the sound is where they have more difficulty” (T3). In addition, teachers report that as a result, students tend to have spelled reading and letter inversion, which often translates into reading comprehension deficits, and reading fluency deficits: “. . . and of course the level of interpretation, of what he read, also ended up being a little compromised” (T5).

In fact, empirical evidence shows that phonological awareness is one of the most influential variables in the acquisition of reading skills (e.g., Alonzo et al., 2020; Gutiérrez, 2017; Lacal et al., 2018), and that it is therefore crucial to stimulate during the pre-school age (Jiménez et al., 2009; Anthony & Francis, 2005). The difficulties identified by teachers in the process of their students learning to read are in line with the literature (e.g., Alonzo et al., 2020; Lacal et al., 2018), in that they emphasize phonological awareness for the success and development of reading. According to Hayward et al. (2017), students tend to exhibit several errors in phonological awareness, such as adding letters at the beginning, middle, and/or end of the word, omitting phonemes, repeating phonemes, reversing the order of phonemes, as well as substituting a letter or phonemes in a word. Considering that there is an interdependent relationship between decoding ability and comprehension, it would be expected that teachers would report these difficulties at the reading comprehension level, since children who have

difficulty reading words also have difficulty understanding written texts (Ribeiro et al., 2016).

According to teachers' perceptions, the factors that justify deficits in phonological awareness may be varied, namely associated with family characteristics. Teachers recognize that the absence of parental support influences the reading process, leading to increased challenges and difficulties: "(...) *then there are also other issues at the level of family support that is also important*" (T5). From the point of view of the family context, the teachers' perceptions converge on the importance of parental support in the process of acquiring the various skills, and according to the literature, the absence of stimulation opportunities in the early years, associated with poor parental involvement, may contribute to the development of reading difficulties (Battin-Pearson et al., 2000; Chiu and McBride-Chang, 2006; Corso and Meggiato, 2019; Moura et al., 2019; Nicolau and Navas, 2015; Rech and Miranda, 2018; Rosa, 2019). These findings are very relevant considering the importance of systemic approach to learning problems. Teachers, based on their professional experience, verify the importance of parental involvement considering that it plays an important role in the development of children's emergent literacy (Esmaeeli et al., 2019).

Regarding the various weaknesses that specific learning disabilities in reading can bring about, teachers highlight low academic performance, low school motivation, low self-esteem and low self-efficacy: "*unfortunately we have students who have a lot of disabilities, and this makes learning in other areas difficult as well. . . If they don't master reading, they will have difficulties in the other areas.*" (T4); "*The fact that they learn to read is a pride for them and their self-esteem goes up, their well-being, everything.*" (T4). The comprehensibility of the factors underlying the disabilities associated with learning to read and the consequent perceived impact on socioemotional and academic levels are in line with the literature. In fact, as previously mentioned, reading is crucial for students to experience academic success and for the learning of other school contents to occur (Khalid et al., 2017). Therefore, it is also consistent with evidences that reading disabilities directly affect the socioemotional level. Thus, recurrent and ongoing learning disabilities in reading can contribute to a loss of interest in school, a sense of frustration, and can consequently lead to behavioral problems, social isolation, depressive and anxiety symptoms (e.g., Cruz, et al., 2014; Emam, 2018; Razak et al., 2018; Rodrigues et al., 2018; Skues et al., 2016; Watson et al., 2016). This is why it is essential to prevent learning difficulties through an early, universal and systematic perspective. Pre-school primary school teachers play a determinant role in this process.

Main Strategies Adopted by Pre-School and Primary School Teachers to Promote Phonological Awareness

According to their daily practices, the pre-school teachers highlight diverse phonological awareness stimulation activities that aim to stimulate different types of phonological awareness, namely syllabic, intrasyllabic, and phonemic awareness.

According to the practices shared by the pre-school teachers, it was found that there were more strategies for stimulating syllabic awareness and intrasyllabic awareness, and subsequently, strategies for stimulating phonemic awareness were the least mentioned. In fact, studies indicate that the stimulation of phonological awareness should occur gradually, that is, it should start from larger phonological units to the level of smaller phonological units (Freitas et al., 2007; Rios, 2013).

Furthermore, teachers refer that they use story reading, rhymes, exploration of the environment, or specific materials to stimulate phonological awareness:

"They like stories a lot, we also tell a lot of stories, and then we try to explore at the level of sounds" (P7); "*I'm doing a world knowledge thing and that word I have there I can. . . ah. . . for example, we do endless syllabic division sessions*" (P10); "*On the rhyming blanket, with their names, rhymes with their mothers' names, rhymes with objects in the room*" (P8); "*We are going to make the sound box/f/and we are always looking for words with/f/, like/f/, father and go the image of the father or go the real object if it is possible*" (P6).

Besides the daily dynamics, some participants refer that they implement specific phonological awareness stimulation programs: "*In the "A Ler Vamos" and "Matiga" projects it is assumed that it is the kindergarten teacher who develops, but alongside and at the same time*" (P6). In this dynamic process of planning regular and intentional activities to stimulate phonological awareness, the participants highlight the increasing role of technological resources as a facilitator to engage children through games or other playful activities "*meeting interests and tastes*" (P7). One of the teachers even mention that this playful character is important in that "*the goals are achieved; they continue to play without realizing that they are already achieving what we want*" (P7).

In line and continuity with the methodologies and strategies used by the pre-school teachers, the teachers also tend to favor play, such as games, images, songs, and videos, in order not only to promote the acquisition and consolidation of the reading processes, but above all, to motivate students to learn. Thus, focusing on reading fluency, teachers tend to use activities that involve reading single words, reading simple sentences, and reading texts. Associated with the reading of texts to stimulate reading fluency, there are moments of exploration of the texts, in order to promote the students' reading comprehension. Despite the recognition of the cross-sectional gaps in phonological awareness among primary school students, teachers do not seem to prioritize in their planning activities to stimulate and consolidate this pre-academic competence, focusing more on letter identification and differentiation between vowels and consonants.

The stimulation of phonological awareness through games and play constitutes a motivator for the learning process (Cotonhoto et al., 2019). As mentioned above, students with specific learning disabilities in reading tend to be commonly unmotivated to learn (Rech and Skues et al., 2016; Miranda, 2018; Rodrigues et al., 2018). Thus, it is essential to enhance their motivation through playful activities (Rodrigues et al., 2018). These assume a high importance not only in the learning process,

but also in the development from the social and emotional point of view, which is why they have been gaining prominence in the school context to facilitate the development of skills and learning. The use of playful materials, such as games and songs, resources used by the participants, can contribute significantly to the process of stimulating phonological awareness in students, reducing reading difficulties (e.g., Cotonhoto et al., 2019; Júnior and Alencar, 2020).

The narratives of the pre-school and primary school teachers also reflect the need to develop collaborative work in promoting phonological awareness and preventing learning disabilities, with the support of other professionals:

“This ends up being a job, it may not seem like it because I’m alone in the classroom or I’m with the support colleague, but mostly I’m alone with them, but it’s still a team effort.” (T1); “I always prefer to work as a team.” (T2).

Research indicates that the years of experience are significantly associated with the value placed on phonological awareness by educational professionals (Jordan et al., 2019). In addition, their professional development is also associated with the experiences of articulation with other professionals (Cunningham and O’Donnell, 2015). The valorization of the multidisciplinary work identified by the participants may thus substantiate the value that these professionals place on phonological awareness and the practices they adopt, in order to overcome the disabilities identified in the students (Pisheh et al., 2017). In particular, an articulated and sequential work between pre-school teachers and primary teachers seems to be essential, ensuring a positive transition between pre-school and primary school.

CONCLUSION

The present study is a pilot exploratory study integrated into a larger project aimed at promoting school success in a school cluster with high rates of school failure and early school dropout (*Learning with All Project*). Considering that school failure is strongly correlated with learning disabilities in reading in the early years of schooling, and, in turn, the prevention of these specific learning disabilities are largely dependent on the development of phonological awareness, and, finally, the development of this pre-academic skill is dependent on the approach of pre-school teachers and primary teachers, it becomes essential, first of all, to understand what value these professionals attribute to this competence. Thus, the main results of this study show that pre-school teachers and primary school teachers, in general, know and value phonological awareness and identify its relevance to explain the learning difficulties they find in their children. The results also reveal that pre-school teachers tend to implement activities intentionally focused on stimulating this skill. To this end, participants prioritize playful, everyday activities, value technological resources, and recognize the importance of collaborative work in this area. However, it is noteworthy that primary school teachers focus more on reading tasks, putting in second place the basic skills that allow their successful completion, namely phonological awareness. It is

also noteworthy that, according to their perspective, the work of pre-school and primary school teachers seems not to be sufficient to overcome the difficulties experienced, considering the impact of the absence of parental support on learning tasks.

A strength of this study is the in-depth exploration of each participant’s individual perception, allowing, through their narrative but using a standardized interview protocol, an understanding of the value placed on phonological awareness, the risk factors that may explain, and the pedagogical practices that reflect it, in a school context with high rates of school failure and early school dropout. Moreover, the fact that pre-school and primary school teachers were considered in the same study allows a more integrated reading of the results, as well as the specificity of the contexts, which are intended to be articulated.

In future studies, it would be essential to expand the sample, by adding more clusters of pre-school and primary teachers to deepen comprehend the collaborative work developed to face the gaps on phonological awareness. It is also important that future research include observational data collection in order to clarify and confirm teachers’ practices. Thus, simultaneously, it would also be interesting to complement this type of exploratory study with quantitative methodologies that would allow, in a larger sample, to characterize the perceptions and practices associated with the promotion of phonological awareness.

The results shed light to several practical implications. According to an effective multi-tiered approach (Brown-Chidsey and Bickford, 2016), with a view to preventing learning disabilities, it is essential, first and foremost, that pre-school education is valued by the entire political, social, and educational system. Consistent with this valorization, training, capacity building and consulting with early school teachers should be aligned with the current state of research, including the deepening of knowledge and practices at the level of predictors of (in)succes in reading and writing. In this way, pre-school teachers will place greater value on phonological awareness, and will be more likely to intentionalize their pedagogical practices in order to stimulate these dimensions that constitute protective factors against school failure. Consequently, universal screening practices should be adopted, in which phonological awareness should be assessed in all children, seeking to identify weaknesses at this level, and providing the opportunity for intervention and stimulation before transition to primary school. When students move to primary school, collaborative work between professionals is essential to optimize the stimulation of phonological awareness and to understand the needs and academic weaknesses or vulnerabilities of each student, as well as the responsiveness of intervention to effectively prevent learning disabilities. This issue is particularly relevant when the educational contexts are of great adversity (e.g., child maltreatment, early life stress), in which several risk factors converge, and in which families are rarely factors of promotion and stimulation of skills. In this sense, the articulation between family and school, and the parental empowerment and availability of resources for consolidation of skills and learning seems to be crucial.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Scientific Council of the Faculty of Education and Psychology - Catholic University of Portugal. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Alonzo, C. N., McIlraith, A. L., Catts, H. W., and Hogan, T. P. (2020). Predicting Dyslexia in Children with Developmental Language Disorder. *J. Speech Lang. Hear. Res.* 63 (1), 151–162. doi:10.1044/2019_JSLHR-L-18-0265
- Alshaboul, Y. (2018). Jordanian Pre-service EFL Teachers' Perspectives about Phonological Awareness: Contributions to reading Development. *Aje* 5 (2), 173–188. doi:10.30958/aje.5-2-5
- Alves De Moura, A., Martins Rodrigues, F. N. C., Lopes Soares, S., and Simões Ferreira, H. (2019). A psicopedagogia na alfabetização de crianças com dificuldades de aprendizagem. *Rpge* 23 (1), 85–102. doi:10.22633/rpge.v23i1.11493
- Amorim, A. N., Jeon, L., Abel, Y., Felisberto, E. F., Barbosa, L. N. F., and Dias, N. M. (2020). Using Escribo Play Video Games to Improve Phonological Awareness, Early Reading, and Writing in Preschool. *Educ. Res.* 49 (3), 188–197. doi:10.3102/0013189x20909824
- Anthony, J. L., and Francis, D. J. (2005). Development of Phonological Awareness. *Curr. Dir. Psychol. Sci.* 14 (5), 255–259. doi:10.1111/j.0963-7214.2005.00376.x
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., and Hawkins, J. D. (2000). Predictors of Early High School Dropout: A Test of Five Theories. *J. Educ. Psychol.* 92 (3), 568–582. doi:10.1037/0022-0663.92.3.568
- Bazeley, P., and Jackson, K. (2013). *Qualitative Data Analysis with NVIVO*. nd Ed. Los Angeles: SAGE Publications.
- Bratsch-Hines, M. E., Vernon-Feagans, L., Varghese, C., and Garwood, J. (2017). Child Skills and Teacher Qualifications: Associations with Elementary Classroom Teachers' Reading Instruction for Struggling Readers. *Learn. Disabilities Res. Pract.* 32, 270–283. doi:10.1111/ldrp.12136
- Brown-Chidsey, R., and Bickford, R. (2016). *Practical Handbook of Multi-Tiered Systems of Support: Building Academic and Behavioral Success in Schools*. New York: The Guilford Press.
- Buchelt Rech, M. E., and Dos Anjos Miranda, M. (2018). Dislexia: A contribuição da Psicopedagogia no que se refere aos distúrbios de aprendizagem que afeta o desenvolvimento da leitura e da escrita. *Riai* 4 (3), 131–143. doi:10.17561/riai.v4.n3.8
- Chiu, M. M., and McBride-Chang, C. (2006). Gender, Context, and Reading: A Comparison of Students in 43 Countries. *Scientific Stud. Reading* 10 (4), 331–362. doi:10.1207/s1532799xssr1004_1
- Corso, L., and Meggiato, A. (2019). Quem são os alunos encaminhados para acompanhamento de dificuldades de aprendizagem? *Psicopedagogia* 36 (109), 57–72. Available at: <https://cdn.publisher.gn1.link/revistapsicopedagogia.com.br/pdf/v36n109a07.pdf>.
- Cotonhoto, L., Rossetti, C., and Missawa, D. (2019). A importância Do jogo e da brincadeira na prática pedagógica. *Revista Construção Psicológica* 27 (28), 37–47. Available at: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1415-69542019000100005&lng=pt&tlng=pt.
- Cruz, J. S., Almeida, M., Pinto, P., Constante, P., Macedo, A., Amaral, J., et al. (2014). Contribuição da literacia emergente para o desempenho em leitura no final Do 1.º CEB. *Análise Psicológica* 32 (3), 245–257. doi:10.14417/ap.749
- Cunningham, A., and O'Donnell, C. (2015). "29 Teachers' Knowledge about Beginning Reading Development and Instruction," in *The Oxford*

AUTHOR CONTRIBUTIONS

LV and MC contributed to conception and design of the study. FM, CP and IC organized the database. MC performed the qualitative analysis. LV, MC and FM wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This publication was funded by Norte 2020 and Fundo Social Europeu (NORTE-07-4740-FSE-000637).

- Handbook of reading*. Editors A. Pollatsek and R. Treiman. 1st ed. (Newgen: Publisher), 447–462. Available at: https://www.researchgate.net/profile/Anne-Cunningham-9/publication/275581846_Teachers'_Knowledge_about_Beginning_Reading_Development_and_Instruction/links/553fac870cf2320416eb6753/Teachers-Knowledge-about-Beginning-Reading-Development-and-Instruction.pdf.
- Didion, L., Toste, J. R., and Filderman, M. J. (2020). Teacher Professional Development and Student Reading Achievement: A Meta-Analytic Review of the Effects. *J. Res. Educ. Effectiveness* 13 (1), 29–66. doi:10.1080/19345747.2019.1670884
- Direção-Geral da Educação (2016). Orientações Curriculares para a Educação Pré-Escolar. Ministério da Educação e Ciência. Available at: https://www.dge.mec.pt/ocepe/sites/default/files/Orientacoes_Curriculares.pdf.
- Eloranta, A. K., Närhi, V. M., Eklund, K. M., Ahonen, T. P. S., and Aro, T. I. (2018). Resolving reading Disability-Childhood Predictors and Adult-Age Outcomes. *Dyslexia* 25 (1), 20–37. doi:10.1002/dys.1605
- Emam, M. M. (2018). Academic and Social Attributions as Predictors of Emotional and Behavioural Difficulties in Students Referred for Learning Disabilities and Typically Achieving Students: the Female Profile in Oman. *Emotional Behav. Difficulties* 23 (3), 326–339. doi:10.1080/13632752.2018.1455629
- Esmaceli, Z., Kyle, F., and Lundetrae, K. (2019). Contribution of Family Risk, Emergent Literacy and Environmental Protective Factors in Children's reading Difficulties at the End of Second-Grade. *Reading and Writing* 32, 2375–2399. doi:10.1007/s11145-019-09948-5
- Ferraz, I., Viana, F., and Pocinho, M. (2020). A "maturidade" para aprender a ler: contributos para uma reflexão. *Calidoscópio* 18 (1), 3–19. doi:10.4013/cld2.020.181.01
- Freitas, M., Alves, D., and Costa, T. (2007). *O conhecimento da língua: desenvolver a consciência fonológica*. Lisbon: Ministério da Educação.
- Ghanaat Pisheh, E. A., Sadeghpour, N., Nejatyjahromy, Y., and Mir Nasab, M. M. (2017). The Effect of Cooperative Teaching on the Development of Reading Skills Among Students with Reading Disorders. *Support Learn.* 32 (3), 245–266. doi:10.1111/1467-9604.12168
- Gonçalves, S., Gonçalves, J., and Marques, C. (2021). *Manual de investigação qualitativa*. Lisbon: Pactor.
- Gutiérrez, R. (2017). Facilitators of the Learning Process of Writing in Early Ages. *Anales De Psicología* 33 (1), 32–39. doi:10.6018/analesps.33.1.229611
- Hayward, D. V., Annable, C. D., Fung, J. E., Williamson, R. D., Lovell-Johnston, M. A., and Phillips, L. M. (2017). Beyond the Total Score: A Preliminary Investigation into the Types of Phonological Awareness Errors Made by First Graders. *Early Child. Educ J* 45 (6), 809–820. doi:10.1007/s10643-016-0829-3
- Jiménez, J. E., Naranjo, F., O'Shanahan, I., Muñetón, M., and Rojas, E. (2009). Pueden Tener Dificultades Los Niños Que Leen Bien? *Revista española de pedagogía* 67 (242), 45–60. doi:10.6018/analesps.33.1.229611
- Jordan, R. L. P., Garwood, J., and Trathen, W. (2019). Assessing General Education and Special Education Majors' Self-Efficacy for Teaching Reading efficacy for Teaching Reading. *Learn. Disabilities Res. Pract.* 34 (4), 185–193. doi:10.1111/ldrp.12207
- Júnior, S., and Alencar, F. (2020). O ensino da gramática na Educação Básica: o lúdico em sala de aula. *Caderno de ensino, linguagens e suas tecnologias* 1 (1), 95–105. Available at: <https://revistascientificas.ifrr.edu.br/revista/index.php/celte/article/view/1471/0>.

- Khalid, N. M., Buari, N. H., and Chen, A.-H. (2017). Comparison of Oral Reading Errors between Contextual Sentences and Random Words Among Schoolchildren. *Ies* 10 (1), 47–55. doi:10.5539/ies.v10n1p47
- Lacal, M., Madrid, D., and Estrada-Vidal, L. (2018). Factores predominantes en el aprendizaje de la iniciación a la lectura. *Revista Mexicana de Investigación Educativa* 23 (79), 1121–1147. Available at: http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S1405-66662018000401121&lng=es&nrm=iso.
- Melby-Lervåg, M., Lyster, S. A., and Hulme, C. (2012). Phonological Skills and Their Role in Learning to Read: A Meta-Analytic Review. *Psychol. Bull.* 138 (2), 322–352. doi:10.1037/a0026744
- Memisevic, H., Dedic, A., Biscevic, I., Hadzic, S., Pasalic, A., and Malec, D. (2020). Identifying Predictors of reading Speed and reading Comprehension in Bosnian. *Appl. Neuropsychol. Child.*, 1–10. doi:10.1080/21622965.2020.1815023
- Mr Watson, S., Gable, R., and Morin, L. (2016). The Role of Executive Functions in Classroom Instruction of Students with Learning Disabilities. *Int. J. Sch. Cog Psychol.* 03 (1), 167. doi:10.4172/2469-9837.1000167
- Newbury, J., Justice, L. M., Jiang, H. H., and Schmitt, M. B. (2020). Cognitive, Noncognitive, and Home Environment Correlates of Reading Difficulties in Primary-Grade Students with Language Impairment. *J. Speech Lang. Hear. Res.* 63 (6), 1933–1946. doi:10.1044/2020_jslhr-19-00363
- Nicolau, C. C., and Navas, A. L. G. P. (2015). Avaliação das habilidades preditoras Do sucesso de leitura em crianças de 1º e 2º anos Do ensino fundamental. *Rev. CEFAC* 17 (3), 917–926. doi:10.1590/1982-021620157214
- Porta, M. E., and Ramirez, G. (2019). The Impact of an Early Intervention on Vocabulary, Phonological Awareness, and Letter-Sound Knowledge Among Spanish-speaking Kindergarteners. *Int. J. Sch. Educ. Psychol.*, 8, 65–79. doi:10.1080/21683603.2018.1558137
- Razak, N. H. A., Johari, K. S. K., Mahmud, M. I., Zubir, N. M., and Johan, S. (2018). General Review on Cognitive Behavior Play Therapy on Childrens' Psychology Development. *Ijarped* 7 (4), 134–147. doi:10.6007/IJARPED/v7-i4/4842
- Ribeiro, I., Viana, F., Baptista, A., Choupina, C., Santos, S., Brandão, S., et al. (2016). Ainda Estou a Aprender. As tecnologias no apoio à avaliação e à intervenção nas dificuldades na aprendizagem da leitura. Maia: Almedina.
- Rios, C. (2013). *Programa de promoção do desenvolvimento da consciência fonológica Psico & Soma*. 2a ed.
- Rodrigues, A., Ribeiro, C., and Couceiro Figueira, A. P. (2018). Perturbação da Aprendizagem Específica com défice na leitura em alunos dos 1º e 2º ciclos Do ensino básico. Professores de educação especial e pais em colaboração. *Rieoei* 78 (2), 97–12. doi:10.35362/rie7823182
- Rosa, M. (2019). Dificuldades de aprendizagem na alfabetização. *Revista Praxis Pedagógica* 2 (1), 135–148.
- Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers*. Los Angeles: SAGE publications.
- Skues, J. L., Cunningham, E. G., and Theiler, S. S. (2016). Examining Prediction Models of Giving up within a Resource-Based Framework of Coping in Primary School Students with and without Learning Disabilities. *Int. J. Disabil. Develop. Educ.* 63 (2), 224–245. doi:10.1080/1034912X.2015.1070951
- Wanzek, J., Otaiba, S., and McMaster, K. (2019). *Intensive Reading Interventions for the Elementary Grades*. New York The Guilford Press.
- Washburn, E., Mulcahy, C., Musante, G., and Joshi, R. (2017). Novice Teachers' Knowledge of Reading-related Disabilities and Dyslexia. *Learn. Disabilities: A Contemp. J.* 15 (2), 169–191.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Veríssimo, Costa, Miranda, Pontes and Castro. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vocabulary Repetition Following Multisensory Instruction Is Ineffective on L2 Sentence Comprehension: Evidence From the N400

Reza Pishghadam¹, Haniyeh Jajarmi², Shaghayegh Shayesteh^{1*}, Azin Khodaverdi¹ and Hossein Nassaji³

¹ Department of English, Ferdowsi University of Mashhad, Mashhad, Iran, ² Department of English, Bahar Institute of Higher Education, Mashhad, Iran, ³ Department of English, University of Victoria, Victoria, BC, Canada

OPEN ACCESS

Edited by:

Olivia López Martínez,
University of Murcia, Spain

Reviewed by:

Andrés Antonio González-Garrido,
University of Guadalajara, Mexico
Sabela Fondevila,
Complutense University of Madrid,
Spain

*Correspondence:

Shaghayegh Shayesteh
shayesteh@um.ac.ir

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 09 May 2021

Accepted: 03 January 2022

Published: 28 January 2022

Citation:

Pishghadam R, Jajarmi H,
Shayesteh S, Khodaverdi A and
Nassaji H (2022) Vocabulary
Repetition Following Multisensory
Instruction Is Ineffective on L2
Sentence Comprehension: Evidence
From the N400.
Front. Psychol. 13:707234.
doi: 10.3389/fpsyg.2022.707234

Putting the principles of multisensory teaching into practice, this study investigated the effect of audio-visual vocabulary repetition on L2 sentence comprehension. Forty participants were randomly assigned to experimental and control groups. A sensory-based model of instruction (i.e., emotioncy) was used to teach a list of unfamiliar vocabularies to the two groups. Following the instruction, the experimental group repeated the instructed words twice, while the control group received no vocabulary repetition. Afterward, their electrophysiological neural activities were recorded through electroencephalography while doing a sentence acceptability judgment task with 216 sentences under acceptable (correct) and unacceptable (pragmatically violated) conditions. A one-way analysis of variance (ANOVA), a multivariate analysis of variance (MANOVA), and a Bayesian repeated-measures ANOVA were used to compare the behavioral and neurocognitive responses [N400 as the main language-related event-related brain potential (ERP) effect] of the two groups. The results showed no significant N400 amplitude difference in favor of any of the groups. The findings corroborated the ineffectiveness of two repetitions preceded by multisensory instruction on L2 sentence comprehension.

Keywords: multisensory instruction, emotioncy, event-related brain potentials (ERPs), N400, repetition

INTRODUCTION

Comprehension in general and sentence comprehension, in particular, have been the pinnacle of many cognitive studies on L2 (e.g., Newman et al., 2012; Zheng and Lemhöfer, 2019). In such studies, sentence processing has been examined from multiple perspectives, including syntactic (Embick et al., 2000) and semantic (Dapretto and Bookheimer, 1999; Bookheimer, 2002). The semantic processing of a sentence relies, to a large extent, on the processing of individual words of that sentence. Therefore, vocabulary retention plays a significant role, hastening or hindering this process.

To improve vocabulary learning and retention, different strategies have been employed. As for one, central to vocabulary learning as a gradual process is the concept of repetition (Nation, 2015).

Repetition is known to be mechanical or meaningful depending on the teaching methodology applied by teachers. Along with the changes in language teaching methodology from audio-lingual classroom drills grounded in the theory of behaviorism to communicative approaches, vocabulary repetition was constantly shaped and reshaped. Meaningful repetition practices took priority over simple mechanical ones to make learning more enduring (Horst, 2013; Kartchava and Nassaji, 2019; Hidalgo and Garcia Mayo, 2021).

Although researchers jointly agree that learning depends on the degree of any type of repetition with more repetitions leading to the better learning of the points (e.g., Thalheimer, 2003; Chen and Truscott, 2010; Laufer and Rozovski-Roitblat, 2011), there has been considerable debate over the optimal number of repetitions that ensures vocabulary learning and boosts comprehension (Peters, 2014; Nation, 2015; Liu, 2018). While Horst et al. (1998); Waring and Takaki (2003), and Webb (2007) respectively found 8, 10, and 12 repetitions as the optimal number, Vidal (2011) minimized the frequency to two and three meaningful repetitions in a reading context.

To capture the repetition effect, different approaches have been adopted. Unlike conventional approaches, which basically target learners' performance and achievement, most recent studies have endeavored to employ neurocognitive tasks to get more reliable results. By virtue of this objective inspection, a few studies documented that repetition may have no effect (e.g., Amir Kassim et al., 2018) or even a negative effect (e.g., Peterson and Mulligan, 2012; Mulligan and Peterson, 2013) on human memory. As for one, Amir Kassim et al. (2018) deduced that, unlike visual and a combination of auditory and visual repetitions, two auditory repetitions have no effect on the participants' recognition memory. Not only that, Peterson and Mulligan (2012) reported a negative repetition effect for the participants who went through a list of cue-target pairs twice compared to those who studied the pairs once only.

To further substantiate the findings, neurolinguists set out to record and examine the brain activity of the learners through electroencephalography (EEG) and event-related brain potential (ERP) techniques, evidencing that the human brain responds emphatically to any type of repetition (Van Strien et al., 2007). A review of the related literature reveals that such neurophysiological studies have mainly investigated the word repetition effect on the basis of pertinent ERP component modulations during the process of repetition (Henson, 2003; Maccotta and Buckner, 2004; Van Strien et al., 2007). Yet, the missing chain in the literature is how these word repetitions affect the overall comprehension of the learners.

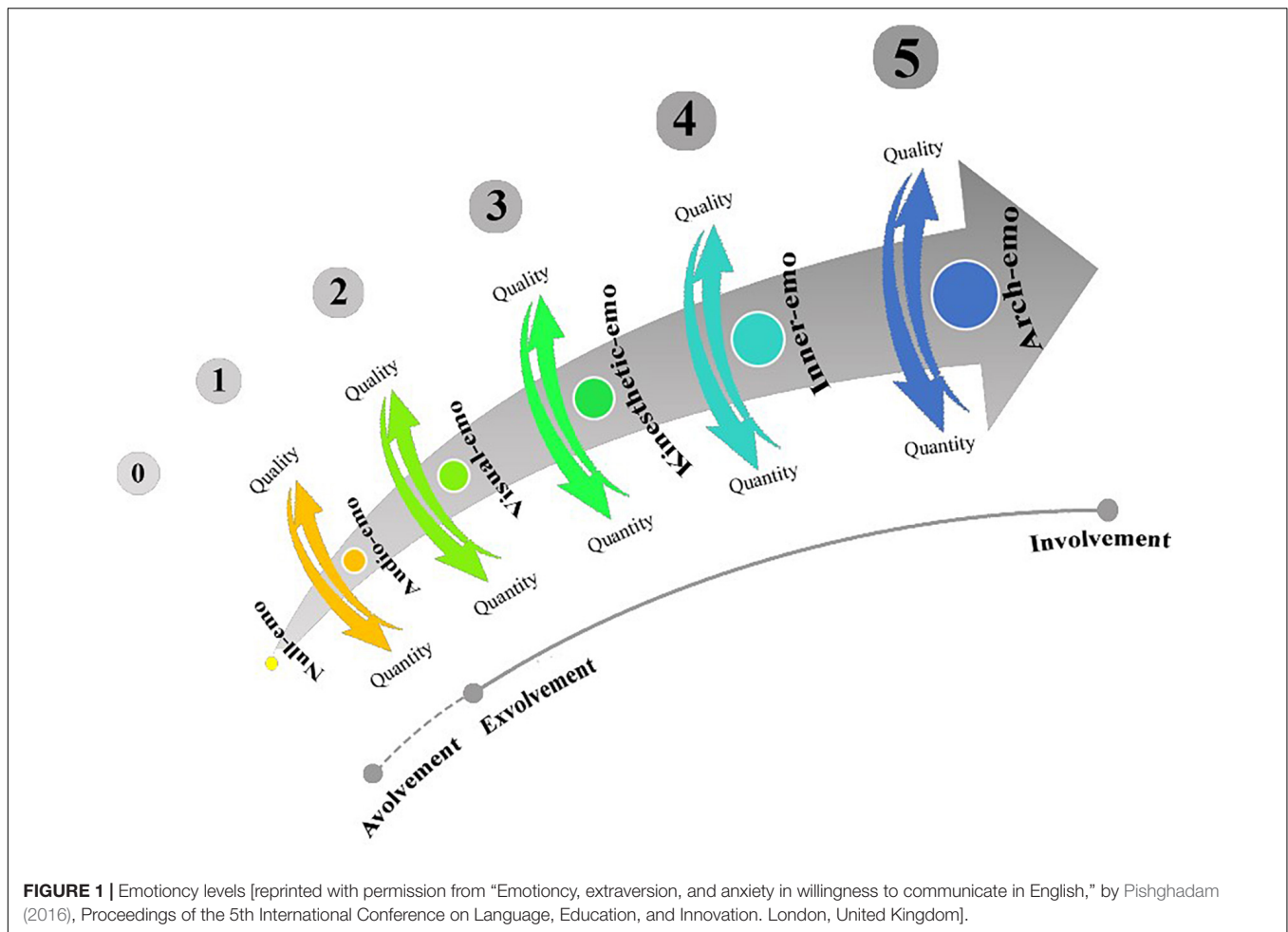
To delve into the electrophysiological underpinning of sentence comprehension, researchers (e.g., Hagoort et al., 2004; Hald et al., 2006; Kos et al., 2012) have designed different sentence acceptability judgment tasks with semantically violated (sentences with the word knowledge violation, e.g., a caper is kind.) and pragmatically violated sentences (sentences with the world knowledge violation, e.g., a caper is sweet.). According to their findings, the neurocognitive mechanism of sentence comprehension manifests itself in a series of ERP components, with N400 as the most general component, providing insights

into the neurobiology of meaning. The N400 effect, with its peak around 400 ms following the stimulus, is sensitive to semantic modifications (Xu et al., 2015; Payne et al., 2019). The amplitude of this negative-going deflection is basically defined by the degree of congruence between a word and its sentential context and the load of cognitive endeavor required to access the semantic memory (Kutas and Federmeier, 2000). The component is similarly influenced by other variables such as verbal working memory (e.g., Brown and Hagoort, 1999), word frequency (Dambacher et al., 2006), presentation modality (Kutas and Federmeier, 2011), and word priming (McRae et al., 2005). These amplitude changes may bring about different degrees of comprehension.

For improved comprehension, which is closely linked to vocabulary retention, not only vocabulary repetition practices but the nature of instruction may serve a pivotal role. It is commonly believed that a rudimentary path to deep processing and enhanced learning is the involvement of the senses. Underpinning the importance of senses in learning a language (Massaro, 2004), sensory teaching, pioneered by Montessori (1912), has been used by educators and teachers believing that senses, either in isolation or in different combinations, give way to inclusive learning, which engages all the learners with different needs (Hockings, 2010; Katai, 2011). Brain research findings have similarly corroborated the effect of multisensory instruction (MSI) on brain performance, particularly sentence comprehension, which is improved by the involvement of more sensory information channels and neural structures as a result of the interaction of more senses (Shams and Seitz, 2008; Pishghadam et al., 2020, 2021; Shayesteh et al., 2020).

What we hypothesize in this study is that, given the effectiveness of MSI in engendering in-depth learning (Baines, 2008), vocabulary repetition is likely to be redundant for sentence comprehension. To verify that, we used a validated sensory-based model of instruction, coined as *emotioncy* (a blend of emotion + frequency), and combined the five senses of auditory, visual, tactile, olfactory, and gustatory (see Pishghadam et al., 2017, 2020, 2021; Karami et al., 2019; Makiabadi et al., 2019; Shayesteh et al., 2020; Boustani et al., 2021). The model (**Figure 1**) presents us with two major combinations of the senses, namely *exvolvement* (i.e., a combination of auditory, visual, and tactile/kinesthetic) and *involvement* (i.e., a combination of auditory, visual, and tactile/kinesthetic, olfactory, and gustatory).

To decide upon the number of vocabulary repetitions following our MSI, we drew upon the findings of a recent, relevant study conducted by Jajarmi et al. (2020). They adopted a bisensory (auditory + visual) approach according to the emotioncy model and taught a list of unknown English vocabulary items to a group of language learners. They used different numbers of repetitions to eventually come up with the minimum number of effective repetitions. Quite in line with Hintzman (1970), Nelson (1977), and Phaf (2012), using paper and pencil tests, Jajarmi et al. (2020) reported two repetitions as the threshold for making significant changes in vocabulary learning and that, one, three, four, and five repetitions make no further contribution to vocabulary learning as a result of bisensory instruction.



For the purpose of this study, and to examine if sentence comprehension is influenced by two vocabulary repetitions (as the minimum number of effective repetitions) following the MSI, we selected six vocabulary items of which the learners had no previous knowledge. We taught the words to a control group and an experimental group of participants. The control group received the MSI only, whereas the experimental group had the MSI followed by two audio-visual repetitions of the vocabulary items (see section “The Instruction” for details on the procedure). Thereafter, we compared the immediate behavioral and cognitive performance of the two groups, on a sentence acceptability judgment task (with correct and pragmatically violated sentences), for any probable neural response differences associated with sentence comprehension. In order to evaluate the differences, the ERP technique was employed. Based on the previous studies acknowledging the efficiency of using multiple senses in the process of learning and comprehension, and the ERP studies recognizing the N400 as an indicator of semantic access difficulty, we predicated that two vocabulary repetitions following the MSI may not reduce the N400 amplitude and facilitate semantic access during sentence comprehension. Therefore, we expect to observe no N400 amplitude difference between the control and experimental groups, concluding that

MSI is a working theory that is not influenced by two vocabulary repetitions.

MATERIALS AND METHODS

Participants

Forty-five (33 female and 12 male) native speakers of Persian, with English as their foreign language, volunteered to take our pretests, 3 of whom were not recruited to participate in the ERP experiment due to their pretest results. Moreover, the data for two of the participants were discarded because of excessive eye movement and muscle artifact. The participants’ age ranged from 18 to 30 years ($M = 21.7$, $SD = 2.6$). They were all right-handed (Oldfield, 1971), neurologically healthy, and had normal or corrected-to-normal vision. All of them were at the intermediate level of language proficiency, and their working memory score ranged from 10 to 12 ($M = 11.3$, $SD = 1.3$) (Wechsler, 1981). They neither had participated in the pilot tests nor had any knowledge of the six selected vocabulary items they were supposed to learn. For the purpose of this study, the participants were randomly assigned to a control group (G1, $N = 20$) and an experimental group (G2, $N = 20$). The participants gave written informed

consent under a protocol approved by the Ferdowsi University of Mashhad Ethics Committee before the experiment and took part in the research according to their willingness to participate. They received either course credits or gifts for their participation.

Materials

Pretest Materials

The Emotioncy Scale

In order to make sure that the participants had no knowledge of the selected items for the experiment, an emotioncy scale was used (Borsipour, 2016). Each item measured the participants' familiarity with the target words through a 6-point Likert scale with (1) not familiar; (2) heard; (3) heard and seen; (4) heard, seen, and touched; (5) heard, seen, touched, and used; and (6) heard, seen, touched, used, and done research on. The participants who had prior experiences with any of the six words were excluded in this phase.

The Oxford Quick Placement Test

The Oxford Quick Placement Test (OQPT) (Allan, 1992) was administered to measure the participants' English proficiency level. This test has two parts, each containing 40 and 20 items, respectively. The items are in multiple-choice and cloze test formats, and the time to respond to the questions is 30 min. In this test, the obtained scores of 30–40 represent intermediate proficiency level in English.

The Digit Span Subtest of the Wechsler Adult Intelligence Scale III

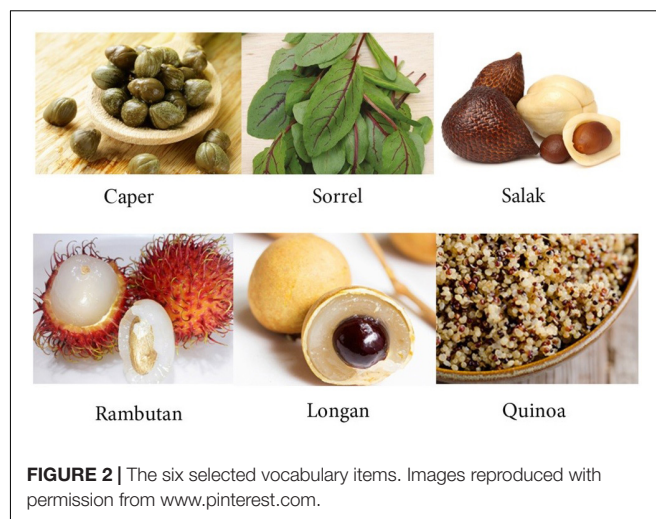
Since working memory and specifically the phonological loop plays a substantial role in vocabulary learning and vocabulary retention (Gillam, 2002), we used the digit span subtest of the Wechsler Adult Intelligence Scale III (WAIS-III) (Wechsler, 1981) as a measure of homogeneity. Given that based on the results, the mean span for those who took the test was 11 with a standard deviation of 1, we selected those participants within the limited range of 10–12.

The Edinburgh Inventory of Handedness

The Edinburgh inventory of handedness (Oldfield, 1971), as a measure of hand laterality, was used to select right-handed individuals. The inventory includes 10 questions about writing, drawing, throwing, using scissors, using a toothbrush, using a knife (without a fork), using a spoon, using a broom (upper hand), striking a match, and opening a box, along with two supplementary questions: "which foot do you prefer to kick with?" and "which eye do you use when using only one?" According to the scale, the participants who did more than two of the mentioned activities with their left hand were excluded from the study.

Stimulus Materials (Vocabulary Items)

In order to choose the six vocabulary items, a list of 48 words (along with their Persian translation) which were the names of some edible things, including fruits, plants, and vegetables, was culled and put into the emotioncy scale (Borsipour, 2016). It should be mentioned that the translations of the words were not cognates in the L1 of the participants. Then the scale was



randomly administered to 150 respondents (87 females and 63 males), who were different from the main participants of the study. Finally, six words of which 95% of the respondents had no prior knowledge, were selected for the MSI. The words were *caper*, *longan*, *sorrel*, *salak*, *rambutan*, and *quinoa* (Figure 2).

Procedure

The Sentence Acceptability Judgment Task

A sentence acceptability judgment task was constructed according to a framework presented by Hagoort et al. (2004), Hald et al. (2006), and Kos et al. (2012), using Psychophysics Toolbox Version 3 (PTB-3) for MATLAB (version 2015a, The MathWorks, MA, United States). The task required the participants to judge the acceptability of the sentences they saw word by word on the screen. The six instructed vocabulary items were embedded in a number of 144 sentences (72 sentence pairs with 3–8 words each). Each pair comprised acceptable and unacceptable conditions of a sentence: a non-violated correct sentence (Co) and a sentence with pragmatic violation (Pr). In order to avoid conditioning, 72 unrelated sentences of similar length, complexity, and structure (36 correct and 36 pragmatically violated), including the names of animals, fruits, and objects, were added as fillers, making 216 sentences in total (Table 1). Each sentence pair was identical except for one word only (i.e., the critical word), which appeared at the sentence-final position. The critical words were matched across the two conditions in terms of average length in characters, word

TABLE 1 | Example Sentences of the two different conditions in the sentence pairs in addition to the filler sentences.

Sentence type	Condition	Example sentences
Target	Correct	A salak looks like a <i>fig</i> .
	Pragmatically violated	A salak looks like a <i>cherry</i> .
Filler	Correct	A monkey has a tail.
	Pragmatically violated	A monkey has a horn.

The critical words are in italics. The target words are boldfaced.

class, bigram frequency, and cloze probability (checked by two native speakers).

The task consisted of three 10-min experimental blocks separated by two short breaks. In each block, each sentence was presented word by word in the center of the computer screen (**Figure 3**). The words were boldfaced in black lower case Times New Roman letters with 58-point font size against a light gray background. The first word of each sentence was capitalized, and the final word of each sentence was presented with a period. The viewing distance was about 100 cm for each participant. Each experimental block started with a 600 ms baseline before the stimulus onset. Each word was presented for 750–850 ms (randomly varied to avoid the predictability of the response time (RT) for participants in terms of their reaction times) followed by a blank screen for 300 ms as an inter-stimulus interval (before the appearance of the next word). After the final word, there was a blank screen for 2800 ms in which the participants were supposed to decide on the truthfulness of the sentences by pressing a key. They were asked to press the right arrow key on the computer keyboard if the sentence was correct, the down arrow key if the sentence was pragmatically wrong, and press no key if they did not know the response. After the response window, an eye image

was displayed in the center of the screen for 3000 ms, allowing the participants to move their eyes and blink intentionally to prevent eye fatigue. There was a 300-ms blank page between the eye disappearance and the start of the next trial.

The Instruction

The data collection was split into pre-experimental (i.e., instruction) and experimental (i.e., ERP recording) phases. During the pre-experimental phase, each participant learned the six vocabulary items through the MSI. For the purpose of the instruction, along with the real fruits and vegetables, a PowerPoint presentation, which contained the name and some different pictures of the items, and a photo booklet were used to make it similar to a classroom environment. The participants received sensory instruction through inner emotioncy, which is the integration of auditory, visual, kinesthetic, smell, and taste, for the six words. The same amount of time was allocated to the instruction of each word. The whole instruction took approximately 20 min. **Table 2** presents a sample instruction of one of the words and the information transferred to the participants.

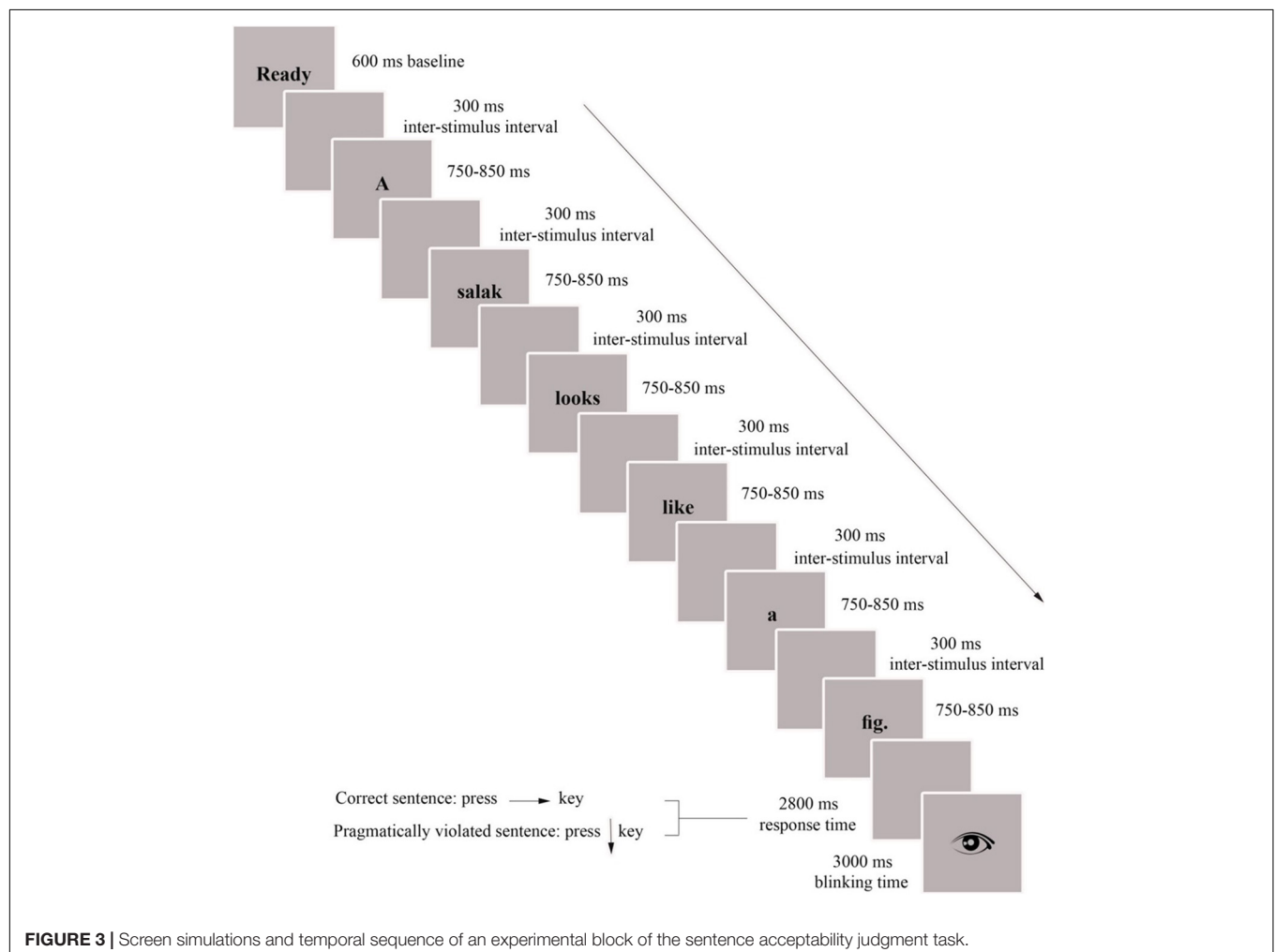


TABLE 2 | A sample instruction for an involved word (using a combination of auditory, visual, and tactile/kinesthetic, olfactory, and gustatory senses).

Salak (the participants had a salak with a plate and a knife to cut and smell. The instructor was providing the necessary information about the fruit at the same time).

Look at this fruit. This is a Salak. Salak has an alternative name which is "snake fruit." Guess why. Aha . . . because . . . you see . . . the skin looks like that of a snake, doesn't it? The skin is very thin but inedible. It is brown like a walnut. You see . . . the shape is almost like a fig. Tell me about the size. . . as you see it is as big as a lemon. Now peel it very gently. Try not to hurt the flesh. A salak has three big lobes. The lobes look like garlic. Does the fruit have any seeds? Cut the lobes to find it if any. . . Aha you see . . . There is a seed in one of the lobes only. How about the smell? Tasty? Taste it. It is very juicy.

After the instruction, the participants of G1 directly went for the ERP recording, yet the participants of G2 received two audio-visual repetitions for each word. For the first repetition, the participants of G2 were asked to look at the real objects on the desk and repeat the names after the instructor when the instruction was almost over. For the second repetition, they went through the same procedure (i.e., repetition) 15 min later when they had the EEG cap on. There was a time interval of 15 min between the second repetition and the ERP experiment, during which the participants went through a practice block of the task (see section "Electroencephalography Recording") and got ready for the recording. According to Cepeda et al. (2006), the lag between the repetitions and the retention interval need to approximately match for optimal memory performance.

Electroencephalography Recording

The participants were tested individually in a sound-attenuated and dimly illuminated room. Prior to the main experiment, they went through a practice block of 20 items to get acquainted with the task requirements.

The EEG was recorded from 23 active Ag/AgCl-sintered electrodes mounted on an elastic electrode cap (g.GAMMAcap from g.tec medical engineering GmbH). The electrodes were placed according to the 10–20 international system of the American Electroencephalographic Society over midline sites at Fz, FCz, Cz, Pz, and Oz; frontal sites at AF3, AF4, F3, F4, F7, and F8; fronto-central sites at FC3 and FC4; fronto-temporal sites at FT7 and FT8; central sites at C3 and C4; parietal sites at P3, P4, P7, and P8; and occipital sites at PO7 and PO8. The optimal electrode arrangement was determined according to similar studies in the field (e.g., Salmon and Pratt, 2002; Hagoort et al., 2004; Van Berkum et al., 2005; Hald et al., 2006). All electrodes were referenced to the left mastoid and re-referenced to the average of the left and right mastoids. Vertical and horizontal eye movements were monitored *via* three additional electrodes placed above and below the left eye and on the left outer canthus. Electrode impedances were kept below 5 k Ω . The EEG and EOG signals were digitized online with a sampling frequency of 250 Hz and were amplified using the 32-channel wireless g.Nautilus EEG system (gtec, Austria), with a bandpass filter between 0.1 and 70 Hz and a notch filter of 50 Hz.

Data Analysis

All recorded EEG signal data were imported into MATLAB (MathWorks, 2015b). To analyze the waveforms, MATLAB along with the EEGLAB (an extension of MATLAB software) were used. The EEG data were bandpass filtered between 0.5 and

60 Hz. Afterward, the data were re-referenced to the mean of the linked mastoids. Poor EEG channels were replaced with their interpolated version applied to the remaining channels. No more than 2 channels were interpolated for each participant, with the majority of the interpolated channels positioned at parieto-occipital and occipital sites. High amplitude eye blinks and muscle artifacts were then removed using the Artifact Subspace Reconstruction (ASR) algorithm from the EEGLAB. The remaining high frequencies were eliminated using a low pass filter with a cut-off frequency of 25 Hz. Next, epochs from –200 to 1100 ms, with respect to the onset of the critical word, were segmented to a 200 ms pre-onset baseline. A linear detrend algorithm (using the 200 ms before the stimulus onset to 3 s after) was applied to the epoched data to further remove drifts. Noisy epochs with potentials exceeding $\pm 70 \mu\text{V}$ were rejected. Finally, all remaining trials (see Table 3 for the descriptive statistics) were averaged.

For the critical words of the sentences, N400 was analyzed. Based on the findings of the previous literature (e.g., Danko et al., 2014; Yang et al., 2014; Molinaro et al., 2016; Volz et al., 2019), the N400 component was quantified as the mean amplitude in a

TABLE 3 | Descriptive statistics for the number of averaged epochs.

Condition	Group	Min	Max	Mean (out of 72 items)	SD
Correct	G1	49	59	55.68	1.23
	G2	50	57	54.23	2.87
Pragmatically violated	G1	50	60	56.01	0.85
	G2	51	59	56.94	2.21

Overall there were 72 epochs to average for every condition of each group. Yet, we only averaged the ones to which the participants gave right answer.

TABLE 4 | Descriptive statistics of response accuracy scores and response times for G1 and G2.

	Linguistic condition	Group ^a	Mean (for 72 items)	SD
Response accuracy	Correct	G1	59.05	5.64
		G2	58.85	6.86
	Pragmatically violated	G1	60.80	7.38
		G2	61.95	5.40
	Correct	G1	0.96	0.21
		G2	0.96	0.25
Response time (s)	Pragmatically violated	G1	0.99	0.22
		G2	0.99	0.24

^aN = 20.

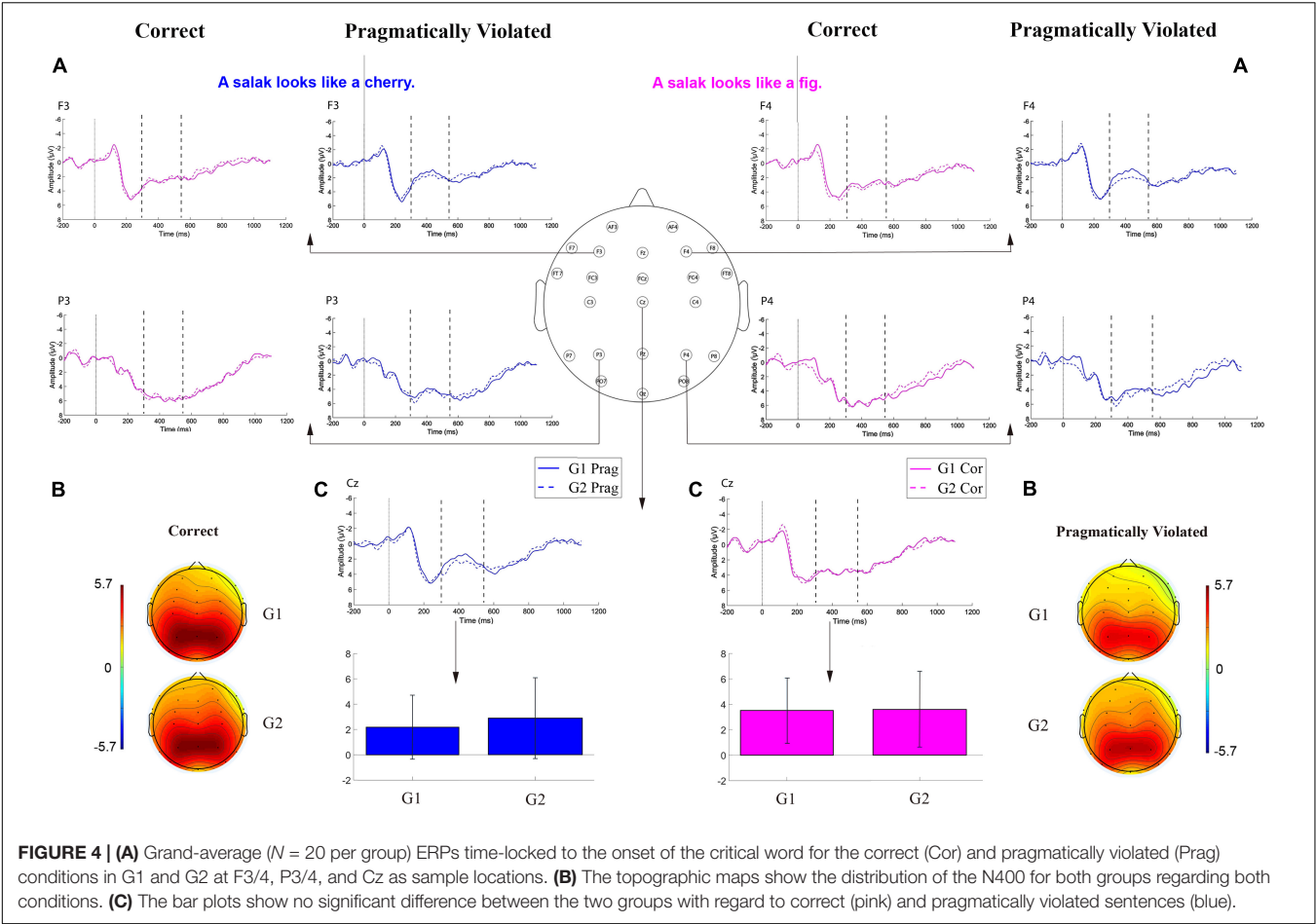


FIGURE 4 | (A) Grand-average ($N = 20$ per group) ERPs time-locked to the onset of the critical word for the correct (Cor) and pragmatically violated (Prag) conditions in G1 and G2 at F3/4, P3/4, and Cz as sample locations. **(B)** The topographic maps show the distribution of the N400 for both groups regarding both conditions. **(C)** The bar plots show no significant difference between the two groups with regard to correct (pink) and pragmatically violated sentences (blue).

TABLE 5 | Bayesian repeated measures model comparison for the N400.

Models	P(M)	P(M data)	BF _M	BF ₁₀	Error %
Electrode + condition	0.05	0.70	42.93	1.00	
Electrode + condition + group	0.05	0.17	3.87	0.25	2.48
Electrode + condition + group + condition × group	0.05	0.11	2.41	0.16	3.86
Electrode + condition + group + electrode × group	0.05	2.41e−5	4.34e−4	3.42e−5	2.37
Electrode + condition + group + electrode × group + condition × group	0.05	1.77e−5	3.18e−4	2.51e−5	8.90
Electrode + condition + electrode × condition	0.05	1.19e−5	2.14e−4	1.69e−5	2.31
Electrode + condition + group + electrode × condition	0.05	3.15e−6	5.67e−5	4.47e−6	3.93
Electrode + condition + group + electrode × condition + condition × group	0.05	2.39e−6	4.30e−5	3.39e−6	11.57
Electrode + condition + group + electrode × condition + electrode × group	0.05	4.55e−10	8.19e−9	6.45e−10	3.92
Electrode	0.05	3.98e−10	7.17e−9	5.66e−10	2.03
Electrode + condition + group + electrode × condition + electrode × group + condition × group	0.05	3.16e−10	5.70e−9	4.49e−10	5.04
Electrode + Group	0.05	9.95e−11	1.79e−9	1.41e−10	2.15
Electrode + condition + group + electrode × condition + electrode × group + condition × group + electrode × condition × group	0.05	3.13e−14	5.64e−13	4.45e−14	3.00
Electrode + group + electrode × group	0.05	1.27e−14	2.29e−13	1.80e−14	2.31
Condition	0.05	9.08e−101	1.63e−99	1.28e−100	2.53
Condition + group	0.05	1.97e−101	3.55e−100	2.80e−101	2.31
Condition + group + condition × group	0.05	7.56e−102	1.36e−100	1.07e−101	2.80
Null model (including subject)	0.05	2.52e−107	4.55e−106	3.58e−107	2.01
Group	0.05	5.72e−108	1.03e−106	8.12e−108	2.26

latency window of 300–550 ms after the onset of the critical word and relative to a 200 ms baseline.

For the behavioral data analysis, a one-way between-groups analysis of variance (ANOVA) was performed to examine the main effect of repetition on the participants' acceptability judgment of the correct and pragmatically violated sentences. For the ERP data analysis, as to the main effect of repetition, a multivariate analysis of variance (MANOVA) along with a Bayesian Repeated Measures ANOVA were run to test for the null hypothesis. All the statistical procedures were computed using an alpha level of 0.05, and to reduce the chance of a Type 1 error, Bonferroni adjustment was applied.

RESULTS

Behavioral Results

Descriptive statistics of response accuracy (RA) scores and RTs for G1 and G2 are given in **Table 4**.

A one-way between-groups ANOVA was performed to examine the main effect of repetition on the participants' acceptability judgment of the correct and pragmatically violated sentences. The results of Levene's test for homogeneity of variances showed that the assumption of homogeneity of variance was not violated ($p > 0.05$). The difference between the two groups did not reach significance regarding the main effect of repetition on RA in terms of both correct [$F(1,38) = 0.21, p = 0.92, \eta_p^2 = 0.14$] and pragmatically violated [$F(1,38) = 0.32, p = 0.58, \eta_p^2 = 0.18$] conditions.

Similarly, the results of RTs failed to show a significant difference between G1 and G2 in the amount of time it took them to judge the truthfulness of the sentences in terms of either of the conditions, that is, correct [$F(1,38) = 0.33, p = 0.97, \eta_p^2 = 0.21$] and pragmatically violated [$F(1,38) = 0.19, p = 0.89, \eta_p^2 = 0.19$].

Event-Related Brain Potential Results

At the next step, we analyzed the participants' brain activity, recorded during the task, to find out whether they were in line with the behavioral results. **Figure 4** shows grand-average ERPs time-locked to the onset of the critical word for the correct and pragmatically violated conditions in G1 and G2 at F3, F4, P3, P4, and Cz as sample electrode sites from the 23 inspected locations on the scalp. The grand-average waveforms for G1 and G2 showed different neural correlates of sentence processing, including a broadly distributed negative-going deflection (representing the N400) starting at about 300 ms after the onset of the critical word peaking at 400 ms. To find the location of the maximum N400 amplitude, we grouped the 23 electrodes into three regions of interest (ROIs): anterior (AF3/4, F3/4, F7/8, FT7/8, Fz), central (FC3/4, FCz, C3/4, Cz), and posterior (P3/4, P7/8, Pz, PO7/8, Oz). The F test result [$F(2,78) = 17.02, p = 0.000; \eta_p^2 = 0.30$] revealed that the negativity in the anterior areas ($M = 1.73 \mu\text{v}$) was significantly larger than that of the central ($M = 2.98 \mu\text{v}$) and posterior regions ($M = 3.97 \mu\text{v}$).

Consequently, the two groups were compared on the N400 effect at the 23 electrodes. To this end, a one-way between-groups MANOVA was run to investigate group differences in the N400 mean amplitude in terms of the two linguistic conditions (i.e., correct and pragmatically violated). Preliminary assumption testing noted no violations of normality and homogeneity of variance-covariance matrices ($p > 0.05$). The multivariate effect of group regarding the N400 mean amplitudes signaled no significant effect for repetition in terms of the groups [$F(8, 31) = 0.39, p = 0.97$; Wilks' Lambda = 0.40; $\eta_p^2 = 0.60$].

To ensure the lack of difference between the mean amplitudes, the Bayesian Repeated Measures ANOVA [factors: electrode (23), condition (2); between-subject factor: group (G1 & G2)] was used. As **Tables 5, 6** show, Bayesian analyses support the null hypothesis, indicating no main or interaction effect of the group (i.e., repetition).

DISCUSSION

Different studies reinforce that sentence comprehension is influenced by various factors, including MSI (Pishghadam et al., 2020, 2021; Shayesteh et al., 2020; Boustani et al., 2021). Given that this type of instruction is believed to activate several regions of the brain, it is considered one of the most effective ways of teaching. In order to explore if this whole-brain instruction technique may further improve learners' L2 sentence comprehension, if it is followed by two vocabulary repetitions, we adopted a neurocognitive approach and used the ERP method.

The overall behavioral (RA & RT) and electrophysiological findings of the study supported our preliminary hypothesis that, unlike Jajarmi et al.'s (2020) findings, two repetitions may not actually boost the L2 vocabulary knowledge obtained through the MSI. To teach L2 vocabulary items to adults mainly, teachers generally adopt a unisensory approach (audition only), representing what Shayesteh et al. (2019) refer to as *thin education*, or integrate auditory and visual senses (bisensory instruction) and disregard senses of touch, taste, and smell due to limited time and instructional facilities. The influence of vocabulary repetition for this type of learning may not probably be analogous to the one following the MSI. Thus, confirming Nation's (2001, p. 115) proposition that "the nature of the original learning" determines later retrieval, we suggest that the nature of instruction and learning modulates, to

TABLE 6 | Analysis of effects.

Effects	P(incl)	P(excl)	P(incl data)	P(excl data)	BF _{incl}
Electrode	0.73	0.26	1.00	0.00	∞
Condition	0.73	0.26	1.00	4.98e-10	7.16e+8
Group	0.73	0.26	0.29	0.70	0.15
Electrode \times condition	0.31	0.68	1.74e-5	1.00	3.78e-5
Group \times electrode	0.31	0.68	4.18e-5	1.00	9.06e-5
Group \times condition	0.31	0.68	0.11	0.88	0.29
Group \times electrode \times condition	0.05	0.94	3.13e-14	1.00	5.64e-13

a considerable extent, the effect of subsequent instructional practices – like vocabulary repetition – employed to deepen student learning. Consequently, unlike the studies reporting that learning requires multiple repetitions (e.g., Webb, 2007), we put forward that the (two)repetition effect following multisensory instructional practices seems to be ineffective on L2 sentence comprehension. The justification may probably have its roots in the neurophysiological mechanism underlying the process of learning through multiple senses. The knowledge of words builds up through sensory experiences. That is, different kinds of input that enter the brain through different modalities cumulate and form a comprehensive whole (Shams and Seitz, 2008; Katai, 2011). Throughout this process, greater sensory information gateways and neural networks are activated, and more extended areas of the brain (including sensory-specific and multisensory convergence zones) are engaged (Driver and Noesselt, 2008; Goswami, 2008). Repetition, on the other hand, reinforces the neural connections across the synapses, strengthening the link between form and meaning (Gathercole and Baddeley, 1990), and adds to the quality of knowledge (Nation, 2001), which, in fact, facilitates later retrieval and comprehension. However, given that activating a small proportion of neurons in large scale networks may not affect the network output (Parker, 2010), we assume that the effect of two repetitions, with a short delay period, is so subtle that it does not modify the extensive network of neurons activated through the multiple senses approach.

Holding the view that senses may have a profound impact on vocabulary learning and sentence comprehension, we intend to suggest that, rather than challenging the irrefutable role of repetition as a step forward in language learning, the substantial role of senses in this process needs to be underlined, more than the past, by incorporating sensory-based models, such as emotioncy (Pishghadam et al., 2017; Karami et al., 2019; Makiabadi et al., 2019), into the regular curriculum. Such models lead the old multisensory movement toward entering a new phase which may open up new vistas for teachers and educators. Despite the time-consuming nature of multisensory education, it is strongly recommended that teachers make use of senses in their teaching practices, believing that the considerable merits compensate for the extra efforts exerted by teachers. Another pedagogical implication of the current study could be that two vocabulary repetitions seem redundant after multisensory learning, as opposed to bisensory learning (Jajarmi et al., 2020), since we witnessed no significant change between the cognitive reactions of the two groups. However, raising the number of repetitions may perhaps produce very different results. Therefore, repetition should not be entirely overlooked; instead, it should be more carefully probed and applied.

In the end, there are a few points that need to be taken into careful consideration in future endeavors. First and foremost is that, in addition to the effect of repetition happening in close temporal succession, the effect of spaced repetition with increasingly larger intervals on multisensory vocabulary learning should be meticulously verified since it is believed to produce improved long-term results (Rawson and Kintsch, 2005). Therefore, spacing of the repetitions may allow for probable differences in later performance, generating different results. Moreover, given that the task we used in this study checked the participants' receptive vocabulary knowledge only, further research is required to investigate the effect of two repetitions on productive vocabulary knowledge as well. Last but not least, as a complementary action, future studies need to compare the cognitive processes underlying repetition after bisensory and MSI. It is needless to say that increasing the sample size may lead to more reliable conclusions.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ferdowsi University of Mashhad Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RP conceived and designed the experiments. HJ, SS, and AK performed the experiments. HJ and AK analyzed the data. RP, HJ, SS, and AK contributed to reagents, materials, and analysis tools. HJ, SS, and HN wrote the manuscript. RP and HN reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Iran's National Elites Foundation (INEF), the Cognition and Sensory Emotion Lab, Ferdowsi University of Mashhad, and a grant-in-aid of research from the Cognitive Sciences and Technologies Council (CSTC) in 2016 (Code 4251).

REFERENCES

Allan, D. (1992). *The Oxford Quick Placement Test*. Oxford: Oxford University Press.

Amir Kassim, A., Rehman, R., and Price, J. M. (2018). Effects of modality and repetition in a continuous recognition memory task: repetition has no effect on auditory recognition memory. *Acta Psychol.* 185, 72–80. doi: 10.1016/j.actpsy.2018.01.012

- Baines, L. (2008). *A Teacher's Guide to Multisensory Learning: Improving Literacy by Engaging the Senses*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188.
- Borsipour, B. (2016). *Emotioncy and Willingness to Read: A Case of Iranian EFL Learners*. Ph.D. thesis. Mashhad: Ferdowsi University of Mashhad.
- Boustani, N., Pishghadam, R., and Shayesteh, S. (2021). Multisensory input modulates P200 and L2 sentence comprehension: a one-week consolidation phase. *Front. Psychol.* 12:746813. doi: 10.3389/fpsyg.2021.746813
- Brown, C. M., and Hagoort, P. (eds) (1999). *The Neurocognition of Language*. Oxford: Oxford University Press.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354
- Chen, C., and Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Appl. Linguist.* 31, 693–713. doi: 10.1093/applin/amq031
- Dambacher, M., Kliegl, R., Hofmann, M., and Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Res.* 1084, 89–103. doi: 10.1016/j.brainres.2006.02.010
- Danko, S. G., Boytsova, J. A., Solovjeva, M. L., Chernigovskaya, T. V., and Medvedev, S. V. (2014). Event-related brain potentials when conjugating Russian verbs: the modularity of language procedures. *Hum. Physiol.* 40, 237–243. doi: 10.1134/s0362119714030050
- Dapretto, M., and Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron* 24, 427–432. doi: 10.1016/s0896-6273(00)80855-7
- Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron* 57, 11–23. doi: 10.1016/j.neuron.2007.12.013
- Embrick, D., Marantz, A., Miyashita, Y., O'Neill, W., and Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6150–6154. doi: 10.1073/pnas.100098897
- Gathercole, S. E., and Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *Br. J. Psychol.* 81, 439–454. doi: 10.1111/j.2044-8295.1990.tb02371.x
- Gillam, R. B. (2002). The role of working memory in vocabulary development. *Perspect. Lang. Learn. Educ.* 9, 7–11.
- Goswami, U. (2008). Principles of learning, implications for teaching: a cognitive neuroscience perspective. *J. Philos. Educ.* 42, 381–399. doi: 10.1111/j.1467-9752.2008.00639.x
- Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441.
- Hald, L. A., Bastiaansen, M. C., and Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain Lang.* 96, 90–105. doi: 10.1016/j.bandl.2005.06.007
- Henson, R. N. A. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/s0301-0082(03)00086-8
- Hidalgo, M. A., and Garcia Mayo, M. P. (2021). The influence of task repetition type on young EFL learners' attention to form. *Lang. Teach. Res.* 25, 565–586. doi: 10.1177/1362168819865559
- Hintzman, D. L. (1970). Effects of repetition and exposure duration on memory. *J. Exp. Psychol.* 83, 435–444. doi: 10.1037/h0028865
- Hockings, C. (2010). *Inclusive Learning and Teaching in Higher Education: A Synthesis of Research*. New York, NY: Higher Education Academy.
- Horst, J. S. (2013). Context and repetition in word learning. *Front. Psychol.* 4:149. doi: 10.3389/fpsyg.2013.00149
- Horst, M., Cobb, T., and Meara, P. (1998). Beyond a clock orange: acquiring second language vocabulary through reading. *Read. Foreign Lang.* 11, 207–223.
- Jajarmi, H., Pishghadam, R., and Moghimi, S. (2020). Barrasi tedade moasere tekrare vaje dar yadgiriye vajegane zabane englisi be onvane zabane khareji [Investigating the effective number of word repetitions on learning the vocabulary of English as a foreign language]. *Lang. Transl. Stud.* 54, 1–26.
- Karami, M., Pishghadam, R., and Baghaei, P. (2019). A probe into EFL learners' emotioncy as a source of test bias: insights from differential item functioning analysis. *Stud. Educ. Eval.* 60, 170–178. doi: 10.1016/j.stueduc.2019.01.003
- Kartchava, E., and Nassaji, H. (2019). The role of task repetition and learner self-assessment in technology-mediated task performance. *Int. J. Appl. Linguist.* 170, 180–203. doi: 10.1075/itl.19013.kar
- Katai, Z. (2011). Multi-sensory method for teaching-learning recursion. *Comput. Appl. Eng. Educ.* 19, 234–243. doi: 10.1002/cae.20305
- Kos, M., Van den Brink, D., and Hagoort, P. (2012). Individual variation in the late positive complex to semantic anomalies. *Front. Psychol.* 3:318. doi: 10.3389/fpsyg.2012.00318
- Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Laufer, B., and Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: the effects of task type, word occurrence and their combination. *Lang. Teach. Res.* 15, 391–411.
- Liu, S. (2018). Review on incidental vocabulary acquisition through reading. *Theory Pract. Lang. Stud.* 8, 702–711. doi: 10.17507/tpls.080622
- Maccotta, L., and Buckner, R. L. (2004). Evidence for neural effects of repetition that directly correlate with behavioral priming. *J. Cogn. Neurosci.* 16, 1625–1632. doi: 10.1162/0899829042568451
- Makiabadi, H., Pishghadam, R., Naji Meidani, E., and Khajavy, G. H. (2019). Examining the role of emotioncy in willingness to communicate: a structural equation modeling approach. *Rev. Psicodidact.* 24, 120–130.
- Massaro, D. W. (2004). "From multisensory integration to talking heads and language learning," in *The Handbook of Multisensory Processes*, eds G. A. Calvert, C. Spence, and B. E. Stein (Cambridge, MA: Massachusetts Institute of Technology), 311–332. doi: 10.1007/s00426-005-0031-5
- McRae, K., Hare, M., Elman, J. L., and Ferretti, T. R. (2005). A basis for generating expectancies for verbs from nouns. *Mem. Cogn.* 33, 1174–1184. doi: 10.3758/bf03193221
- Molinari, N., Su, J. J., and Carreiras, M. (2016). Stereotypes override grammar: social knowledge in sentence comprehension. *Brain Lang.* 155, 36–43. doi: 10.1016/j.bandl.2016.03.002
- Montessori, M. (1912). *The Montessori Method*. New York, NY: Cambridge Press.
- Mulligan, N. W., and Peterson, D. J. (2013). The negative repetition effect. *J. Exp. Psychol.* 39, 1403–1416. doi: 10.1037/a0031789
- Nation, I. S. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge Press.
- Nation, I. S. (2015). Principles guiding vocabulary learning through extensive reading. *Read. Foreign Lang.* 27, 136–145.
- Nelson, T. O. (1977). Repetition and depth of processing. *J. Verb. Learn. Verb. Behav.* 16, 151–171. doi: 10.3758/BF03213234
- Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., and Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *J. Cogn. Neurosci.* 24, 1205–1223. doi: 10.1162/jocn_a_00143
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Parker, D. (2010). Neuronal network analyses: premises, promises and uncertainties. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2315–2328. doi: 10.1098/rstb.2010.0043
- Payne, B. R., Stites, M. C., and Federmeier, K. D. (2019). Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading. *Psychophysiology* 56, 1–15. doi: 10.1111/psyp.13432
- Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Lang. Teach. Res.* 18, 75–94. doi: 10.1177/1362168813505384
- Peterson, D. J., and Mulligan, N. W. (2012). A negative repetition effect in episodic memory. *J. Exp. Psychol.* 38, 1786–1791. doi: 10.1037/a0028220
- Phaf, R. H. (2012). *Learning in Natural and Connectionist Systems: Experiments and a Model*. Berlin: Springer.

- Pishghadam, R. (2016). "Emotioncy, extraversion, and anxiety in willingness to communicate in English," in *Proceedings of the 5th International Conference on Language, Education, and Innovation*, eds W. A. Lokman, F. M. Fazidah, I. Salahuddin, and I. A. W. Mohd (London: Infobase Creation Sdn Bhd), 1–5. doi: 10.1002/9781118784235.eelt0980
- Pishghadam, R., Baghaei, P., and Seyednozadi, Z. (2017). Introducing emotioncy as a potential source of test bias: a mixed Rasch modeling study. *Int. J. Test.* 17, 127–140. doi: 10.1080/15305058.2016.1183208
- Pishghadam, R., Daneshvarfar, F., and Shayesteh, S. (2020). Oscillatory neuronal dynamics during L2 sentence comprehension: the effects of sensory enrichment and semantic incongruency. *Lang. Cogn. Neurosci.* 36, 903–920. doi: 10.1080/23273798.2021.1886312
- Pishghadam, R., Jajarmi, H., and Shayesteh, S. (2021). Sense combinations influence the neural mechanism of L2 comprehension in semantically violated sentences: insights from emotioncy. *J. Neurolinguist.* 58:100962. doi: 10.1016/j.jneuroling.2020.100962
- Rawson, K. A., and Kintsch, W. (2005). Rereading effects depend on time of test. *J. Educ. Psychol.* 97, 70–80. doi: 10.3389/jpsyg.2018.02517
- Salmon, N., and Pratt, H. (2002). A comparison of sentence- and discourse-level semantic processing: an ERP study. *Brain Lang.* 83, 367–383. doi: 10.1016/S0093-934X(02)00507-2
- Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends Cogn. Sci.* 12, 411–417. doi: 10.1016/j.tics.2008.07.006
- Shayesteh, S., Pishghadam, R., and Khodaverdi, A. (2020). FN400 and LPC responses to different degrees of sensory involvement: a study of sentence comprehension. *Adv. Cogn. Psychol.* 16, 45–58. doi: 10.5709/acp-0283-6
- Shayesteh, S., Pishghadam, R., and Moghimi, S. (2019). Extending the boundaries of multisensory teaching: an introduction to the dual-continuum model of sensorial education. *Int. J. Innov. Creat. Change* 4, 64–82.
- Thalheimer, W. (2003). *The Learning Benefits of Questions*. Available online at: <https://www.worklearning.com/wp-content/uploads/2017/10/Learning-Benefits-of-Questions-2014-v2.0.pdf> (accessed April 21, 2018).
- Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *J. Exp. Psychol.* 31, 443–467. doi: 10.1037/0278-7393.31.3.443
- Van Strien, J. W., Verkoeijen, P. P. J. L., Van der Meer, N., and Franken, I. H. A. (2007). Electrophysiological correlates of word repetition spacing: ERP and induced band power old/new effects with massed and spaced repetitions. *Int. J. Psychophysiol.* 66, 205–214. doi: 10.1016/j.ijpsycho.2007.07.003
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Lang. Learn.* 61, 219–258. doi: 10.1111/j.1467-9922.2010.00593.x
- Volz, K., Stark, R., Vaitl, D., and Ambach, W. (2019). Event-related potentials differ between true and false memories in the misinformation paradigm. *Int. J. Psychophysiol.* 135, 95–105. doi: 10.1016/j.ijpsycho.2018.12.002
- Waring, R., and Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Read. Foreign Lang.* 15, 130–163.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Appl. Linguist.* 28, 46–65. doi: 10.1093/applin/aml048
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale: WAIS-R Manual*. New York, NY: Harcourt Brace Jovanovich.
- Xu, G., Zhong, W., Jin, H., and Mo, L. (2015). An ERP study on how subsequent sentence context can influence previous world knowledge constraints. *J. Neurolinguist.* 33, 96–103. doi: 10.1016/j.jneuroling.2014.09.003
- Yang, J. C., Chi, L., Teichholtz, S., Schneider, A., Nanakul, R., Nowacki, R., et al. (2014). ERP abnormalities elicited by word repetition in fragile X-associated tremor/ataxia syndrome (FXTAS) and amnesic MCI. *Neuropsychologia* 63, 34–42. doi: 10.1016/j.neuropsychologia.2014.08.001
- Zheng, X., and Lemhöfer, K. (2019). The "semantic P600" in second language processing: when syntax conflicts with semantics. *Neuropsychologia* 127, 131–147. doi: 10.1016/j.neuropsychologia.2019.02.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Pishghadam, Jajarmi, Shayesteh, Khodaverdi and Nassaji. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



School Belonging and Reading Literacy: A Multilevel Moderated Mediation Model

Yuting Tan, Zhengcheng Fan, Xiaoman Wei and Tao Yang*

Collaborative Innovation Center of Assessment Toward Basic Education Quality, Beijing Normal University, Beijing, China

OPEN ACCESS

Edited by:

Olivia López Martínez,
University of Murcia, Spain

Reviewed by:

Qianqian Pan,
The University of Hong Kong, Hong
Kong SAR, China
Mwarumba Mwavita,
Oklahoma State University,
United States

*Correspondence:

Tao Yang
yangtao@bnu.edu.cn

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 16 November 2021

Accepted: 10 January 2022

Published: 02 February 2022

Citation:

Tan Y, Fan Z, Wei X and Yang T (2022)
School Belonging and Reading
Literacy: A Multilevel Moderated
Mediation Model.
Front. Psychol. 13:816128.
doi: 10.3389/fpsyg.2022.816128

School belonging is of great significance to students' physical and mental health development, especially academic improvement. However, the mechanism of the influence of school belonging on student academic achievement should be further explored, especially reading performance. Based on ecological systems theory and self-determination theory, the present research constructs a multilevel design to examine a moderated mediation model in which school belonging as a level-1 predictor, mastery goal orientation as a level-1 mediator and school disciplinary climate as a level-2 moderator jointly impact reading literacy. Results of the study were based on the questionnaires from 11,364 (5,455 girls and 5,909 boys) 15-year-olds nested in 332 schools in mainland China that participated in PISA 2018. The cross-sectional analysis indicated that: (1) school belonging had a direct and positive effect on student reading literacy; (2) the relationship between school belonging and reading literacy was prominently mediated by mastery goal orientation; (3) both school disciplinary climate level and strength could negatively moderate the latter half path of "school belonging → mastery goal orientation → reading literacy." Implications, limitations, and future research directions are discussed.

Keywords: PISA 2018, school belonging, reading literacy, mastery goal orientation, school disciplinary climate

INTRODUCTION

Academic achievement is an important indicator to evaluate student learning status and the quality of school education. Thus, the study of academic achievement and its influencing factors has always been the general concern of the whole society (Poropat, 2009). Nowadays, increasing researchers focus on the unique role of emotional factors in the cognitive development of students, such as school belonging (Liu and Lu, 2011; Kiefer et al., 2015; Froiland et al., 2016; Okilwa, 2016; Reynolds et al., 2017; Abdollahi and Noltemeyer, 2018; Korpershoek et al., 2019; Dixson, 2020; Huang, 2020; Koyuncu and Firat, 2020; Li et al., 2020). A substantial body of studies conducted in the last decade generally supported the point that school belonging is positively linked with students' learning motivation, mental resilience (Kim and Kim, 2013), growth mindset (Dixson, 2020), and academic success (Korpershoek et al., 2019), while significant negative relations with school belonging and school dropout, absenteeism, violence and, bullying, etc., also were proved (Gini et al., 2018; Korpershoek et al., 2019; Huang, 2020).

Despite prior research literature have established the positive relationship between a student's school belonging and academic achievement (Kiefer et al., 2015; Okilwa, 2016; Reynolds et al., 2017; Abdollahi and Noltemeyer, 2018; Korpershoek et al., 2019; Huang, 2020), several theoretical

voids remain. Firstly, domain specificity has been observed in a lot of literature with specific effects of predictors (e.g., gender, socioeconomic status) on different subject domains such as science, reading, and math (Hinnant et al., 2009; Reynolds et al., 2017). However, most existing studies focused on the impact of school belonging on academic achievement that represented a combination of several subjects (Okilwa, 2016; Abdollahi and Noltemeyer, 2018; Dixon, 2020; Li et al., 2020), such as the combination of mathematics, reading, and science (Li et al., 2020) or the combination of mathematics and reading (Okilwa, 2016). Additionally, some studies had also assessed this issue in mathematics (Ho, 2005; Froiland et al., 2016; Reynolds et al., 2017). Thus, there is a need to explore the possible different relations between school belonging and academic outcomes in other disciplinary domains, especially in reading. One reason is that reading literacy is not only a foundation for achievement in other subject areas within the educational system but also a prerequisite for successful participation in most areas of adult life (OECD, 2019; Koyuncu and Firat, 2020). Another is that reading is considered to be a learning experience, which is infused with more emotions compared with other subjects. Nevertheless, little research had been done to focus on the role of emotional factors in reading achievement (Hamedi et al., 2020; Zaccoletti et al., 2020). To fill this gap, the first purpose of this study is to identify the effect of school belonging on student reading performance with a large-scale survey.

Secondly, previous studies have paid more attention to the direct impact of school belonging on academic achievement (Ho, 2005; Liu and Lu, 2011; Okilwa, 2016; Reynolds et al., 2017; Korpershoek et al., 2019; Dixon, 2020; Koyuncu and Firat, 2020), or regarded school belonging as a mediating variable affecting academic outcomes by some variables (Froiland et al., 2016; Huang, 2020; Li et al., 2020), but how and when school belonging affect academic achievement is little known. Several empirical investigations have shown that motivation has been a dominant perspective to explain academic achievement. What's more, reading motivation can better explain the differences in students' reading comprehension than other variables (Rogiers et al., 2020; Ma et al., 2021). Moreover, multiple prior studies supported that school belonging can significantly predict students' learning motivation (e.g., mastery goal orientation) (Neel and Fuligni, 2013; Korpershoek et al., 2019). For instance, a meta-analysis showed that mastery goal orientation, considered to be a proactive and spontaneous learning tendency, was significantly correlated with school belonging, while other kinds of motivations were not (Korpershoek et al., 2019). To the best of the knowledge of the authors, the mediating effect of mastery goal orientation between these two variables has not yet been explored. Thus, another objective of this research is to understand the internal mechanism of this issue, since this could lead to comprehension of how emotional and motivational factors are related to changes in reading achievement.

From the view of constructivism, the specific interaction between learners and environmental factors influences learning outcomes (Chi et al., 2018). For instance, Reynolds et al. (2017) suggested that school climate and school identification (connectedness, belonging, relatedness) are conceptualized as

distinct but related concepts and should be investigated in the same model to explain academic development. Speaking of the school environment, one of the key environmental elements is the disciplinary climate (Cheema and Kitsantas, 2014, 2016; Sortkær and Reimer, 2016; Chi et al., 2018), which refers to the degree to which noise and disorder are suppressed, teachers have more time to cover the curriculum and use diverse teaching strategies, and students have time and opportunities to concentrate on academic tasks (Cheema and Kitsantas, 2014, 2016). And existing findings have revealed the moderating effect of disciplinary climate on the relationships between other variables and academic outcomes, such as mathematics performance (Sortkær and Reimer, 2016) and science achievement (Chi et al., 2018), but no similar exploration has been done in the reading context. In this sense, we also need to understand the nature of the interaction between school disciplinary climate and students' internal elements to influence reading literacy. As a result, the extant studies on the link between school belonging and reading literacy will consider the motivational mechanism and climatic boundaries.

Over the past few decades, school belonging has been the most frequently investigated in the United States and other Western countries (Kiefer et al., 2015; Froiland et al., 2016; Okilwa, 2016; Reynolds et al., 2017; Abdollahi and Noltemeyer, 2018; Korpershoek et al., 2019). However, the related research conducted in the Eastern context is still very limited. Considering the differences in cultural and administrative norms between Western and Eastern contexts (Ning et al., 2015), it thus remains unclear whether their theoretical and empirical findings are generalizable to Eastern cultures. For example, in China, cultivating high-performing, motivated, and well-behaved learners now and for life is considered as a combined duty of schooling, which is different from western school systems (Ning, 2020). Thus, we need to continue to study this issue deeper in China to confirm the generalizability of the findings to other cultural and geographic backgrounds. Therefore, this current study aims to provide a deeper understanding of the internal mechanism and conditional boundary underlying the association between student school belonging and reading literacy in a Chinese mainland sample of Programme for International Student Assessment (PISA) 2018.

THEORETICAL FOUNDATIONS AND HYPOTHESIS DEVELOPMENT

School Belonging and Reading Literacy

School belonging in educational settings is an important index to evaluate the learning and life quality of students in schools (Huebner, 2004). It has consistently been defined as "the extent to which students feel personally accepted, respected, included, and supported by others in the school social environment" (Goodenow, 1993). According to the self-determination theory (SDT), there are three basic psychological needs of individuals: autonomy, competence, and relationship (Ryan and Deci, 2002), in which relationship needs point out the individual's need for an interpersonal relationship. The sense of belonging is

a basic psychological need of individuals (Korpershoek et al., 2019). Consistent with this point, researchers have suggested that students with a higher sense of belonging to school are more likely to show positive psychological outcomes (Korpershoek et al., 2019) and these positive psychological outcomes are associated with academic success (Kiefer et al., 2015; Okilwa, 2016; Reynolds et al., 2017; Abdollahi and Noltemeyer, 2018; Korpershoek et al., 2019), including Chinese students' reading literacy (Huang, 2020).

Although school belonging is often found to be associated with academic achievement, there are some inconsistencies in these findings (Froiland et al., 2016; Dixon, 2020; Koyuncu and Firat, 2020), including Chinese students (Liu and Lu, 2011; Li et al., 2020). For instance, Liu and Lu (2011) found that whether in the initial state or the growth trajectories of the high school transition period, urban public school students' school belonging could non-significantly predict the changes in their academic achievement that was evaluated by the standardized composite scores of Chinese, math, and English. And another study in rural China also reported that school belonging did not predict junior school students' standardized scores of reading, math, and science (Li et al., 2020). As mentioned above, a possible reason for these inconsistent findings in the Chinese samples may be due to the disciplinary differences among reading, math, and science, which were overlooked by the use of combined scores of several subjects. Sample representativeness can also result in this inconsistency. To be specific, a sample was recruited from two urban public high schools located in eastern China (Liu and Lu, 2011) and another sample was represented a rural area of southwestern China (Li et al., 2020), which limit the generalizability of the findings to the whole education situation of the Chinese mainland. Therefore, more empirical research that used a more representative sample should be conducted to get a coherent understanding of the direct effect of school belonging on reading achievement in the Chinese context. Based on SDT and literature analysis, the following assumption was made:

Hypothesis 1: School belonging can significantly and positively predict reading literacy.

Mastery Goal Orientation

Achievement goal orientation theory is considered as one of the most crucial frameworks for understanding learning motivation (Elliot, 1999, 2005), and is often used to predict students' various academic outcomes and comprehend their academic difficulties (Patrick et al., 2011; Was et al., 2017; Diaconu-Gherasim et al., 2019; Zheng et al., 2020). It is generally believed that there are two different patterns: performance goal orientation and mastery goal orientation (Dweck, 1986). Students who demonstrate performance goal orientation believe that learning is to gain high social comparison and external rewards (Elliot, 1999, 2005), while students who tend to mastery goal orientation believe that learning is to acquire new knowledge or skills, gain academic recognition and achieve self-development (Patrick et al., 2011; Was et al., 2017). It can be seen from the above analysis that students with mastery goals generally perform better in academic assessment, and multiple empirical studies also support this point

(Shim et al., 2008; King and McInerney, 2016; Diaconu-Gherasim et al., 2019; Korpershoek et al., 2019; Theis et al., 2020).

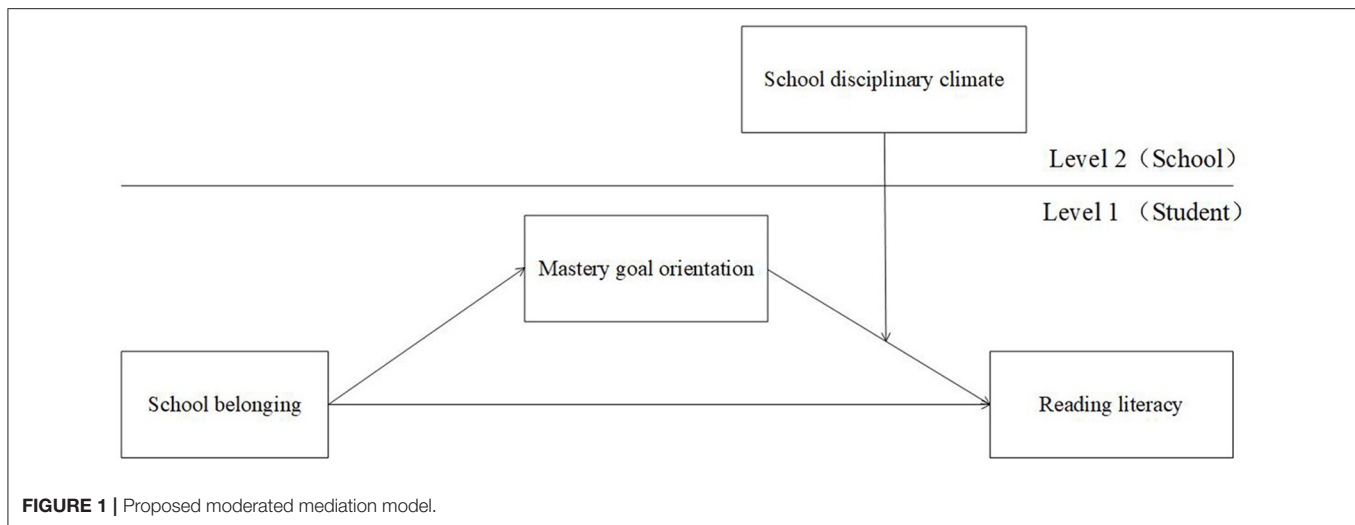
According to prior studies, students under high levels of school belonging tend to prefer mastery goals, rather than performance goals (Aerts et al., 2015; Korpershoek et al., 2019). And which is demonstrated by SDT, that is external environmental factors influence individual behavior and mental health (e.g., academic achievement) through the mediating effect of internal psychological needs (Jang et al., 2009). Based on SDT and the aforementioned research-derived relationships between school belonging, mastery goal orientation, and academic performance, we found ample evidence that indicated mastery goal orientation could represent a potential mediating role between school belonging and reading literacy. Therefore, the following hypothesis was proposed:

Hypothesis 2: Mastery goal orientation mediates the link between school belonging and reading literacy.

School Disciplinary Climate

Disciplinary climate is conceptualized as the perceptions that students hold on the consistency of classroom rules and how teachers address behavioral problems during class (Cheema and Kitsantas, 2014). And the school disciplinary climate focuses on school discipline and school order from the lens of school members' shared perceptions (Guo et al., 2018). Based on the organizational psychology perspective, climate subsumes two main aspects: climate level and climate strength (Schneider et al., 2013). Climate level refers to the quality of climate, which can be described as positive or negative; climate strength reflects the consensus or agreement on individual perceptions, which can be described as strong or weak (Schneider et al., 2002, 2013, 2017). Accordingly, a positive effect of climate level is usually expected when the climate strength is strong (Schneider et al., 2017).

Ecological systems theory also indicated that the school, as an important environment in the microsystem, has an increasingly strong influence on students (Bronfenbrenner, 1992). From the perspective of evidence, most school disciplinary climate studies have paid more attention to climate level and little focus on climate strength (Guo et al., 2018). These studies mainly supported the positive association between disciplinary climate level and academic behavior (Arum and Velez, 2012; Frempong et al., 2012; Cheema and Kitsantas, 2014; Ning et al., 2015; Jenkins and Ueno, 2017; Chi et al., 2018; Guo et al., 2018; Ning, 2019, 2020), including reading achievement (Ning et al., 2015). Meanwhile, Guo et al. (2018) highlighted the positive effect of disciplinary climate strength on student reading performance. However, to the best knowledge of authors, few studies have identified the impact of school disciplinary climate on students' academic achievement from the lens of both climate level and climate strength. In addition, the moderating effect of disciplinary climate has been found in the relationship between other variables and learning outcomes (Sortkær and Reimer, 2016; Chi et al., 2018). For instance, Sortkær and Reimer (2016) found that the interaction between gender and disciplinary climate significantly influences student mathematics performance, and Chi et al. (2018) showed that disciplinary



climate could moderate the association between inquiry-based science activities and student science achievement for both genders. Based on the above findings of other disciplines, the following hypothesis was raised in the reading context:

Hypothesis 3: School disciplinary climate level moderates the indirect effect of school belonging and reading literacy through mastery goal orientation.

Hypothesis 4: School disciplinary climate strength moderates the indirect effect of school belonging and reading literacy through mastery goal orientation.

To sum up, this study constructed a cross-level moderated mediation model based on ecological systems theory and SDT (see **Figure 1**). The research aimed to explore three main research questions: (1) whether school belonging can directly predict student reading literacy, (2) whether mastery goal orientation has a mediating effect on the relationship between school belonging and reading literacy, and (3) whether the school disciplinary climate has a moderating effect on this mediating model. Testing the above questions would help to clarify the motivational mechanism (i.e., mastery goal orientation) and conditional boundaries (i.e., school disciplinary climate level and strength) underlying the association between school belonging and reading literacy for Chinese mainland students, and provide theoretical basis and empirical support for improving middle school students' reading literacy.

MATERIALS AND METHODS

Participants

PISA is a triennial international survey designed to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students, together with their background and non-cognitive characteristics and school context. PISA 2018 is the seventh test of the project and also provides multiple reference standards for student evaluation by collecting relevant variables of students' families, teachers, and schools. In the two-stage

unequal probability sampling design adopted by PISA, a total of 362 schools were selected from Beijing, Shanghai, Jiangsu, and Zhejiang provinces in China. At each sampled school, as many as 35 students were randomly selected. After preliminary data cleaning and weighting by PISA, 12058 (5,775 girls and 6,283 boys) students that represented 992,302 15-year-old students in mainland China were remained. In addition, considering the requirements of sufficient statistical power (Kreft and de Leeuw, 1998; Kim and Hong, 2020), 30 schools with <30 students were excluded in this study so that each school has a maximum of 35 students and a minimum of 30 students after data cleaning. And the missing data on student-level measures was also deleted for HLM analysis (Enders, 2010), resulting in a total of 11,364 students from 332 schools (5,455 girls and 5,909 boys) were remained.

Measures

Reading Literacy

According to PISA 2018 reading framework, "reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society (OECD, 2019)." Based on this definition, reading literacy assessment in PISA 2018 was developed by examining three cognitive processes, such as locating information, understanding, evaluating, and reflecting in both simple and multiple texts. Both multiple-choice and open constructed-response items were used. To be clear, participants' reading literacy was assessed in the official language of their home country (here: Chinese). Ten plausible values of reading literacy were displayed in the results of PISA 2018, in which the results were standardized with an average value of 500 and a standard deviation of 100 for OECD countries. In this study, we showed the pooled results that used 10 plausible values for reading literacy to analyze (OECD, 2009).

School Belonging

School belonging was represented by the level of acceptance, respect, and support those students feel in their schools (OECD,

2019), the BELONG index in the PISA 2018 student database. This is a standardized index with an average value of zero and a standard deviation of one for the OECD countries. Technically, the BELONG index was calculated by averaging students' levels of agreement with six statements regarding their feelings in their schools, such as "I make friends easily at school." The response scale ranges from 1 to 4, indicating from "strongly agree" to "strongly disagree," respectively. In addition, items 1, 4, and 6 were entitled reverse questions. And these items were reversed so that a higher score corresponds to a higher level of school belonging. The Cronbach's raw α for BELONG in China is 0.84 and the McDonald's ω_h for MASTGOAL in China is 0.84, which indicates an acceptable level of internal consistency.

Mastery Goal Orientation

Mastery goal orientation was represented by the tendency to mastery goals that students show in their study and life (OECD, 2019), the MASTGOAL index in the PISA 2018 student database. This is a standardized index with an average value of zero and a standard deviation of one for the OECD countries. Technically, the MASTGOAL index was calculated by averaging students' levels of agreement with three statements regarding their tendency to mastery goals, such as "My goal is to learn as much as possible." The response scale ranges from 1 to 5, indicating from "not at all true of me" to "extremely true of me," respectively. All items were positive so that a higher score corresponds to a higher level of mastery goal orientation. The Cronbach's raw α for MASTGOAL in China is 0.72 and the McDonald's ω_h for MASTGOAL in China is 0.76, which indicates an acceptable level of internal consistency.

School Disciplinary Climate

Disciplinary climate is measured by the extent to which students miss learning opportunities due to disruptive behavior in the reading classroom (OECD, 2019). In this study, school disciplinary climate represented the average level of classroom disciplinary climate in a school, with a higher score indicating an orderly climate. The DISCLIMA index was calculated by averaging students' levels of agreement with five statements in their reading lessons, such as "Students don't listen to what the teacher says." The response scale ranges from 1 to 4, indicating from "every lesson" to "never or hardly ever," respectively. All items were positive so that higher scores correspond to a better disciplinary climate. The Cronbach's raw α for DISCLIMA in China is 0.89 and the McDonald's ω_h for DISCLIMA in China is 0.90, which indicates an acceptable level of internal consistency.

In addition to the aforementioned scales, the economic, social, and cultural status index (ESCS) and GENDER which were listed in PISA 2018 database, were used in the analyses. ESCS is a standardized index with an average value of zero and a standard deviation of one for the OECD countries. Three indices were used in the construction of the ESCS index, including the highest occupational status of parents (HISEI), the highest educational level of parents in years of education (PARED), and home possession (HOMEPOS). GENDER was coded as a dummy variable with a value of 1 for girls ($n = 5,455$) and 0 for boys ($n = 5,909$).

Analytical Strategy

Considering the nested feature of the data (i.e., students in the same school shared the same school disciplinary climate), multilevel modeling (MLM) was employed to analyze the data of the present research. The hypothetical model was tested with the MLmed Beta 2 macro for SPSS software (Hayes and Rockwood, 2020). Utilizing this analytic approach, a multilevel moderated mediation model was proposed. It estimated (see **Figure 1**) school disciplinary climate as a level-2 moderator of the level-1 indirect effect of school belonging on reading literacy via mastery goal orientation.

Before the formal analysis, the intra-class correlation coefficient (ICC) values of the moderator (i.e., school disciplinary climate) were calculated to estimate the dependence magnitude (Cohen, 1988). The ICC agreements were calculated to examine whether disciplinary climate could significantly explain the variance in individual responses [ICC(1)] and to assess the reliability of school-level means [ICC(2), Bliese, 2000]. R_{wg} was also used to test the polymerization criteria for school disciplinary climate (James et al., 1993). The results showed that $ICC(1) = 0.13 (>0.12)$, $ICC(2) = 0.84 (>0.47)$, and $R_{wg} = 0.88 (>0.7)$, which indicates that a decent proportion of the total variation found in all student ratings can be attributed to the fact that students are nested within schools.

Consistent with previous research, the school mean of students' responses on disciplinary climate is used as an index for disciplinary climate level (Chan, 1998), and within-school standard deviation of disciplinary climate scores is used to indicate climate strength (Roberson et al., 2007). Given that the standard deviation is an index of variability, it is multiplied by -1 , so that a higher score represents a higher climate strength (Guo et al., 2018).

RESULTS

Confirmatory Factor Analysis

We use 14 response items measuring the three constructs (six items of school belonging, three items of mastery goal orientation, and five items of school disciplinary climate) to test the CFA in Mplus 8.0 to merge the variables gradually and examine changes in fitting degrees to test the discriminant validity of the model. The fitting indices of the one-factor model were not good: $\chi^2/df = 434.03$, $CFI = 0.51$, $TLI = 0.42$, $RMSEA = 0.20$ and $SRMR = 0.17$. Secondly, we combined school belonging and mastery goal orientation, and the fitting indices of the two-factor model were still not satisfactory: $\chi^2/df = 172.84$, $CFI = 0.81$, $TLI = 0.77$, $RMSEA = 0.12$ and $SRMR = 0.08$. At last, the result of three-factor model indicated a good fit to the data: $\chi^2/df = 8.31$, $CFI = 0.99$, $TLI = 0.99$, $RMSEA = 0.03$ and $SRMR = 0.02$. And the loading of each item in the three-factor model were shown in **Table 1**.

Then, Harman's One-Factor Test was conducted to examine the common method biases. And the result reported that the first factor accounted for 31.54% ($<40\%$) of the total variance, which indicated that the influence of the homologous coefficient of variance was not serious.

Descriptive Statistical Analysis

In this study, both disciplinary climate level and disciplinary climate strength were at the school level (i.e., level-2), whereas school belonging, mastery goal orientation, and reading literacy were at the student level (i.e., level-1). The control variables (i.e., gender and ESCS) were included in the model as student-level variables. Descriptive statistics and correlations between the variables are shown in **Table 2**. The significant coefficients of pairwise correlation range between -0.03 and 0.54 .

Before testing the hypotheses, we examined the variations in reading literacy across the level. The result shows that level-1 variance was 53.64% of the total variance, and level-2 variance was 46.36% of the total variance. These results indicated a significant variance at level-2 for reading literacy (James, 1982). Thus, using a multilevel model was appropriate.

Hypothesis Testing

MLM was employed to verify the hypothesis of the present research. There are two models to be tested because there are two moderators (i.e., school disciplinary climate level and school

disciplinary climate strength). The results of the moderated mediation analysis are reported in **Table 3** and **Figures 2, 3**. As seen in **Figures 2, 3**, no matter which moderator was considered, the direct effect of school belonging on reading literacy at the student level was always significant and positive [effect = 1.690, 95% CI (0.397, 2.983); effect = 1.630, 95% CI (0.336, 2.923)], hence hypotheses 1 was supported.

Hypothesis 2 tested the mediating effect of mastery goal orientation on the link between school belonging and reading literacy. As shown in **Table 3**, no matter which moderator was considered, the indirect effect of school belonging on reading literacy via mastery goal orientation at the student level was always significant and positive [effect = 1.008, 95% CI (0.694, 1.320); effect = 0.999, 95% CI (0.684, 1.311)], so that hypotheses 2 was verified.

Hypothesis 3 and 4 tested the moderated roles of school disciplinary climate level and strength on the indirect effect of school belonging and reading literacy through mastery goal orientation. **Figure 2** shows that a significant and negative interaction effect was found in school disciplinary climate level moderated the link between mastery goal orientation and reading literacy [$b = -1.630$, 95% CI (-2.759 , -0.410)]. Similarly, **Figure 3** also shows that a significant and negative interaction effect was observed in which school disciplinary climate strength moderated the link between these two variables [$b = -1.705$, 95% CI (-2.945 , -0.464)]. The results in **Table 3** show that both indexes of moderated mediation were significant and negative [effect = -0.374 , 95% CI (-0.660 , -0.096); effect = -0.392 , 95% CI (-0.682 , -0.108)]. The conditional indirect effect of school belonging on reading literacy at a low level of school disciplinary climate level [effect = 1.382, 95% CI (0.972, 1.807)] was higher than the effect at a high level of school disciplinary

TABLE 1 | The loading of each item.

Scale	Items' number & loading					
Mastery goal orientation	1	2	3			
	0.51	0.82	0.83			
School disciplinary climate	1	2	3	4	5	
	0.72	0.71	0.72	0.88	0.86	
School belonging	1	2	3	4	5	6
	0.79	0.57	0.86	0.69	0.54	0.79

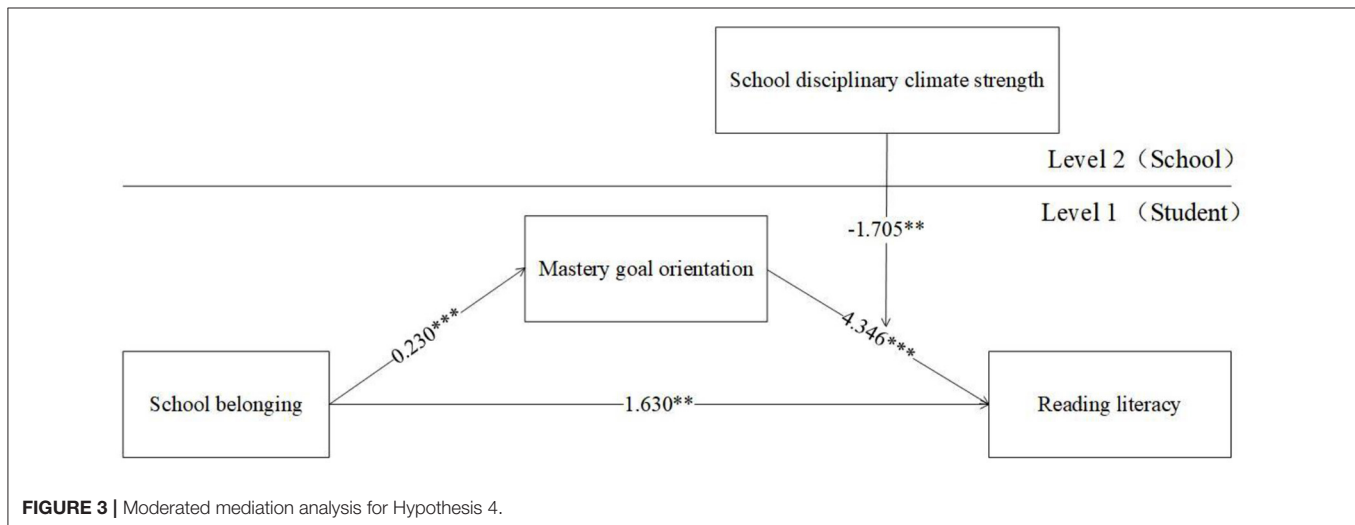
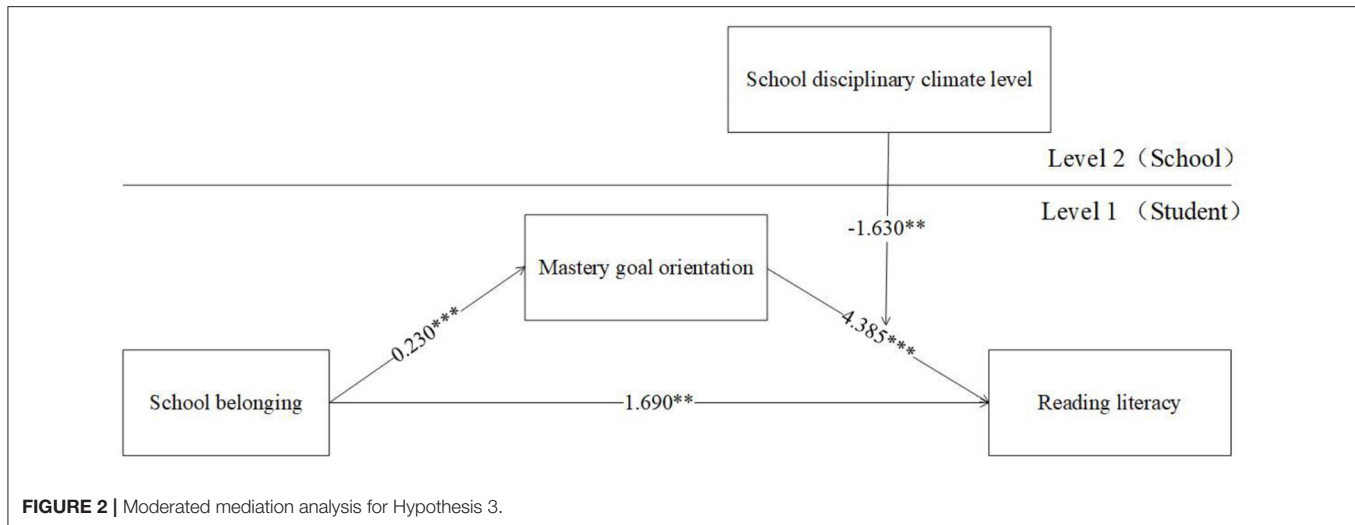
TABLE 2 | Descriptive statistics and correlation coefficient ($n = 11,364$).

	M	SD	1	2	3	4	5	6
1 Gender	0.48	0.50	-					
2 ESCS	-0.34	1.07	0.02*	-				
3 School belonging	-0.15	0.91	-0.03**	0.14***	-			
4 Mastery goal orientation	0.06	0.91	0.02**	0.16***	0.26***	-		
5 School disciplinary climate level	0.82	0.42	0.05***	0.29***	0.14***	0.14***	-	
6 School disciplinary climate strength	-0.94	0.17	0.04***	0.10***	0.03***	0.04***	0.54***	-
7 Reading literacy	564.58	88.26	0.08***	0.36***	0.07***	0.14***	0.38***	0.27***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 3 | Indirect effect and moderated mediation effect.

	School disciplinary climate level				School disciplinary climate strength			
	Effect	SE	LL	UL	Effect	SE	LL	UL
Indirect effect	1.008	0.157	0.694	1.320	0.999	0.157	0.684	1.311
Index of moderated mediation	-0.374		-0.660	-0.096	-0.392		-0.682	-0.108
Low (-1SD)	1.382	0.209	0.972	1.807	1.390	0.209	0.980	1.877
High (+1SD)	0.633	0.216	0.202	1.057	0.607	0.220	0.168	1.038



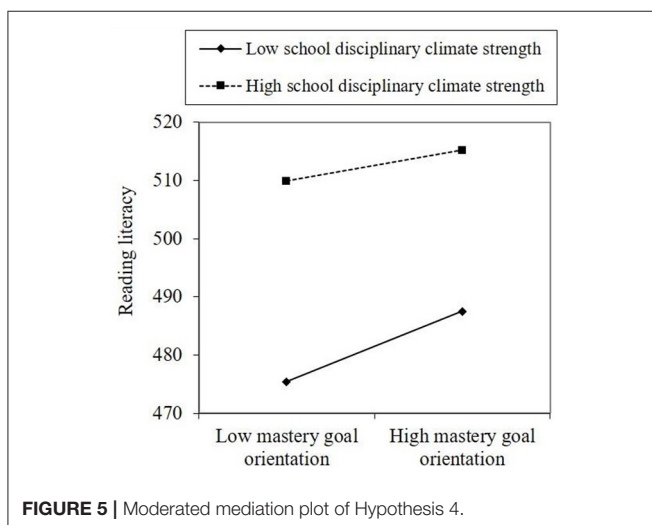
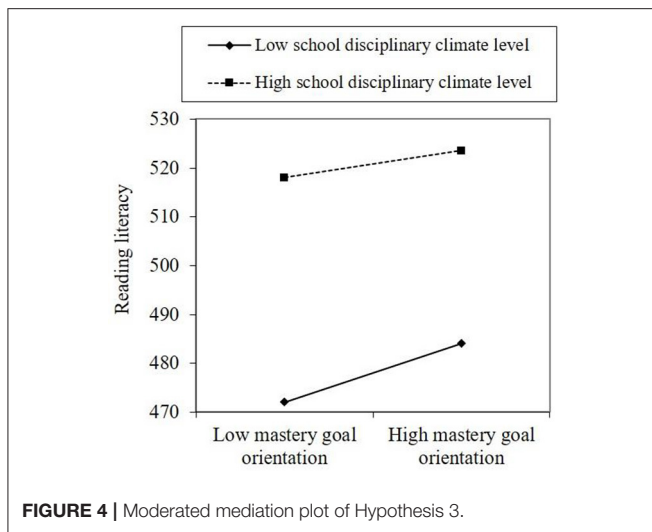
climate level [effect = 0.633, 95% CI (0.202, 1.057)], Hypotheses 3 was proved (see **Table 3** and **Figure 4**). Similarly, the conditional indirect effect of school belonging on reading literacy at a low level of school disciplinary climate strength [effect = 1.390, 95% CI (0.980, 1.877)] was higher than the effect at a high level of school disciplinary climate strength [effect = 0.607, 95% CI (0.168, 1.038)], Hypotheses 4 was proved (see **Table 3** and **Figure 5**).

DISCUSSION

How to improve students' academic achievement is an essential consideration for all school educators and policymakers, as well as students' well-being and non-cognitive development. Based on ecological systems theory and SDT, the present research provides valuable insights into how school belonging affects Chinese mainland students' reading literacy and also gives consideration to the impact of mastery goal orientation as a

critical mediator. Furthermore, we also reveal the interplay between school disciplinary climate (i.e., climate level and climate strength) and mastery goal orientation influencing students' reading literacy by performing a multilevel moderated mediation analysis. And this present study contributes to the existing literature as outlined below.

In the first place, the positive direct effect of school belonging on reading literacy was clarified and confirmed in the Chinese context. This finding was in line with theoretical expectations, most previous works that were conducted in Western counties (Kiefer et al., 2015; Okilwa, 2016; Reynolds et al., 2017; Abdollahi and Noltemeyer, 2018; Korpershoek et al., 2019) and one study that especially centered on Chinese students' reading literacy in PISA (Huang, 2020), but was inconsistent with some results that focused on a small sample of Chinese students, whose academic achievement represented a combination of three subjects (Liu and Lu, 2011; Li et al., 2020). Draw on the above analysis, both discipline differences (Hinnant et al., 2009; Reynolds et al., 2017) and sample representativeness could lead to the



inconsistency in these studies. In consistent with SDT (Ryan and Deci, 2002), one explanation for the positive relationship may lie in the higher level of school belonging provided for students by respect and care from teachers, peers, and schools, which meet students' basic psychological needs, heighten their motivation (Neel and Fuligni, 2013), self-efficacy (Dixson, 2020), and other internal psychological factors (Korpershoek et al., 2019), and even hinder the influence of negative factors such as academic pressure and anxiety (Abdollahi et al., 2020), as a result, it can improve academic outcomes. Once again, the result of the current study affirms that valid and abundant emotional interactions with classmates and teachers can not only enhance adolescents' reading comprehension and performance, but also promote their cognitive development and academic prospects.

Regarding the second hypothesis, we empirically support the idea that mastery goal orientation partially mediates the positive relationship between school belonging and reading

literacy, serving as a more powerful explanatory mechanism. In consistent with SDT, school belonging has the potential to encourage students and to transform their emotional energy into better mastery goals and then advanced reading literacy. It seems reasonable to assume that positive relationships with classmates and teachers, which imply a high sense of belonging to school, could increase more positive emotions such as hope, happiness, and pride and less negative emotions such as boredom and anger (Pekrun et al., 2006, 2009), and strengthen their perception of mastery goals (Patrick et al., 2011). In addition, a plethora of studies noted that mastery goal orientation can effectively facilitate students' metacognitive strategies (King and McInerney, 2016), working memory (Wolters, 2004) and classroom participation (Datu and Park, 2019; Yi et al., 2020), self-management (Skinner et al., 2009) and other positive variables associated with academic achievement, can also resist negative factors such as academic pressure and anxiety (Zheng et al., 2020), so that can improve academic achievement (Wolters, 2004; King and McInerney, 2016; Diaconu-Gherasim et al., 2019; Theis et al., 2020). Therefore, this result may explain the nonsignificant direct effect of school belonging on Chinese students' academic achievement (Liu and Lu, 2011; Li et al., 2020). It is not that the sense of belonging to school does not directly predict academic development, but indirectly through other variables such as mastery goal orientation. In summary, this finding contributes to the literature by verifying a theoretical model of the mechanism between individual emotional characteristics and motivational factors in achieving academic success.

At last, perhaps the most important and novel finding of this study was that both school disciplinary climate level and climate strength could moderate the link between mastery goal orientation and reading literacy, and the indexes of moderated mediation were all significant and negative. This result was consistent with the neural findings, for example, Braver et al. (2014) have reported that an orderly school disciplinary climate is extremely crucial in the motivation-achievement relationship for adolescents, considering that most adolescents' prefrontal cortex, which is in charge of self-control, has not fully matured. And in line with the ecological systems theory, this study admitted the importance of the interaction between the learning environment (i.e., school disciplinary climate) and student internal characteristics (i.e., mastery goal orientation) for academic promotion. Conceptually, in a strong climate, explicit expectations can lead people to act similarly. While people behave differently in a weak climate, the organizational goals and values are too obscure for them to form consistent perceptions (Schneider et al., 2002, 2013, 2017). Especially, only when students' perception of school disciplinary climate is both positive and consistent, can the school disciplinary climate be acceptable for school members (Guo et al., 2018). In other words, there is a significant correlation between the level and strength of school disciplinary climate. Therefore, these two similar moderated mediation indexes are understandable and logical.

To be specific, the conditional indirect effect of school belonging on reading literacy at a low level of school disciplinary

climate was greater than the effect at a high level of school disciplinary climate. Further analysis revealed that the reason is that the impact of the mastery goal orientation on student reading literacy can be weakened by both a more positive school disciplinary climate level and strong climate strength, but be enhanced by both a more negative school disciplinary climate level and weak climate strength. On the one hand, it is possible that a highly ordered school disciplinary climate would undermine student sense of self-determination, intrinsic motivation (Kover and Worrell, 2010), and autonomy (Ning, 2020), especially in East Asian cultures, where an orderly learning atmosphere and explicit rules are established to hinder distraction, noise, and disorder in the classroom (Ning, 2019, 2020), which could easily lead to the weakening of students' tendency to mastery goals. On the other hand, although the negative and weak school disciplinary climate always means loose school management, students under higher self-management ability and autonomy are more likely to tend to mastery goals and then take the initiative to utilize abundant resources to make academic progress. This may explain why both higher school disciplinary climate level and strength would weaken the relationship between mastery goal orientation and reading literacy, while a strong link between these two variables was observed at both low disciplinary climate level and strength. On the basis of this finding, the importance of interaction between individual motivational factors and external environment is highlighted in achieving academic success in the Chinese context. However, research in this area is limited and further works are urgently needed to investigate the potential boundary conditions of the effect of school belonging on learning outcomes to understand the changes in reading behavior.

And these results enrich the research on the relationship between school belonging and academic achievement and provide a practical basis for policymakers, school administrators, and teachers to improve the reading literacy of students and entire educational quality. In the first place, the empirical results support the valuable idea that school belonging is significant for students' development. Therefore, it is urgent to take effective measures to help students who may have such problems develop a sense of belonging to school. Secondly, students with mastery goal orientation are more likely to channel positive and powerful emotional energy into academic achievement. Hence, the educational significance of this finding is emphasized by the fact that mastery goal orientation is a factor that is relatively more manipulable compared to some other vital predictors of academic success such as socioeconomic status. However, mastery goal orientation seems to pose a differential impact on reading literacy depending on the different levels of disciplinary climate level and strength. The pursuit of an orderly disciplinary climate would regulate the bad behavior of students and provide a powerful safeguard for teaching and learning, but may impair the initiative and autonomy of students with mastery goals, and even weaken the link between positive psychological factors (i.e., mastery goal orientation) and academic achievement. Certainly, we need more relevant research to assist policymakers, school administrators, and teachers to address the issue of creating a more temperate disciplinary environment.

Limitations and Implications for Future Research

Three limitations of the present research should be noted to provide directions for future research. Firstly, we used the data from PISA 2018 that was published by OECD. These results represent the education situation of 15-year-old students from Beijing, Shanghai, Jiangsu, and Zhejiang provinces in China, which does not necessarily represent the education status of the whole country. In addition, because the definition and optimal level and strength of a school discipline climate may vary from culture to culture, the successful findings and practices on school disciplinary climate found in one country may not be suitable for the schools in other countries (Guo et al., 2018). Therefore, we need more effort in other subjects and cultures. Secondly, in the cross-sectional design of PISA 2018, it is limited to make directional assumptions only from the data in this study. Thus, we cannot fully reveal the causal relationship between these variables, so that we urge future research to replicate our findings using multiple sources and multiple time points in data collection. Finally, PISA examined a wide range of variables in their student self-reported questionnaire, which is prone to common method deviation. Although we have proved that the possible common method deviation is not serious and the tools have higher reliability and validity according to the statistical results, such situations should still be avoided in future studies. Future research could overcome this limitation by considering objective methods (e.g., observations and interviews) to evaluate the relationships between variables.

CONCLUSION

According to ecological systems theory and SDT, from both motivational and environmental point of view, the present research establishes a cross-level model to explain how and when school belonging impact students' reading literacy in the Chinese context. After performing the cross-sectional analysis, the results indicated that: (1) school belonging had a direct and positive effect on student reading literacy; (2) the relationship between school belonging and reading literacy was prominently mediated by mastery goal orientation; (3) both school disciplinary climate level and strength could negatively moderate the latter half path of "school belonging → mastery goal orientation → reading literacy." Notwithstanding the preliminary nature of our findings, we have some confidence in these conclusions as they are based on a sample from a large-scale survey in the Chinese context, affording a high degree of external validity. Nevertheless, there is still much work that remains to be done in this emerging line of research and future work needs to consolidate and extend our findings.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <https://www.oecd.org/pisa/data/2018database>.

AUTHOR CONTRIBUTIONS

YT and TY designed this study. YT analyzed the data and wrote this article. ZF, XW, and TY reviewed the study and performed substantial suggestions. All authors contributed to the article and approved the submitted version.

REFERENCES

- Abdollahi, A., and Noltemeyer, A. (2018). Academic hardness: Mediator between sense of belonging to school and academic achievement? *J. Educ. Res.* 111, 345–351. doi: 10.1080/00220671.2016.1261075
- Abdollahi, A., Panahipour, S., Tafti, M. A., and Allen, K. A. (2020). Academic hardness as a mediator for the relationship between school belonging and academic stress. *Psychol. Sch.* 57, 823–832. doi: 10.1002/pits.22339
- Aerts, S., Houtte, M. V., Dewaele, A., Cox, N., and Vincke, J. (2015). School motivation in secondary schools: A survey of LGB and heterosexual students in Flanders. *Youth Soc.* 47, 412–437. doi: 10.1177/0044118X12467657
- Arum, R., and Velez, M. (2012). *Improving Learning Environments: School Discipline and Student Achievement in Comparative Perspective* (1st ed.). Stanford, CA: Stanford University Press. doi: 10.11126/stanford/9780804778039.003.0001
- Bliese, P. D. (2000). “Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis,” in *Multilevel Theory, Research, and Methods in Organizations*, eds K. J. Klein and S. W. J. Kozlowski (San Francisco, CA: Jossey-Bass).
- Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., et al. (2014). Mechanisms of motivation-cognition interaction: Challenges and opportunities. *Cogn. Affect. Behav. Neurosci.* 14, 443–472. doi: 10.3758/s13415-014-0300-0
- Bronfenbrenner, U. (1992). “Ecological systems theory,” in *Six Theories of Child Development: Revised Formulations and Current Issues*, ed R. Vasta (London: Jessica Kingsley Publishers).
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *J. Appl. Psychol.* 83, 234–246. doi: 10.1037/0021-9010.83.2.234
- Cheema, J. R., and Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: A look at gender and racial-ethnic differences. *Int. J. Sci. Mathematics Educ.* 12, 1261–1279. doi: 10.1007/s10763-013-9454-4
- Cheema, J. R., and Kitsantas, A. (2016). Predicting high school student use of learning strategies: The role of preferred learning styles and classroom climate. *Educ. Psychol.* 36, 845–862. doi: 10.1080/01443410.2014.981511
- Chi, S., Liu, X., Wang, Z., and Han, S. W. (2018). Moderation of the effects of scientific inquiry activities on low SES students' PISA 2015 science achievement by school teacher support and disciplinary climate in science classroom across gender. *Int. J. Sci. Educ.* 40, 1284–1304. doi: 10.1080/09500693.2018.1476742
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York, NY: Routledge.
- Datu, J. A. D., and Park, N. (2019). Perceived school kindness and academic engagement: The mediational roles of achievement goal orientations. *School Psychol. Int.* 40, 456–473. doi: 10.1177/0143034319854474
- Diaconu-Gherasim, L. R., Tepordei, A.-M., Mairean, C., and Rusu, A. (2019). Intelligence beliefs, goal orientations and children's academic achievement: Does the children's gender matter? *Educ. Stud.* 45, 95–112. doi: 10.1080/03055698.2018.1443796
- Dixon, D. D. (2020). How hope measures up: Hope predicts school variables beyond growth mindset and school belonging. *Curr. Psychol.* doi: 10.1007/s12144-020-00975-y
- Dweek, C. S. (1986). Motivational processes affecting learning. *Am. Psychol.* 41, 1040–1048. doi: 10.1037/0003-066X.41.10.1040
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educ. Psychol.* 34, 169–189. doi: 10.1207/s15326985ep3403_3
- Elliot, A. J. (2005). “A conceptual history of the achievement goal construct,” in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweek (New York, NY: Guilford Publications).
- Enders, C. K. (2010). *Applied Missing Data Analysis (Illustrated edition)*. New York, NY: The Guilford Press.
- Frempong, G., Ma, X., and Mensah, J. (2012). Access to postsecondary education: can schools compensate for socioeconomic disadvantage? *Higher Educ.* 63, 19–32. doi: 10.1007/s10734-011-9422-2
- Froiland, J. M., Davison, M. L., and Worrell, F. C. (2016). Aloha teachers: Teacher autonomy support promotes Native Hawaiian and Pacific Islander students' motivation, school belonging, course-taking and math achievement. *Soc. Psychol. Educ.* 19, 879–894. doi: 10.1007/s11218-016-9355-9
- Gini, G., Marino, C., Pozzoli, T., and Holt, M. (2018). Associations between peer victimization, perceived teacher unfairness, and adolescents' adjustment and well-being. *J. Sch. Psychol.* 67, 56–68. doi: 10.1016/j.jsp.2017.09.005
- Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychol. Sch.* 30, 79–90. doi: 10.1002/1222520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X
- Guo, S., Li, L., and Zhang, D. (2018). A multilevel analysis of the effects of disciplinary climate strength on student reading performance. *Asia Pacific Educ. Rev.* 19, 1–15. doi: 10.1007/s12564-018-9516-y
- Hamed, S. M., Pishghadam, R., and Fadardi, J. S. (2020). The contribution of reading emotions to reading comprehension: The mediating effect of reading engagement using a structural equation modeling approach. *Educ. Res. Policy Pract.* 19, 211–238. doi: 10.1007/s10671-019-09256-3
- Hayes, A. F., and Rockwood, N. J. (2020). Conditional process analysis: Concepts, computation, and advances in the modeling of the contingencies of mechanisms. *Am. Behav. Sci.* 64, 19–54. doi: 10.1177/0002764219859633
- Hinnant, J. B., O'Brien, M., and Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *J. Educ. Psychol.* 101, 662–670. doi: 10.1037/a0014306
- Ho, E. S. C. (2005). Effect of school decentralization and school climate on student mathematics performance: The case of Hong Kong. *Educ. Res. Policy Practice* 4, 47–64. doi: 10.1007/s10671-005-1546-7
- Huang, L. (2020). Exploring the relationship between school bullying and academic performance: The mediating role of students' sense of belonging at school. *Educ. Stud.* doi: 10.1080/03055698.2020.1749032
- Huebner, E. S. (2004). Research on assessment of life satisfaction of children and adolescents. *Soc. Indic. Res.* 66, 3–33. doi: 10.1023/B:SOCI.0000007497.57754.e3
- James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *J. Appl. Psychol.* 67, 219–229. doi: 10.1037/0021-9010.67.2.219
- James, L. R., Demaree, R. G., and Wolf, G. (1993). rwg: An assessment of within-group interrater agreement. *J. Appl. Psychol.* 78, 306–309. doi: 10.1037/0021-9010.78.2.306
- Jang, H., Reeve, J., Ryan, R. M., and Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *J. Educ. Psychol.* 101, 644–661. doi: 10.1037/a0014241
- Jenkins, A., and Ueno, A. (2017). Classroom disciplinary climate in secondary schools in England: What is the real picture? *Br. Educ. Res. J.* 43, 124–150. doi: 10.1002/berj.3255
- Kiefer, S. M., Alley, K. M., and Ellerbrock, C. R. (2015). Teacher and peer support for young adolescents' motivation, engagement, and school belonging. *RMLE Online* 38, 1–18. doi: 10.1080/19404476.2015.11641184
- Kim, D. H., and Kim, J. H. (2013). Social relations and school life satisfaction in South Korea. *Soc. Indic. Res.* 112, 105–127. doi: 10.1007/s11205-012-0042-8
- Kim, S., and Hong, S. (2020). Comparing methods for multilevel moderated mediation: a decomposed-first strategy. *Struct. Eq. Model. Multidisciplinary J.* 25, 661–677. doi: 10.1080/10705511.2019.1683015

FUNDING

This study was supported by the Independent project of Collaborative Innovation Center of Assessment toward Basic Education Quality (China) (BJZK-2020A1-20007).

- King, R. B., and McInerney, D. M. (2016). Do goals lead to outcomes or can it be the other way around?: causal ordering of mastery goals, metacognitive strategies, and achievement. *Br. J. Educ. Psychol.* 86, 296–312. doi: 10.1111/bjep.12107
- Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., and Boer, H., de. (2019). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Res. Papers Educ.* 35, 641–680. doi: 10.1080/02671522.2019.1615116
- Kover, D. J., and Worrell, F. C. (2010). The influence of instrumentality beliefs on intrinsic motivation: a study of high-achieving adolescents. *J. Adv. Acad.* 21, 470–498. doi: 10.1177/1932202X1002100305
- Koyuncu, I., and Firat, T. (2020). Investigating reading literacy in PISA 2018 assessment. *Int. Elect. J. Element. Educ.* 13, 263–275. doi: 10.26822/iejee.2021.189
- Kreft, I. G., and de Leeuw, J. (1998). *Introducing Multilevel Modeling*. London: Sage. doi: 10.4135/9781849209366
- Li, L., Chen, X., and Li, H. (2020). Bullying victimization, school belonging, academic engagement and achievement in adolescents in rural China: A serial mediation model. *Child. Youth Serv. Rev.* 113:104946. doi: 10.1016/j.childyouth.2020.104946
- Liu, Y., and Lu, Z. (2011). Trajectories of Chinese students' sense of school belonging and academic achievement over the high school transition period. *Learn. Individ. Differ.* 21, 187–190. doi: 10.1016/j.lindif.2010.12.007
- Ma, L., Luo, H., and Xiao, L. (2021). Perceived teacher support, self-concept, enjoyment and achievement in reading: A multilevel mediation model based on PISA 2018. *Learn. Individ. Differ.* 85:101947. doi: 10.1016/j.lindif.2020.101947
- Neel, C. G.-O., and Fuligni, A. (2013). A longitudinal study of school belonging and academic motivation across high school. *Child Dev.* 84, 678–692. doi: 10.1111/j.1467-8624.2012.01862.x
- Ning, B. (2019). Examining the importance of discipline in Chinese schooling: An exploration in Shanghai, Hong Kong, Macao, and Taipei. *Asia Pacific Educ. Rev.* 20, 489–501. doi: 10.1007/s12564-018-9563-4
- Ning, B. (2020). Discipline, motivation, and achievement in mathematics learning: An exploration in Shanghai. *Sch. Psychol. Int.* 41, 595–611. doi: 10.1177/0143034320961465
- Ning, B., Damme, J. V., Noortgate, W. V. D., Yang, X., and Gielen, S. (2015). The influence of classroom disciplinary climate of schools on reading achievement: A cross-country comparative study. *School Effect. School Improve.* 26, 586–611. doi: 10.1080/09243453.2015.1025796
- OECD (2009). *PISA Data Analysis Manual: SPSS (2nd ed)*. OECD Publishing.
- OECD (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing. doi: 10.1787/b25efab8-en
- Okilwa, N. S. A. (2016). Exploring school- and home-related protective factors for economically disadvantaged middle school students. *J. At-Risk Issues* 19, 34–46.
- Patrick, H., Kaplan, A., and Ryan, A. (2011). Positive classroom motivational environments: Convergence between mastery goal structure and classroom social climate. *J. Educ. Psychol.* 103, 367–382. doi: 10.1037/a0023311
- Pekrun, R., Elliot, A. J., and Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *J. Educ. Psychol.* 98, 583–597. doi: 10.1037/0022-0663.98.3.583
- Pekrun, R., Elliot, A. J., and Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *J. Educ. Psychol.* 101, 115–135. doi: 10.1037/a0013383
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychol. Bull.* 135, 322–338. doi: 10.1037/a0014996
- Reynolds, K. J., Lee, E., Turner, I., Bromhead, D., and Subasic, E. (2017). How does school climate impact academic achievement? An examination of social identity processes. *School Psychol. Int.* 38, 78–97. doi: 10.1177/0143034316682295
- Roberson, Q. M., Sturman, M. C., and Simons, T. L. (2007). Does the measure of dispersion matter in multilevel research? A Comparison of the relative performance of dispersion indexes. *Organizational Res. Methods* 10, 564–588. doi: 10.1177/1094428106294746
- Rogiers, A., Van Keer, H., and Merchie, E. (2020). The profile of the skilled reader: An investigation into the role of reading enjoyment and student characteristics. *Int. J. Educ. Res.* 99:101512. doi: 10.1016/j.ijer.2019.101512
- Ryan, R. M., and Deci, E. L. (2002). "An overview of self-determination theory: An organismic-dialectical perspective," in *Handbook of Self-Determination Research*, eds E. L. Deci and R. M. Ryan (The University of Rochester Press).
- Schneider, B., Ehrhart, M. G., and Macey, W. H. (2013). Organizational climate and culture. *Annu. Rev. Psychol.* 64, 361–388. doi: 10.1146/annurev-psych-113011-143809
- Schneider, B., González-Romá, V., Ostroff, C., and West, M. A. (2017). Organizational climate and culture: Reflections on the history of the constructs in the Journal of Applied Psychology. *J. Appl. Psychol.* 102, 468–482. doi: 10.1037/apl0000090
- Schneider, B., Salvaggio, A. N., and Subirats, M. (2002). Climate strength: A new direction for climate research. *J. Appl. Psychol.* 87, 220–229. doi: 10.1037/0021-9010.87.2.220
- Shim, S. S., Ryan, A. M., and Anderson, C. J. (2008). Achievement goals and achievement during early adolescence: Examining time-varying predictor and outcome variables in growth-curve analysis. *J. Educ. Psychol.* 100, 655–671. doi: 10.1037/0022-0663.100.3.655
- Skinner, E. A., Kindermann, T. A., Connell, J. P., and Wellborn, J. G. (2009). "Engagement and disaffection as organizational constructs in the dynamics of motivational development," in *Handbook of Motivation at School*, eds K. R. Wenzel and A. Wigfield (Portland, OR: Routledge/Taylor & Francis Group).
- Sortkær, B., and Reimer, D. (2016). *Disciplinary Climate and Student Achievement: Evidence From Schools and Classrooms*. AU Library Scholarly Publishing Services. doi: 10.7146/aul.154.126
- Theis, D., Sauerwein, M., and Fischer, N. (2020). Perceived quality of instruction: The relationship among indicators of students' basic needs, mastery goals, and academic achievement. *Br. J. Educ. Psychol.* 90, 176–192. doi: 10.1111/bjep.12313
- Was, C. A., Al-Harthi, I., Stack-Oden, M., and Isaacson, R. M. (2017). Academic identity status and the relationship to achievement goal orientation. *Elect. J. Res. Educ. Psychol.* 7, 627–652. doi: 10.25115/iejrep.v7i18.1363
- Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *J. Educ. Psychol.* 96, 236–250. doi: 10.1037/0022-0663.96.2.236
- Yi, H., Tian, L., and Huebner, E. S. (2020). Mastery goal orientations and subjective well-being in school among elementary school students: The mediating role of school engagement. *Eur. J. Psychol. Educ.* 35, 429–450. doi: 10.1007/s10212-019-00431-x
- Zaccoletti, S., Altoe, G., and Mason, L. (2020). The interplay of reading-related emotions and updating in reading comprehension performance. *Br. J. Educ. Psychol.* 90, 663–682. doi: 10.1111/bjep.12324
- Zheng, J., Jiang, N., and Dou, J. (2020). Autonomy support and academic stress: A relationship mediated by self-regulated learning and mastery goal orientation. *New Waves-Educ. Res. Dev. J.* 23, 43–63.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Tan, Fan, Wei and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Role of Digital Technologies to Promote Collaborative Creativity in Language Education

Moisés Selfa-Sastre^{1*}, Manoli Pifarré², Andreea Cujba², Laia Cutillas¹ and Enric Falguera¹

¹Department of Specific Didactics, Universitat de Lleida, Lleida, Spain, ²Department of Psychology, Universitat de Lleida, Lleida, Spain

OPEN ACCESS

Edited by:

Olivia López Martínez,
University of Murcia, Spain

Reviewed by:

Hongzhi (Veronica) Yang,
The University of Sydney, Australia
Javier Corbalán,
Universidad de Murcia, Spain

*Correspondence:

Moisés Selfa-Sastre
Moises.selfa@udl.cat

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 04 December 2021

Accepted: 13 January 2022

Published: 09 February 2022

Citation:

Selfa-Sastre M, Pifarré M, Cujba A,
Cutillas L and Falguera E (2022) The
Role of Digital Technologies to
Promote Collaborative Creativity in
Language Education.
Front. Psychol. 13:828981.
doi: 10.3389/fpsyg.2022.828981

The importance of cultivating creativity in language education has been widely acknowledged in the academic literature. In this respect, digital technologies can play a key role in achieving this endeavour. The socio-cultural conceptualization of creativity stresses the role of communication, collaboration and dialogical interaction of creative expression in language education. The objective of this paper is to study the literature focusing on cases of collaborative creativity and technology embedded in language education. To this end, we carry out a systematic revision of state-of-the-art literature consisting of 26 blind peer-reviewed empirical studies selected from several databases that address our main research question, namely, which specific roles and forms of digital technology can be identified in the existing literature that support collaborative creativity in language education. Results show that the features of digital technology unfold a range of learning opportunities in language education and can play three different roles in promoting collaborative creativity: (1) as a tutoring device that guides the implementation of key co-creation skills; (2) as a tool that enables and shapes the development of co-creative thinking skills; and (3) as a medium that creates rich and resourceful environments to stimulate the emergence of collective creative processes. The paper also reveals that these three roles can be performed using a wide range of interactive technologies that encourage students to participate in a rich, co-creative language learning experience and equip learners with key competences to approach complex problems in a globalised and hyper-connected world. Finally, this paper may contribute to developing future language technology-enhanced learning projects capable of promoting key collaborative and creative processes.

Keywords: creativity, collaboration, technology, language, education

INTRODUCTION

Creativity has been identified as the backbone of the skills needed to participate fully in the 21st century society, both in professional and everyday situations that require innovative responses. Also, creativity is regarded as one of the key competences to approach complex problems in a globalised and hyper-connected world (Gretter and Yadav, 2016; Henriksen et al., 2016). Consequently, over the last decade and globally, educational policies have been

implemented aimed at including in curricula competences and contents that promote creativity and innovation (Van de Oudeweetering and Voogt, 2018).

Recent educational research agrees that creativity is the ability to generate novel, appropriate and valuable ideas that can lead to producing original and valuable products or learning outcomes (Rojo, 2019). Therefore, the definition of creativity includes two characteristics: on the one hand *originality* refers to novelty, infrequency and uniqueness; on the other hand *usefulness* refers to utility, appropriateness, fitness or valuableness for the community (Hernández-Torrano and Ibrayeva, 2020).

Recent trends consider creative acts to be socio-cultural in nature and origin. From this perspective, the actions aiming to generate original and valuable products are seen as a social, rather than as an individual phenomenon. Since cultural traditions, social practices and social artefacts regulate and transform the human mind (Shweder and Sullivan, 1990; Glăveanu, 2018), there is a profound interdependence between individuals and their socio-cultural context. The transactions, interactions and activities between these two systems are at the root of creativity. According to Vygotsky et al. (1996), context-mediated action is the socio-cultural genesis of mental functions. Thus, the socio-cultural conceptualisation of creativity emphasises the role of intersubjectivity, communication and dialogical interaction in creative expression (Glăveanu, 2010). There is no doubt that creativity emerges from the close and binding relationship between language and thought.

The focus of educational research on socio-cultural aspects of creativity has led to the coinage of such concepts as *collaborative creativity*, *collaborative creativity*, *group creativity* or *distributed creativity* (Sawyer, 2012; Glăveanu, 2014). Creativity has been defined as a social process (Sawyer, 2012) that emerges and develops among group situations. Indeed, great innovations are often the result of group work, and social judgement and communication play an important role in developing creative products (Glăveanu, 2018). In this respect, collaborative creativity can be considered as the emergence of shared ideas between two or more individuals (Sakr, 2018). Also, Sun et al. (2022) highlight that group processes, such as sharing, negotiation, group communication and interaction processes, are decisive factors of collaborative creativity. Tanggaard (2020) claims that the situated, social nature of creative practices requires a basic dimension of *togetherness* because we create with the support and engagement of others, and the support of tools and artefacts created by former generations. Sakr (2018) claims the importance of taking into consideration the affective dimensions of collaborative creativity. In this paper, we aim to explore how collaborative creativity supported with technology is promoted in language education.

In the specific case of language education, creativity processes emerge when there is a requirement to meet the challenges posed by language teaching. In this article, collaborative creativity in language education is characterised as the collaborative construction of an original and valuable product that gives a creative answer to language learning challenges. The difference between solving these problems individually or collaboratively is abysmal. While in the former, there is

a univocal and unidirectional dialogue, in the latter, i.e. when two or more subjects are involved in the creation of creative thinking, productivity and collaboration between equals is fostered. This results in rich communication of experiences in pursuit of a common good: in other words, 'students are engaged in higher level thinking activities such as problem solving and discussion of complex ideas' (Shuler et al., 2010, p. 11). There is, therefore, an engagement between people to solve what Montalvo (2011) calls *linguistic enigmas*. In this scoping review, we only take into consideration those pieces of research that promote creativity in collaborative environments.

Nowadays, we witness the rapid development of digital and interactive technologies connecting people in multiuser working spaces where users can interact, share and externalise their ideas in open spaces, interplaying with others' voices in different and multiple multimodal channels. As a result of this active online dialogue, new, dynamic and co-created knowledge can emerge (Pifarré, 2019). On this issue, Wegerif (2015) argues that technology shapes human thinking and impacts on how we think and interact with others. Therefore, technology can play an important role in mediating students' creative actions as well as engaging them into meaning-making and collaborative knowledge creation (Säljö, 1999). Also, Mercer et al. (2019) state that we think *with* and *through* artefacts that constitute mediational means endowed with affordances and constraints.

Previous research has characterised distinct features of interactive technology that can play a role in resourcing, promoting and shaping co-creative dialogues (Major et al., 2018). Certainly, digital technologies open up new possibilities for creating and visualising the links between language and thought that allow for a multimodal representation of creative ideas. Ntelioglou et al. (2014) emphasise that the multimodal interaction of technologies facilitates 21st century education in that it promotes broader literacy beyond simple literacy skills by incorporating multiple modes of meaning-making and communication (e.g. auditory, visual, linguistic, spatial and body modes) on the one hand; on the other, the multimodal interaction of technologies provides pedagogical support for learners to optimise their language and literacy learning. For example, digital storytelling has been used to create collaborative storeys as well as to favour language learning specially, learning an L2 in multilingual learning (Anderson et al., 2018; Andayani, 2019; Tyrou, 2021).

Technology can play a crucial role in solving language learning problems. In fact, there is a growing application of digital and interactive technologies to language education. Our interest focuses not only on learning what types of digital environments have been mostly used to stimulate collaborative creativity in language education, but also on reviewing how technology has been used to foster collaborative creativity processes and what results have been obtained from its use.

Our review paper aims to fill this research gap and provide new knowledge on how research in the field of language education has used technology to promote collaborative creativity processes. In this line, identifying the most salient features of existing research may provide an insight into further research

on new pedagogies involving creativity in language education with the use of digital technologies.

THE STUDY

Over the last decade, language education has gradually incorporated the use of a wide range of interactive mobile technology widely used in scientific research. The introduction of such technology in language classrooms has generated opportunities and challenges in the design of learning scenarios that promote collaborative creativity competences. Surprisingly, there are no review studies that analyse the role of technology in supporting and organising collaborative and creative processes in the teaching and learning of language content. Such studies could provide valuable knowledge for outlining theoretical frameworks and pedagogical guidelines to better design language teaching and learning projects that use technology to promote creativity.

To fill this research gap, this article offers a review of educational studies that combine the use of pedagogy and digital technologies to cultivate collaborative creativity competencies in language teaching and learning. This review focuses on research conducted at compulsory and post-compulsory education levels during 2008–2021, as will be seen in the Results and discussion sections.

Our research question is ‘Which specific roles and forms of digital technology can be identified in the existing literature that support collaborative creativity in language education?’

MATERIALS AND METHODS

Bibliographical Search and Criteria for the Selection of Studies

We carried out a scoping review to identify the most relevant studies on the development of collaborative creativity in language education by means of digital technology. This systematic review aims to gather evidence on the role of digital technology in promoting social creativity in language education and carry out a quantitative and qualitative analysis. Our analysis aims to identify and evaluate existing studies rather than use statistical techniques (e.g. metaanalysis) that combine results of these studies to obtain global measuring parameters. Our methodological framework follows the PRISMA statement, a recent guide of systematic reviews launched by Page et al. (2021). The databases consulted are Web of Science (WoS), Scopus and Google Scholar, as they are the most widely acknowledged sources of reference to obtain updated systematic reviews (Codina, 2018: 30–33). We carried out searches in these databases through the advanced search function entering the following keywords and phrases: ‘language education’, ‘creativ*’, ‘collab*’, ‘techno*’ and ‘learning’.

On the first search attempt, we realised that the word ‘technology’ could be restrictive as some articles used keywords, such as ‘computer’, ‘digital’ or ‘video’. In view of this, we decided to replace the word ‘technology’ with the keywords: ‘comput*’,

‘digital’, ‘Web’, ‘video’, ‘blog’, ‘Wiki’ and ‘podcast’. As a result, we managed to collect an appropriate selection of studies for our research. It should be noted that our search covered the last 13 years (from 2008 to 2021), as during these years, there has been an increased integration of technology in language teaching and learning classrooms. In this respect, creativity and collaboration among peers in teaching and learning flourish when learners need to give a novel and original reply to the group.

We established inclusion and exclusion source registers for the systematic review. They are indicated in **Table 1**.

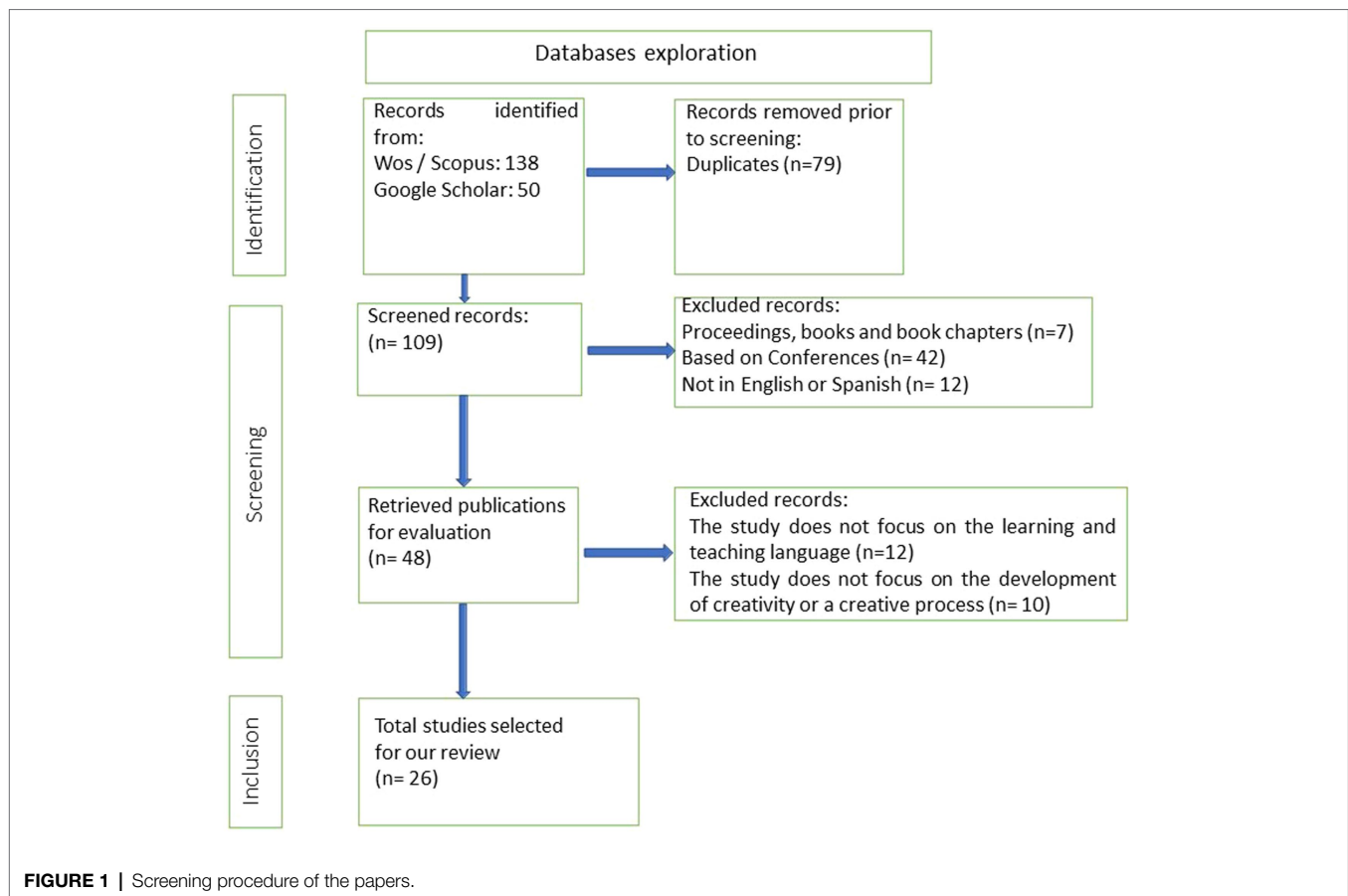
Two independent reviewers conducted all stages of study selection for this review; discrepancies in papers that partially met the criteria were re-examined by a third reviewer and resolved by consensus. Our initial pool was 188 articles, of which 79 were duplicates as they appeared in all three databases consulted. Through the first screening, we discarded the following three types of documents: (a) studies from conferences and/or communications: only blind peer-reviewed articles were considered; (b) books and book chapters were discarded due to the difficulties of having open access to these documents; and (c) articles that were not written in English or Spanish. **Figure 1** summarises the screening procedure and the inclusion and exclusion criteria for the paper review selection.

After this first search, 48 scientific articles were read and assessed for their suitability to achieve the objectives of our research. From these, 12 of them were discarded as they were not directly related to the domain of language teaching and learning. In addition, 10 more articles did not directly discuss research linked to the development of collaborative creativity in language teaching and learning processes. Finally, we narrowed down a selection of 26 relevant studies in line with the research question of this study. A description of all the selected papers is shown in **Table 2**.

We then categorised the 26 studies as they constituted our research basis. After reading, checking and discussing the 26 selected papers, we agreed on the following: firstly, including the paper in one specific category or role of technology. Secondly, analysing how each type of technology was used in each paper to promote collaborative creativity. Thirdly, checking on the discrepancies solved using a consensus-based approach.

TABLE 1 | Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Publications were included if they report on collaborative creativity processes with technology	Conference proceedings as we focused on blind peer-reviewed publications
Publications that focus on teaching and learning processes in language education	Books and book chapters were excluded because of accessibility difficulties
Publications that were peer-reviewed	Publications that were not written in Spanish or English
Publications focused on primary and secondary education as well as university studies	
Papers published between 2008–2021	



RESULTS

Table 2 provides an overview of the core data extracted from the selected studies. In order to identify the possible roles that technology could play in promoting students' collaborative creativity in language education, we were inspired by the different ways of conceptualising the relationship between technology and teaching thinking and creativity developed by Loveless (2007) and Wegerif (2015). For the purposes of this study, as many as three different roles of technology in promoting students' collaborative creativity in language education were identified as: (1) technology as a tutor that induces and models the execution of key co-creative processes for solving language challenges; (2) technology as a tool whose utilisation and appropriation of its characteristics by the students becomes an instrument to think creatively and collaboratively during language learning; and (3) technology as a medium or an environment that prompts the development of key collaboration and creativity processes.

Figure 2 displays the results of the roles that technology plays in promoting students' collaborative creativity in language education and the forms of technology used. As shown in **Figure 2**, technology as a tool is the most frequent role among the studies reviewed ($n=12$). This role was introduced using a wide range of digital technologies. In any case, audio and video platforms are the most common forms of technology

used as instruments that facilitate thinking co-creatively. Out of these 26 studies, 11 of them promoted the use of technology in language education as a medium that provokes co-creation. Fewer studies were found with a tutor role technology ($n=3$).

Six forms of digital technology were identified in the studies reviewed in order to promote collaborative creativity in language education for all students. The studies analysed used mainly audio and video platforms ($n=10$). A few studies used Blog ($n=5$) and web-based environments ($n=4$) and fewer studies used wiki ($n=3$) or mobile technology ($n=3$). Finally, a limited number of studies introduced word processing ($n=1$). Audio and video platforms were the most frequent ones. When digital technologies were implemented as a tutor, Blog ($n=2$) was the main one.

DISCUSSION

Recent research has shown that interactive technologies provide a set of tools than can enrich the learning context and nurture collaborative creativity processes (Henriksen et al., 2016). A specific technology imposes certain constraints, establishes preconditions for students' behaviours and opens up a range of learning opportunities. Because of this, there is a need to analyse how technology is used to promote collaborative creativity in language education.

TABLE 2 | Summary of review studies reporting on collaborative creativity practices with technology in the teaching and learning of languages.

Authors	Technology Role	Technology Type	Creative processes	Main results
Armstrong and Retterer, 2008	Medium	Blog	Collaborative writing in a blog, as a second language learning (Spanish)	Students who have written the longest texts have improved their oral expression in Spanish After participating in the blog writing experience, 100% of the students feel more confident when writing in Spanish
Lund and Rasmussen, 2008	Medium	Wiki	Collaborative writing to produce wiki content that describes a typical British town Creativity: the invention of a city and the aspect of its presentation in the wiki	During the project development, the students have gone through writing processes as interpretation, construction and reconstruction Coordination with group members has created a group task identity and has made students commit to the task
Mak and Coniam, 2008	Medium	Wiki	Collaborative writing to produce wiki content that describes the different facilities and features of their school. Creativity: the aspect of the brochure	Using the wiki as an online collaborative writing environment, students have improved their written expression skills: expanding, rearranging and correcting texts
Rojas-Drummond et al., 2008	Medium	Word processing	Co-construction of texts and multimedia products	Establishment of intertextual and intercontextual relations between texts through the use of ICTs Development of dialogic and textual production strategies Appropriation of diverse cultural artefacts for the construction of knowledge
Yang and Huang, 2008	Medium	Web and Apps	Sharing oral and written productions among peers to learn an L2	Difficulty of integrating technology into the L2 classroom at initial levels of L2 teaching and learning Proposals for concrete integration of technology in the L2 classroom
Comas-Quinn et al., 2009	Tool	Mobile phones Blog	Creation of digital texts to learn about an L2 and its culture	Acquisition of the linguistic and cultural components of an L2 through the creation of digital texts and narratives
Kukulska-Hulme, 2009	Tool	Mobile phones	Mobile learning as a new way to learn L2	The mobile device facilitated the communication between students and teacher to learn the meaning of words, outside the classroom
Lee, 2009	Tool	Blog Podcast (Audacity, iMovie)	Collaborative writing to enhance cultural awareness	This article reports a Spanish–American telecollaborative project through which students created blogs and podcasts for intercultural exchanges in light of socio-cultural perspectives
Kilickaya, 2010	Medium	Audio and video platform <i>Xtranormal</i>	Creating animated films with voice-over audio	Generation of listening contexts to improve writing, reading and pronunciation of an L2
Montalvo, 2011	Tutor	YouTube Blog	Creative Multimedia Riddles	Improved association of ideas, analysis of metaphors and discovery of analogies Distinct cognitive and sensory experience in multimedia riddle interaction
Akinwamide and Adedara, 2012	Medium	Digital audio and radio platforms	Digitize pedagogy in language teaching and learning	Digitize language learning, to facilitate its study outside the classroom Examples of technological tools as a means of language learning
Contreras Salas, 2012	Tutor	Web 2.0 Social media Wiki Blog Podcast <i>Folksomies</i>	Organising social networks. Sharing videos and photos. Creating wikis	Encouraging creative and participatory work using ICT Continuous review of methodological guidelines for teaching English
Lorenzo et al., 2013	Medium	<i>MMOL (Massively Multiuser Online Learning)</i>	Development of microcontent immersed in collaborative virtual environments	The use of MMOL for second language learning shows that collaboration in a 3D educational context, in combination with the use of communication tools (chat, video chat or VoIP) and intelligent assistants (chatbots or NPCs), has a positive effect on the individual acquisition of language content

(Continued)

TABLE 2 | Continued

Authors	Technology Role	Technology Type	Creative processes	Main results
Mellati and Khademi, 2014	Tool	Web based (<i>Net, Email</i>)	Write an essay in a digital environment and receive and give feedback to peers	Encouraging writing in a digital environment Promotion of E-collaboration and collaborative work: give feedback to classmates' essays
Ntelioglou et al., 2014	Tool	Power Point iMovie iPhoto	Descriptive narratives	High quality of texts More participatory and multimodal learning dynamics. Connection to learners' lives
Mellati and Khademi, 2015	Tool	WhatsApp	Collaborative writing using WhatsApp	The use of the application facilitates a more colloquial expression, which gives a sense of belonging to the group
Stevenson et al., 2015	Tool	<i>Storyboards</i>	Creative linguistic activity	Generating and representing creative ideas for storytelling Technology facilitates the participation of diverse members and the development of multimodal literacy
Cruz and Orange, 2016	Tutor	WebQuest Gloster Kahoot	Creative oral presentations	Technology mentors key creative processes
Naqvi and Al Mahrooqi, 2016	Tool	Digital video	Production of an audiovisual message and write a digital text about a commercial product	The use of digital tools in collaborative and creative environments improves motivation for second language learning
Anderson et al., 2018	Medium	Multilingual Digital Storytelling (MDST)	Multilingual digital writing for multiliteracy development	It demonstrates the importance of an integrated and inclusive approach to languages in the framework of multilingualism
Schmoelz, 2018	Medium	Digital Storytelling (DST)	Creation of digital narratives to work and study the co-creativity	The work with digital narratives encourages co-creativity with a greater emotional involvement of the students and a greater commitment and control over the activity and the final results
Andayani, 2019	Tool	Digital storytelling using Microsoft Power Point or Microsoft Video Maker	Digital storytelling projects created by English student teachers for young learners of English as Foreign Language (EFL)	The use of digital narratives has several benefits for students: integration of technology, implementation of pedagogical theories, increased motivation, reduced anxiety about public speaking, and enhanced creativity
Olivier, 2019	Medium	Short videos	Creation of short videos by students with content related to language learning	Videos can be used as a means to motivate students to critically interact with content and to collaborate with new technologies to learn a language
Chubko et al., 2020	Tool	Digital storytelling (DST)	Intervention with DST created by STEM teachers for students with English as L2	Digital narratives help to improve the learning of scientific content, through collaborative and creative environments, by students of English as L2, at the same level as students who have it as L1
Yang and Yeh, 2021	Tool	Making videos (<i>YouTube</i>)	Production of videos to learn about the cultural component of a language	Making promotional videos on YouTube helped English students as a foreign language to be prepared to become socio-cultural agents to introduce themselves and their local culture to the world
Tyrou, 2021	Tool	Wikis	Wiki-mediated L2 collaborative writing	Collaborative writing through technology promotes, among other things, the improvement of: revision of texts (self-correction and peer correction), collaboration, knowledge of foreign cultures and languages

The qualitative analysis of the papers selected for this review identified three different roles of technology for promoting students' collaborative creativity in language education, namely: tutor, tool and medium. In this section, we address the discussion of the results obtained in relation to these three different roles of technology.

Technology as a Tutor of Co-creative Thought

Digital technologies can be seen as gadgets selected to guide a creative activity on the teaching and learning of a given

linguistic content. From this point of view, technology can act as a tutor that encourages creative thinking by following pre-established guidelines and the design of scripts or prompts that promote the performance of specific creative skills. An example of this use in language learning is developed by Cruz and Orange (2016) and applied to Master studies. For their study, they use the multimedia poster *Gloster* to support and improve oral communication on a topic. In this study, technology promotes the development of key creative processes as it can increase opportunities to explore and play with materials, information and ideas around oral communication in a given

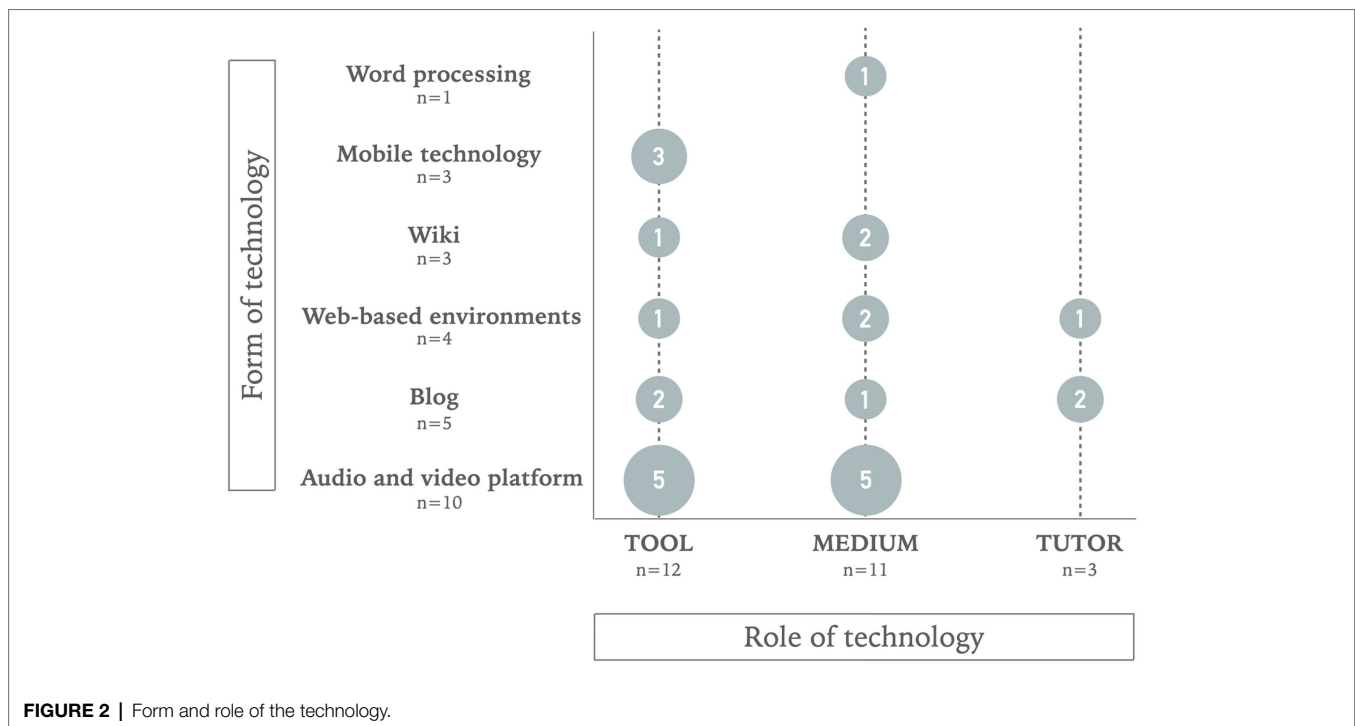


FIGURE 2 | Form and role of the technology.

language. The teacher often plays an important role in the use of technology, as it is the teacher who introduces technology in the classroom to teach curricular content creatively. However, the content-related use of technology can help raise awareness of the ways in which creativity relates to learning curricular knowledge.

In the study by Contreras Salas (2012), Web 2.0 tools are effective in creating information in collaboration with others, organising social networks, sharing videos and photos, and creating wikis, blogs, podcasts and *folksonomies*. The aim of their effective design was to offer a creative *virtual world* in which all the proposals related to the use of ICT were collected in a *Methodology Guide* as a tutorial guide for developing creative thinking. From this guide, teachers, in this case English ones, can draw different possibilities for the use of ICT which are not linear, but transversal and offer multiple possibilities of use and interaction. The result of this study showed that the inclusion of ICT in the teaching of English changed the dynamics of teaching in the learning of this language in 50% of the students. Indeed, students managed time asynchronously as participatory and collaborative work was necessary to carry out learning activities. On the other hand, teachers continuously reviewed the technological learning guidelines contained in the *Methodology Guide*. In this way, the learning process of the students was positively influenced by the methodological approaches of the language teacher.

Similarly, Montalvo (2011) focused on Peruvian primary schoolchildren—years 4, 5 and 6. His study highlights the importance of collaborative learning in linguistic learning situations mediated by digital technologies. These technologies consist in audiovisual riddles which favour the exercise of

creative thinking. Riddles represent a dialogical game between two or more people in which the riddle posed by the sender is tackled by the receivers, thereby establishing a dialogical game between the two. In order to ensure that audiovisual riddles are more widely accepted among new generations, the sender resorts to digital technology, giving them a digital treatment through YouTube. As Igarza (2009: p. 214) points out ‘YouTube is perhaps the Google of the next generation’. In addition to this tool, a blog recorded the response to each of the five riddles posed. Each of the riddles was based on an animated image and a text related to the content of that image. The solution to the task was carried out in groups of two, three or four students, which made it possible to observe the dynamics established among group members when working on the answer to a riddle. The conclusions of the study show that collaborative learning is one of the most functional ways of working in education. It is also closely related to what Montalvo (2011: p. 130) refers to as *collective intelligence*, today associated with Web 2.0 and social networks.

Creativity, peer collaboration and the use of technology have been, as shown above, the object of study, application and analysis in language teaching and learning environments, in which learners play an active role. However, as Contreras Salas (2012) points out in a study with Degree students of Elementary Education specialising in Humanities, Spanish and English of the Universidad Cooperativa de Colombia, based in Bucaramanga, it is necessary to encourage creativity using technology on teachers who have to teach language content. One of the technological avenues most widely explored in language teaching is Web 2.0, as it allows the

teacher to share information and specific guides with the learner through the World Wide Web. Contreras Salas (2012) shows in his research how English teachers create Web 2.0 activities that offer different possibilities of working with ICT.

Technology as a Tool That Facilitates Thinking Co-creatively

Socioculturalism argues that subjects learn to think by internalising the use of cultural tools, such as language or technology, which later become cognitive or critical thinking tools (Vygotsky, 1987). *Instrumental genesis* (Rabardel and Bourmaud, 2003) addresses the connection of human agents and technical artefacts through the concept of instrument. An instrument is a heterogeneous entity, composed of a technical artefact and a human agent. The instrument arises from a double developmental movement, which connects the artefact and its scheme of use. From this instrument, the agents interact and develop a creative product (Overdijk et al., 2012). Technology can thus be seen as a tool used to shape and develop an activity. Instrumentalization therefore changes the tool at the same time as it changes the subject using it.

Digital technologies offer different possibilities to solve language challenges creatively. Users can convert features of digital technologies into instruments for thinking that promote key collaborative and creative processes for solving a language activity.

The qualitative analysis of the papers selected for this review distinguished three different uses of digital technologies as instruments for promoting collaborative creativity actions in language education: (a) as a co-participation and engagement tool; (b) as a multimedia tool that enhances collaborative and creative writing strategies; and (c) a tool that supports linguistic thinking. Next, we will discuss these three uses of technology as a tool.

Technology as a Tool That Enhances Participation and Engagement of All Group Members to Jointly Create Knowledge

Engagement of all group members and being together is a basis of being creative (Tanggaard, 2020). Technology makes it possible to create and narrate a storey collaboratively by means of such tools as iMovie, iPhoto or digital storyboard (Stevenson et al., 2015). These are examples of the use of technology as a tool that facilitates creative linguistic activity of primary and secondary education students. This tool supports the participation of all group members to create joint knowledge and solve a complex creative task.

Kukulska-Hulme (2009) explores second language learning in playful digital environments that encourage peer interaction. This research shows the advantages of mobile learning to learn a second language. Taking the youngest generations as the basis of her research, the author shows how English as a second language (ESL) students develop peer-scaffolding strategies to communicate with each other and learn the meaning of new words with a mobile phone.

Multimedia Technology as a Tool That Supports and Generates Co-creative Writing Strategies

The process of writing a digital storey requires the implementation of new strategies related to generating, communicating and negotiating content in a meaningful way through multimedia information, which facilitates the development of literacy that includes aspects related to *doing* and *being* (Sawyer, 2012). Moreover, these new strategies developed while using technology favour the learning of processes associated with creativity, such as generation of ideas, their development and improvement, selection of the best ideas and their representation (Sun et al., 2022).

In a project with college English student teachers, Andayani (2019) describes the advantages of digital storytelling as a tool to promote collaborative learning skills when learning a second language. This study reports on the benefits in the learning of English of 31 pre-service students asked to create a digital fairy tale using multimedia and interactive technology.

The author concludes that the use of digital narratives has multiple benefits for future English teachers, such as the integration of technology, the implementation of pedagogical theories they had previously studied, the increase in motivation to finish the projects, the reduction of public speaking anxiety and, finally, the possibility of adding music and sound effects, which helped to better dramatise the storey. In post-project interviews, some of the study subjects state that the use of this tool fostered their creativity in designing English teaching and learning activities for their future students. They also point out the motivation and interest generated by the task, as well as the fact that visually supported storeys are easier for English learners to understand.

In another case study with a sample of 30 students from Kyrgyzstan aged between 12 and 16 years, Chubko et al. (2020) used digital storytelling in science with non-native English students. This study is an extension of a previous study, the Indigenous Sky Stories program, conducted with Australian primary school students aged 10–12 years (Ruddell et al., 2016). The conclusions reached in Chubko et al. (2020) are that the creation of digital narratives promotes the literacy of scientific concepts, both in students who master the language of instruction, and in those who do not have this mastery, as demonstrated in an extensive case study, consisting of a sample of over 300 Australian and Kyrgyz students. The authors claim that the use of digital storytelling promoted creativity in constructing a digital narrative and the processes involved in this construction reduced the gap between native and non-native English students.

Naqvi and Al Mahrooqi (2016) discuss an experience with Omani university ESL students. These students were divided into two equal groups: one group of students collaboratively created a digital video showing an audiovisual message in English about a commercial product; a second group, based on this video, carried out a collaborative writing exercise in English on the form and content of the video. After carrying out this exercise in different work sessions, the researchers of this study designed a questionnaire to analyse the learners' impressions of learning English as an L2 using digital videos

and collaboratively written reports on these videos. The results obtained showed positive impressions of the group of learners related to the good use of digital tools in collaborative and creative environments for ESL students.

Comas-Quinn et al. (2009) carried out a study with a group of English college students of the Open University of how mobile phones favour spatial mobility for language learning. In this way, from a shared blog, the mobile device becomes a tool for capturing inputs of any kind related to linguistic and cultural structures of an L2 which can then be shared among peers in a blog. The way to share them is through the creation of digital texts and narratives.

Technology as a Tool That Develops Linguistic Thinking

Different studies claim that interactive technology features related with the co-presence, in one single space of multiple different perspectives, stimulates further thinking. In this space, students can make their ideas visible, externalise their thoughts and represent ideas using multimedia and multimodal facilities. These features of technology can support the generation of new ideas, the connection between seemingly disparate bits of information from divergent perspectives and the construction of a holistic view of the information involved. As a result, different studies claim that the use of technology to solve linguistic challenges co-creatively can develop variables related to linguistic thinking. This is related with the notion of thinking creatively in terms of 'we' and the cultural back as the central axis around which novel ideas are generated and a viable approach for addressing creativity as a culturally diverse capacity (Tanggaard, 2020).

In this line, Ntelioglou et al. (2014) carried out a case study in an inner city elementary school with a large population of recently arrived and Canadian-born linguistically and culturally diverse students from Gambian, Indian, Mexican, Sri Lankan, Tibetan and Vietnamese backgrounds, as well as a recent wave of students from Hungary. The study reports how the use of creative digital tools, such as iMovie and iPhoto for writing descriptive texts, had a positive impact on the expression of personal identity. The texts written by the students included photographs of the selected spaces, descriptions, emotions and experiences of the students in these places. These authors also pointed out that the texts were of very good linguistic quality. Therefore, the students learnt a basic competence in language learning: the written expression of the language. Moreover, the use of technological tools changed the dynamics of learning in the classroom towards more participatory learning processes that included aspects of self-identity and emotions. Therefore, instead of promoting quickly installed functional thinking skills, uniformly defined across cultures, technology promoted creativity as it is a kind of agency in the world, differently defined in various contexts because these require us to act in different creative ways according to the circumstances (Glăveanu et al., 2016).

Lee (2009) describes how through collaborative blogging and collaborative podcasting, university students, from America and Spain, developed their communication and cultural

awareness. The blogs and podcasts created were exchanged between the two cultures with a view to offering and receiving feedback for language correctness. The students did not receive prompts on how they should offer feedback, but instead made their own decisions. At the end of the study, students highlighted that they would not have participated in interactive discussions on linguistic and cultural aspects if they had face-to-face meetings.

Tyrou's (2021) involves 92 university students of Italian as a foreign language in a Wiki environment with a series of activities that encouraged creative thinking, such as visiting virtual museums, and then writing texts collaboratively. The study analysed collaborative writing using Wiki tools in second language teaching. The use of Web 2.0 tools promoted improved learning processes through participation, collaboration and teamwork. This type of collaborative writing mainly improved the process of text revision, favouring both self-correction and peer correction. After analysing student perceptions about web 2.0 technologies for language learning, Tyrou (2021) concludes that 'online collaborative wikis tools can increase knowledge of culture and foreign language, promote teamwork and familiarise our students with new technologies and virtual museums' (p. 53). These are recognised creative values capable to develop students' creative capacity (Tanggaard, 2020).

In this line of work, Mellati and Khademi (2015) explore the possibilities of the WhatsApp mobile application as a tool that favours the sense of belonging to a group. The experience was carried out among 68 Iranian students of English aged between 18 and 35, with an intermediate level of English in the context of a course called Online Mobile Language Learning Course. The experience, in terms of collaborative and creative writing, was very positive. The only drawback observed was that, as the course progressed, a more careless use of the language was observed.

Yang and Yeh (2021) proposed the use of YouTube for teaching and learning the socio-cultural component of the English language. The research was carried out among 71 university students who wanted to learn English. In addition to making videos that were later posted on YouTube, this activity was followed by a critical reflection on the audiovisual production made, in order to reflect with the class group on the socio-cultural component that they wanted to transmit.

Also related to writing texts, Mellati and Khademi (2014) reports on the impact of peer assessment in a technology-based language environment on the quality of creative writing and the development of writing skills. In this study, so-called E-collaboration emerges as a highly intrinsically motivated pathway, as cooperative tasks, specifically based in digital environments, lead to the development of group work and communal learning that positively redounds to individual learning. The results of this research showed that peer learning through Computer-Assisted Language Learning can not only facilitate the development of language skills related to writing texts but also enhance intercultural communicative competence and digital literacy, understood as the ability to locate, organise, understand, evaluate and analyse information using digital technology.

Technology as a Medium That Facilitates an Appropriate Context for Co-creation in Language Education

Dynamic and multimodal interaction within a technology environment affords unique opportunities for learners to co-create in language education. Digital technologies can create rich and resourceful environments capable of acting as a medium which stimulates, orchestrates and supports specific creative processes. Sun et al. (2022) claim that the features of digital technologies can enhance key creative processes, such as emerging of new ideas, identifying connections between seemingly disparate bits of information, fostering collaborations, elaborating the information and promoting imaginative expressions.

The qualitative analysis of the papers selected for this review distinguished three different uses of digital technologies as medium for co-creation: (a) building an immersive and creative experience by providing a wide range of technologies; (b) the use of dedicated technology for building a co-creative writing community and (c) orchestrate the collaborative creativity process. Next, we address the discussion of the results obtained in relation to these two uses of technology as a medium for co-creation.

Building an Immersive and Creative Experience by Providing a Wide Range of Technologies

Technology plays a crucial role when it comes to developing creativity and creative learning environments for second language acquisition. As Lorenzo et al. (2013, p. 1615) state, such an environment 'promotes an immersive, creative and collaborative experience in the process of learning a foreign language'. These virtual learning universes can change the nature of teaching by simultaneously providing a social, immersive and creative experience for second language learners (Canfield, 2008; Chan, 2008; Cooke-Plagwitz, 2008; Jeffery and Collins, 2008).

The purpose of the study by Lorenzo et al. (2013) is the creation of a *Massively Multiuser Online Learning (MMOL)* in university classroom, a didactic strategy that makes use of ICT to improve learning processes in a group of students in face-to-face mode. This integrated platform for massively multiuser learning allows the creation, development and deployment of content and activities for teaching a language in a virtual world. The cooperative, collaborative and socially interactive nature of students as well as teachers is based on a 3D online education environment, which in turn is supported using microcontent immersed in collaborative virtual environments. In other words, the microcontent identified in one of the microformats recognised by the *MMOL* tool is the basic unit of these environments. The *MMOL* microformat may be the same as the one used in Web 2.0, but its meta-description requires further improvement so that it can be intensively reused in any virtualised scenario, or, failing that, adapted to the conditions of a specific context.

The results of this action research with the use of *MMOL* for second language learning show that the possibility to cooperate and collaborate in a 3D educational context, in

combination with the use of communication tools (e.g., chat, video chat or VoIP) and intelligent assistants (chatbots or NPCs), help the learner to accept a role of acceptance and objective criticism for group learning, which has a positive effect on individual acquisition of second language linguistic content.

A specific use of audio and video platforms is the *Xtranormal* environment (Kilickaya, 2010) which allows the creation of animated films with voice audio. The creation of digital products of this type, in an L2 teaching and learning university environment, favours the creation of 3D characters playing different roles for language learning. In this way, listening contexts are generated to improve writing, reading and, in particular, the pronunciation of a second language.

Olivier (2019) explored how the creation of videos can be used to motivate students to interact critically with digital content and participate collaboratively using new technologies in learning a language. The experience was carried out among 82 university students, who produced a total of 50 multimodal creations individually, in pairs or triads. Some of these creations consisted in animations created online, others were animations made with PowerPoint with voice-overs, although they referred to all of these with the umbrella term 'video'. The recorded videos are short videos with the purpose of being open educational resources, on topics provided by the teacher, all related to language learning. No instructions were given to the students on how to plan, write the script or shoot the videos. The conclusions of this research show that the students had to face difficulties not so much related to the use of technology, but rather content selection and condensing information, since the videos were limited in duration. In the same way, the students became real actors in this teaching and learning process, and valued the use of technology as a means of encouraging creativity, as this methodology broke away from traditional practices in language teaching and learning.

Akinwamide and Adedara (2012) designed a platform that provided different digital tools to help teachers working at different levels of teaching and learning digitalize the teaching and learning of a language. Finally, the article by Anderson et al. (2018) presents the findings of a global literacy project based on digital storytelling. They work on multiliteracy through a methodology based collaborative and dialogic ways, allowing for the sharing of divergent thoughts in each community. This research demonstrates the importance of an integrated and inclusive approach to languages in the framework of multiliteracy. Authors conclude that multimodal storytelling develops creative and dialogic thinking.

Dedicated Technology for Building a Co-creative Writing Community

Mak and Coniam (2008) recommended the use of the wiki for language learning. In their study, they used the wiki as an online, co-creative and multimedia environment for writing in English (ESL) with 11-year-old students from Hong Kong, who were not used to working collaboratively. In groups of four students, they participated in a project aiming to describe

the facilities and characteristics of their educational centre to create an advertising brochure for promoting the centre. As the project progressed, the text of the group under analysis improved its quality and complexity. In addition, by writing a collaborative text, the students learned to expand, reorganise and correct their own writing and their group mates' writing. The study highlights that the brochure included creative and original multimedia information.

Lund and Rasmussen (2008) also highlight the use of the wiki as a collaborative and creative technology with high school students from a Norwegian institute who participated in a collaborative writing project to describe a typical English city, within the framework of the subject of ESL. The creative component of this activity was found in the invention of a city based on the real characteristics of British cities using both textual description and images. The students undertook the presentation of their wiki, making use of their imagination. During the development of the project, the students interpreted, constructed and reconstructed writing processes. They went through each process as a result of the following actions: reviewing the wikis of the other groups and becoming aware of the global work of all their colleagues; adapting their texts; and coordinating with the members of their own groups to divide their workload. This led them to create a group task identity and commit them to the task.

Armstrong and Retterer (2008) investigated how the use of the Blog influences foreign language learning (in this case, students of Spanish, of unspecified ages, with an intermediate level of Spanish). Two different activities were planned as: (1) writing a storey among the whole group-class; (2) writing several personal blog posts for each small group. For the first activity, the students created a storey together, over the course of 3 weeks. The teacher started the storey and the different groups of students continued to build it on the basis of the following instruction: each group had to add information to the storey twice a week, but not on the same day, so they had to read the contributions of their peers. In the end, they recorded the storey as if it were a movie. As a result of the study, students who wrote on the blog using a significant number of words, improved their oral expression in terms of accuracy of verb tenses and also increased the complexity of their sentences. An anonymous questionnaire to the students about their blog writing experience showed that 100% of students felt more comfortable writing in Spanish at the end of that experience.

Rojas-Drummond et al. (2008) focused on 6–9-year-old students in Mexico City and how they learned collaboratively in creative writing projects through the use of ICTs. They started from a working context that adopted the model of a learning community. This promoted the social construction of knowledge among all participants. The construction of texts and multimedia products of storeys created by groups of children from fourth to sixth grade, through the innovative educational programme *Learning Together*, revealed the dynamic functioning in educational environments of some central

socio-cultural concepts. Thus, collaborative creativity came across in the writing of texts that involved co-construction of texts; the establishment of intertextual and intercontextual relationships between the texts themselves using ICT; the development of dialogic and textual production strategies; and the appropriation of diverse cultural artefacts for the construction of knowledge.

Orchestrate the Collaborative Creativity Process in Language Education

Schmoelz (2018) reports how secondary education students who use digital narratives to encourage co-creativity show a greater commitment, at the time of planning the writing activity and a high control and effectiveness in the development and resolution of the activity. In the digital storytelling phase, students experience enjoyment and fun that allows a better-constructed storeys. The qualitative study covers 125 students who are interviewed, questioned, recorded and discussed.

CONCLUSION

This paper reviews studies of designs of technology-enhanced learning environments that promote collaborative creativity skills in language education. The final objective of this review has been to capture advanced knowledge for designing future language technology-enhanced learning projects capable of promoting key collaborative and creative processes.

This paper aims to fill a gap in educational research around the use of digital technology to promote collaborative creativity skills. Our selection criteria include four essential research variables to enhance creativity in a global knowledge society: collaboration, creativity, technology and language education. Only 26 studies meet all these criteria.

Although digital and interactive technologies are claimed to create a favourable language learning environment capable of fostering creative and collaborative language learning and writing (Wang and Vásquez, 2012), most of the studies reviewed have not been explicitly designed to improve and evaluate creativity as a social and collaborative endeavour. On the contrary, the importance of creativity in these papers is limited to the creation of a purely digital linguistic product, such as a text, a video or a podcast (Andayani, 2019). Therefore, this review paper can be taken as the basis for future research in language education.

From our review study, we conclude, firstly, that the features of digital and interactive technologies enable the design of powerful and rich language learning environments for knowledge co-creation. These technology-enhanced learning environments open up new opportunities for learners to, collaboratively, generate, modify and evaluate new ideas through online and multimodal interaction.

Secondly, the qualitative analyses of the selected papers conclude that technology can play three important roles to favour co-creativity in language education, namely, tutor, tool and medium. Technology can act as a tutoring device that

guides the implementation of key co-creation skills. Therefore, there is a pattern of work and action that leads to solving language problems with digital tools that promote collaborative work, albeit in a sequenced way: Blogs, Wikis, WebQuest, Kahoot and YouTube as the most popular environments (Contreras Salas, 2012).

Besides, technology can act as a tool that enables and shapes the development of co-creative thinking skills. Therefore, creative thinking arises from the use of technology that shapes the thinking of its users. This is where creative writing of digital narratives in environments, such as IMovie, Iphoto, TOEFL Writing Test (Mellati and Khademi, 2014) or *MMOL* for second language learning emerged.

Furthermore, technology can play the role of the medium that creates rich and resourceful environments to stimulate the emergence of collective creative processes. From this point of view, blogging (Armstrong and Retterer, 2008) or the use of Wikis as online collaborative writing environments (Mak and Coniam, 2008) allows students to improve their co-written writing skills by building texts in digital environments and encouraging e-collaboration between them.

Thirdly, six different forms of technologies have been identified in the reviewed studies that promote co-creativity in language education. They are the following: audio and video platforms, web-based environments, wikis, mobile technology and word processors. These forms of technology support online group learning that enthrall students in active and resourceful-user experience for collaborative knowledge creation.

Finally, our work has its limitations that may have conditioned our results because of having discarded papers that could have contributed to answering our research question. Among these limitations, we highlight the following three: (a) limitation in the type of publications considered: only articles that followed a blind peer review procedure were considered; (b) limitation in the language chosen: only articles written in English and

Spanish were included; and (c) limitation in the search keyword strategy and that these could be insufficient to include key articles in our field of study. However, in an attempt to minimise these limitations, firstly, a systematic review methodology was followed. Secondly, the most significant and prestigious databases in the field of education were consulted: Web of Science, Scopus and Google Scholar.

As a final remark, this paper gives evidence of how technology can support the learning of key linguistic and literary processes, such as: speaking, listening, reading and writing. Furthermore, this paper concludes that co-creativity is an intrinsic phenomenon of literary knowledge. However, there is a need to develop future language technology-enhanced learning projects capable of promoting key collaborative and creative processes in language education. We hope this paper may contribute to reaching this objective.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research has been funded by the Ministry of Science and Innovation of the Government of Spain under Grant EDU2019-107399RB-I00.

REFERENCES

- Akinwamide, T. K., and Adedara, O. G. (2012). Facilitating autonomy and creativity in second language learning through cyber-tasks, hyperlinks and net-surfing. *Engl. Lang. Teach.* 5, 36–42. doi: 10.5539/elt.v5n6p36
- Andayani, R. (2019). "Engaging English student teachers in a digital storytelling project for young learners." in *IOP Conference series: Earth and Environmental Science*; March 1, 2019.
- Anderson, J., Chung, Y. C., and Macleroy, V. (2018). Creative and critical approaches to language learning and digital technology: findings from a multilingual digital storytelling project. *Lang. Educ.* 32, 195–211. doi: 10.1080/09500782.2018.1430151
- Armstrong, K., and Retterer, O. (2008). Blogging as L2 writing: A case study. *AACE J.* 16, 233–251.
- Canfield, D. (2008). Using immersive learning environments in Foreign language classes: second life. *CALICO J.* 26.
- Chan, J. (2008). Developing a meaning-focused and task-based virtual learning reality. *CALICO J.* 26.
- Chubko, N., Morris, J. E., McKinnon, D. H., Slater, E. V., and Lummis, G. W. (2020). Digital storytelling as a disciplinary literacy enhancement tool for EFL students. *Educ. Technol. Res. Dev.* 68, 3587–3604. doi: 10.1007/s11423-020-09833-x
- Codina, L. L. (2018). *Revisión Bibliográfica Sistematizada: Procedimientos Generales y Framework Para Ciencias Humanas y Sociales*. Barcelona: Universitat Pompeu Fabra.
- Comas-Quinn, A., Mardomingo, R., and Valentine, C. (2009). Mobile blogs in language learning: making the most of informal and situated learning opportunities. *ReCALL* 21, 96–112. doi: 10.1017/S0958344009000032
- Contreras Salas, O. L. (2012). Web 2.0 as a pedagogical strategy in the process of English language teaching - learning. *Rastros Rostros* 14, 83–88.
- Cooke-Plagwitz, J. (2008). Conversing in the Metaverse: language teaching and learning in second life. *CALICO J.* 26:13.
- Cruz, M., and Orange, E. (2016). 21st century skills in the teaching of foreign languages at primary and secondary schools. *J. Edu. Technol. Special Issue IETC, ITEC, IDEC, ITICAM*, 1–12.
- Glăveanu, V. P. (2010). Paradigms in the study of creativity: introducing the perspective of cultural psychology. *New Ideas Psychol.* 28, 79–93. doi: 10.1016/j.newideapsych.2009.07.007
- Glăveanu, V. P. (2014). *Distributed Creativity: Thinking outside the Box of the Creative Individual*. Cham/Heidelberg: Springer International Publishing.
- Glăveanu, V. P. (2018). Educating which creativity? *Think. Skills Creat.* 27, 25–32. doi: 10.1016/j.tsc.2017.11.006
- Glăveanu, V., Tanggaard, L., and Wegener, C. (2016). *Creativity – A New Vocabulary*. London: Palgrave Macmillan.
- Gretter, S., and Yadav, A. (2016). Computational thinking and media and information literacy: An integrated approach to teaching twenty-first century skills. *Tech Trends* 60, 510–516. doi: 10.1007/s11528-016-0098-4

- Henriksen, D., Mishra, P., and Fisser, P. (2016). Infusing creativity and technology in 21st century education: A systemic view for change. *J. Edu. Technol. Soc.* 19, 27–37.
- Hernández-Torrano, D., and Ibrayeva, L. (2020). Creativity and education: A bibliometric mapping of the research literature (1975–2019). *Think. Skills Creat.* 35:100625. doi: 10.1016/j.tsc.2019.100625
- Igarza, E. (2009). *Burbujas de Ocio*. Buenos Aires: La Crujía.
- Jeffery, A., and Collins, M. (2008). “Immersive learning and role plays in learning in second life.” in *Proceedings of Society for Information Technology & Teacher Education International Conference*. eds. K. McFerrin. March 3, 2008. (Chesapeake, VA: AACE), 2628–2632.
- Kilickaya, F. (2010). *Text to Speech Animated Movie Creation: Possible Uses in Language Education*. İstanbul: IETC.
- Kukulska-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL* 21, 157–165. doi: 10.1017/S0958344009000202
- Lee, L. (2009). Promoting intercultural exchanges with blogs and podcasting: a study of Spanish–American telecollaboration. *Comput. Assist. Lang. Learn.* 22, 425–443. doi: 10.1080/09588220903345184
- Lorenzo, C.-M., Lezcano, L., and Sánchez-Alonso, S. (2013). Language learning in educational virtual worlds – A TAM based assessment. *J. Univ. Comput. Sci.* 19, 1615–1637.
- Loveless, A. (2007). *Creativity, technology and learning – a review of recent literature*. Available at: <http://archive.futurelab.org.uk/resources/publications-reports-articles/literature-reviews/Literature-Review382>
- Lund, A., and Rasmussen, I. (2008). The right tool for the wrong task? Match and mismatch between first and second stimulus in double stimulation. *Int. J. Comput.-Support. Collab. Learn.* 3, 387–412. doi: 10.1007/s11412-008-9050-8
- Major, L., Warwick, P., Rasmussen, I., Ludvigsen, S., and Cook, V. (2018). Classroom dialogue and digital technologies: A scoping review. *Edu. Info. Technol.* 23, 1995–2028. doi: 10.1007/s10639-018-9701-y
- Mak, B., and Coniam, D. (2008). Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System* 36, 437–455. doi: 10.1016/j.system.2008.02.004
- Mellati, M., and Khademi, M. (2014). Peer evaluation in CMC learning environment and writing skill. *Int. J. App. Ling. Eng. Lit.* 3, 220–228. doi: 10.7575/aiac.ijalel.v3n.5p.220
- Mellati, M., and Khademi, M. (2015). The impacts of distance interactivity on Learners’ achievements in online Mobile language learning: social software and participatory learning. *Int. J. Web Based Learn. Teach. Technol.* 10, 19–35. doi: 10.4018/ijwlt.2015070102
- Mercer, N., Hennessy, S., and Warwick, P. (2019). Dialogue, thinking together and digital technology in the classroom: Some educational implications of a continuing line of inquiry. *Int. J. Educ. Res.* 97, 187–199. doi: 10.1016/j.ijer.2017.08.007
- Montalvo, J. (2011). Adivinanzas audiovisuales para ejercer el pensamiento creativo. *Comunicar* 36, 123–130. doi: 10.3916/C36-2011-03-03
- Naqvi, S., and Al Mahrooqi, R. (2016). ICT and language learning: A case study on student-created digital video projects. *J. Cases Info. Technol.* 18, 49–64. doi: 10.4018/JCIT.2016010104
- Ntelioglou, B. Y., Fannin, J., Montanera, M., and Cummins, J. (2014). A multilingual and multimodal approach to literacy teaching and learning in urban education: A collaborative inquiry project in an inner city elementary school. *Front. Psychol.* 5:533. doi: 10.3389/fpsyg.2014.00533
- Olivier, J. (2019). Short instructional videos as multimodal open educational resources in a language classroom. *J. Edu. Multimedia Hyp.* 28, 381–409.
- Overdijk, M., Van Diggelen, W., Kirschner, P. A., and Baker, M. (2012). Connecting agents and artifacts in CSCS: towards a rationale of mutual shaping. *Int. J. Comput. Support. Collab. Learn.* 7, 193–210. doi: 10.1007/s11412-012-9143-2
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71
- Pifarré, M. (2019). Using interactive technologies to promote a dialogic space for creating collaboratively: A study in secondary education. *Think. Skills Creat.* 32, 1–16. doi: 10.1016/j.tsc.2019.01.004
- Rabardel, P., and Bourmaud, G. (2003). From computer to instrument system: a developmental perspective. *Interact. Comput.* 15, 665–691. doi: 10.1016/S0953-5438(03)00058-4
- Rojas-Drummond, S. M., Albarrán, C. D., and Littleton, K. S. (2008). Collaboration, creativity and the co-construction of oral and written texts. *Think. Skills Creat.* 3, 177–191. doi: 10.1016/j.tsc.2008.09.008
- Rojo, A. M. (2019). *La Investigación de la Creatividad en Traducción*. Comares/Interlingua: Granada.
- Ruddell, N., Danaia, L., and McKinnon, D. (2016). Indigenous sky stories: reframing how we introduce primary school students to astronomy—A type II case study of implementation. *Aus. J. Ind. Edu.* 45, 170–180. doi: 10.1017/jie.2016.21
- Sakr, M. (2018). Multimodal participation frameworks during young children’s collaborative drawing on paper and on the iPad. *Think. Skills Creat.* 29, 1–11. doi: 10.1016/j.tsc.2018.05.004
- Säljö, R. (1999). “Learning as the use of tools: A sociocultural perspective on the human technology link,” in *Learning with Computers: Analysing Productive Interaction*. eds. K. Littleton and P. Light (London: Routledge), 144–161.
- Sawyer, R. K. (2012). *The Science of Human Innovation: Explaining Creativity*. Oxford University Press: New York, NY.
- Schmoelz, A. (2018). Enabling co-creativity through digital storytelling in education. *Think. Skills Creat.* 28, 1–13. doi: 10.1016/j.tsc.2018.02.002
- Shuler, P., Hutchins, G., and LaShell, B. (2010). Student perceptions of tablet computers in a cooperative learning environment. *NACTA J.* 54, 11–17.
- Shweder, R. A., and Sullivan, M. A. (1990). “The semiotic subject of cultural psychology,” in *Handbook of Personality: Theory and Research*. ed. L. A. Pervin (New York: Guilford Publications), 399–416.
- Stevenson, M., Hedberg, J., Highfield, K., and Diao, M. (2015). Visualizing solutions: apps as cognitive stepping-stones in the learning process. *Elec. J. E-Learn.* 13, 366–379.
- Sun, M., Wang, M., Wegerif, R., and Peng, J. (2022). How do students generate ideas together in scientific creativity tasks through computer-based mind mapping? *Comp. Edu.* 176:104359. doi: 10.1016/j.compedu.2021.104359
- Tanggaard, L. (2020). Creating together—moving towards a ‘we-paradigm’ in educating for creativity. *Multi. Edu. Rev.* 12, 4–16. doi: 10.1080/2005615X.2020.1720133
- Tyrou, I. (2021). Online collaborative written activities in L2 for the teaching of language and culture. *European journal of foreign. Lang. Teach.* 5, 34–59. doi: 10.46827/ejfl.v5i4.3744
- Van de Oudewetering, K., and Voogt, J. (2018). Teachers’ conceptualization and enactment of twenty-first century competences: exploring dimensions for new curricula. *Curr. J.* 29, 116–133. doi: 10.1080/09585176.2017.1369136
- Vygotsky, L. S. (1987). *Historia del desarrollo de las funciones psíquicas superiores*. La Habana: Editorial Científico-Técnica.
- Vygotsky, L. S., Cole, M., and Luria, A. R. (1996). *El Desarrollo de los Procesos Psicológicos Superiores*. Barcelona: Crítica.
- Wang, S., and Vásquez, C. (2012). Web 2.0 and second language learning: what does the research tell us? *CALICO J.* 29, 412–430. doi: 10.11139/cj.29.3.412-430
- Wegerif, R. (2015). “Technology and teaching thinking,” in *The Routledge International Handbook of Research on Teaching Thinking*. eds. R. Wegerif, L. Li and J. C. Kaufman (New York: Routledge), 427–440.
- Yang, S. C., and Huang, Y. H. (2008). A study of high school English teachers’ behavior, concerns and beliefs in integrating information technology into English instruction. *Comput. Hum. Behav.* 24, 1085–1103. doi: 10.1016/j.chb.2007.03.009
- Yang, S. H., and Yeh, H. C. (2021). Enhancing EFL learners’ intracultural development as cultural communicators through YouTube video-making. *Technol. Pedagog. Educ.* 30, 557–572. doi: 10.1080/1475939X.2021.1925336

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Selfa-Sastre, Pifarré, Cujba, Cutillas and Falguera. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



An Intervention in Reading Disabilities Using a Digital Tool During the COVID-19 Pandemic

Irene Cadime^{1*}, Iolanda Ribeiro¹, Joana Cruz², Maria do Céu Cosme^{1,3}, Diana Meira^{1,3}, Fernanda Leopoldina Viana⁴ and Sandra Santos⁵

¹ Psychology Research Center, School of Psychology, University of Minho, Braga, Portugal, ² Psychology for Positive Development Research Center–CIPD, Lusíada University of Porto, Porto, Portugal, ³ aPsi–Psychology Association of the University of Minho, Braga, Portugal, ⁴ Research Centre on Child Studies, Institute of Education, University of Minho, Braga, Portugal, ⁵ Centre for Research in Higher Education Policies, Matosinhos, Portugal

OPEN ACCESS

Edited by:

Olivia López Martínez,
University of Murcia, Spain

Reviewed by:

Paula López,
Universidad de León, Spain
S. E. Mol,
Leiden University, Netherlands

*Correspondence:

Irene Cadime
ireneCadime@psi.uminho.pt

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 25 January 2022

Accepted: 19 April 2022

Published: 06 May 2022

Citation:

Cadime I, Ribeiro I, Cruz J,
Cosme MC, Meira D, Viana FL and
Santos S (2022) An Intervention
in Reading Disabilities Using a Digital
Tool During the COVID-19 Pandemic.
Front. Psychol. 13:862383.
doi: 10.3389/fpsyg.2022.862383

In the last decade, ICT-based interventions for developing reading skills in children with reading disabilities have become increasingly popular. This study had three goals: (a) to assess the existence of gains in word reading, oral reading fluency and listening comprehension after a Tier 2 intervention using the digital tool “I’m still learning,” which was delivered partially in a remote modality during the COVID-19 pandemic; (b) to investigate whether the gains depended on the students’ gender, the number of sessions attended and the interventionist; and (c) to investigate parents’ perceptions about the suitability and perceived effects of the intervention. A single group design with pre-test and post-test was used. The intervention was delivered to second graders ($N = 81$) flagged as being at-risk for reading disabilities in a universal screening. The analyses showed significant gains in all three outcome variables after the intervention. The gains did not depend on students’ gender, number of intervention sessions attended or interventionist. Parents’ perceptions of the remote intervention were positive. The study findings highlight the potentialities of using technology-based interventions to foster reading skills and suggest that these may be especially useful during lockdowns.

Keywords: reading disabilities, Tier 2 intervention, digital tool, remote intervention, COVID-19

INTRODUCTION

Response to Intervention (RtI) models provide a line of action regarding assessment and intervention in several areas, including reading: evidence-based interventions are delivered, the effectiveness of those interventions is monitored, and the instruction is adjusted based on how a student responds (Fuchs and Fuchs, 2006; van Norman et al., 2020). Typically, RtI includes three Tiers: Tier 1, which includes universal evidence-based instruction; Tier 2, which encompasses more support for some students in addition to general instruction; and Tier 3 which involves a more intense and personalized intervention, often delivered individually, to only a few students (Gartland and Strosnider, 2020). Students enter Tier 2 when flagged in universal screenings that are used to identify students at-risk for reading disabilities (McAlenney and Coyne, 2011). In RtI models the Tier 2 intervention is delivered in small groups, usually ranging from 2 to 10

students (Balu et al., 2015), and is systematic and tailored to the students' needs (Truckenmiller and Brehmer, 2021). Student progress is regularly monitored and those who fail to demonstrate significant development are moved to Tier 3 intervention (Jenkins et al., 2013; Vaughn and Swanson, 2015).

Reading difficulties have serious implications on students' academic success and motivation, and there has been a growing body of research that evidences the need to develop early interventions that are tailored to the needs of each student to reduce the likelihood of more severe disabilities (Arias-Gundín and Llamazares, 2021) and decrease the differences between students (Pfoest et al., 2014). Several studies support RtI effectiveness in the intervention in reading disabilities (e.g., Suggate, 2010; van Norman et al., 2020). In this study, we will focus specifically on Tier 2 reading interventions. A recent meta-analysis (Gersten et al., 2020) focused on the effects of Tier 2 reading interventions on first to third graders considered at-risk for reading disabilities. Most of the 33 analyzed studies included instruction on phonological awareness, decoding, oral reading fluency and spelling, but intervention in vocabulary and comprehension was seldom described. The results showed larger effect sizes for word or pseudoword reading, compared to reading comprehension and fluency. The observation of larger effects in foundational skills, such as decoding or fluency, compared to comprehension has been consistently reported in other meta-analyses focused on the intervention in reading difficulties in the first years of schooling (e.g., Wanzek et al., 2016). The meta-analysis by Gersten et al. (2020) also showed that interventions including phonological awareness training had significantly smaller effects than interventions without it, suggesting that spending time explicitly teaching phonological awareness is counter-productive after children start to develop decoding skills. Another important result was that various intervention characteristics (e.g., type of interventionist, group size) had little moderating influence on the intervention effects. Previous meta-analyses of the effects of Tier 2 type reading interventions in primary school students had also suggested that the positive effects observed do not depend on the number of intervention hours (Wanzek et al., 2016, 2018). Although these meta-analyses provided important insights on Tier 2 interventions in the first years of reading acquisition, one of their limitations was that they did not analyze whether digital tools were used as part of the intervention, or whether they included complementary homework or some type of remote intervention.

Particularly in the last decade, there has been a significant increase in the use of technology for assessment and intervention with learning disabilities (Dean et al., 2021). Several systematic reviews have shown that a wide diversity of technology is used in this field, from text-to-speech tools to computer-based software designed to improve specific skills (Chai and Chen, 2017; Wood et al., 2018; Dogan and Delialioglu, 2020). One of the most widely used computer programs to promote basic reading skills is GraphoGame, which has been adapted for about two dozen of countries (Ojanen et al., 2015; Ahmed et al., 2020; Dean et al., 2021; Lyytinen et al., 2021). A recent systematic review by McTigue et al. (2020) focused on the effects of GraphoGame on word reading skills. This review concluded that the effects

were mostly small, but a more detailed review of the moderators suggested sizeable effects when there was a high degree of adult participation during the intervention. Some research suggests that new technologies can be successfully integrated in interventions in reading disabilities based on the RtI framework. For example, Duijnen (2021) describes a synchronous online fluency intervention with three struggling readers in second and third grade, with similar reading performance. The students were involved in an 8-week small group intervention, totaling 15 sessions of 45 min each. The results indicated a noticeable increase in word reading accuracy, decoding skills, and reading comprehension after the intervention.

The results of a recent systematic review on technology-based interventions for children with reading difficulties, including 45 studies published between 2010 and 2020, indicated that most interventions were multi-component; that is, they addressed more than one reading component (Alqahtani, 2020). The same review also indicated that in 25 studies (55% of the total) the students worked alone with the intervention tool, whereas in the remaining studies an adult was involved. This proliferation of technology in the intervention in reading disabilities has been accompanied by a concern of how they should be employed to maximize the likelihood of positive outcomes. A recent Delphi study aiming to gather guidelines for good practices in the use of technologies for intervention in dyslexia (Lorusso et al., 2021) suggested an overall positive attitude toward the use of ICT-based interventions, with the flexibility, engagement, and cost-effectiveness being pointed out as some of the advantages of the format. Moreover, the 18 experts involved in this study provided some insights on best practices: (a) the intervention should be started before the third grade and should last up to 6 months; (b) the intervention should target phonological awareness, visual abilities, lexical skills, and grapheme-to-phoneme conversion; and (c) the families' compliance and their ability to support the children and to mediate and supervise the completion of ICT-based activities should be considered (Lorusso et al., 2021).

Social validity refers to the perceptions of the participants in an intervention regarding its goals, procedures, and effects (Wolf, 1978; Foster and Mash, 1999). These perceptions are vital given that they may foster (or hinder) intervention sustainability: if participants and other actors involved, such as teachers and parents, perceive the intervention as important, worthwhile, easy to implement and enjoyable, they will be more likely to adhere to it and intervention will be more likely to go on (Kozleski et al., 2021). However, social validity in reading interventions has seldom been explored. A review by Lindo and Elleman (2010) focused on studies on reading interventions published between 2000 and 2006. They reviewed more than 600 studies and found that only 14 studies included data on social validity, and these were focused either on students' ($n = 4$) or teachers' ($n = 10$) perceptions. Parents' perceptions have not been addressed, although they have been involved in some Tier 2 reading interventions, either as an active part in the intervention or in a less active role, as supporters in additional activities carried at home (Gerzel-Short, 2017; Grindle et al., 2019). Moreover, as suggested in the previously referred Delphi study (Lorusso et al., 2021), the families' compliance is a crucial factor for the success

of ICT-based reading interventions and, this depends to a large extent on how important, feasible and effective they think the intervention is.

The Present Study

The outbreak of the COVID-19 pandemic hastened the need to integrate ICT in interventions for a wide range of disorders. Across the world, lockdowns, illness, quarantine, and prophylactic isolation have limited the access of children to education and intervention services. Consequently, there were some attempts to create interventions that could be delivered remotely to children with learning disabilities. However, most of these consisted of delivering traditional interventions using programs such as Zoom for synchronous remote communication (e.g., Alves and Romig, 2021; Beach et al., 2021; Cruz et al., 2021). In this study, the intervention started face-to-face in November of 2020 but was shifted to a remote modality between January and March of 2021 due to lockdown. We conducted a single-group study with pre-test and post-test and used an e-learning platform for interventions in reading disabilities called “I’m still learning” (Ribeiro et al., 2016). This tool includes tasks for the assessment and intervention in primary school children with, or at-risk for, reading disabilities, focusing specifically on phonological awareness, word reading, oral reading fluency and comprehension. Because this study was conducted with second-grade students, the phonological awareness intervention tasks were discarded. Some of the tasks can be performed independently by the students, but for others, the guidance of an adult is required (e.g., to transition between tasks and to provide more specific feedback). As indicated before, studies that address the social validity in reading interventions are scarce. Although, the intervention addressed in this study was delivered by professionals, parents had a supporting role in the part of the intervention that was conducted remotely. Therefore, besides directly assessing the effects of the intervention in students’ abilities, parents’ perceptions were also addressed to collect evidence of social validity for the remote intervention, supported by the digital tool, and conducted out of school ours. Therefore, the goals of this study were: (a) to explore the existence of gains in students’ word reading, oral reading fluency and listening comprehension at the end of a Tier 2 intervention performed using the referred digital tool; (b) to assess whether the gains depended on the students’ gender, the number of sessions attended and the interventionist; and (c) to investigate parents’ perceptions regarding the remote intervention with their children using the digital tool.

MATERIALS AND METHODS

Participants

The sample comprised 81 second-graders who were flagged as being at-risk for reading disabilities in a universal screening ($N = 528$) conducted in a municipality from the North of Portugal. This universal screening was performed at the beginning of the school year in 27 public schools, comprising 29 classes, and included assessments of letter recognition, reading

fluency, and listening comprehension (Santos et al., 2020b). All assessments were administered by trained teachers in a classroom setting. Students who scored below the cutoff scores in the universal screening (<3 points in oral reading fluency and/or ≤ 7 points in listening comprehension) were flagged as being at-risk. All the selected students had fluency deficits and 21 also had concurrent listening comprehension deficits. Only children who recognized letters and who were, at least already capable of identifying words composed of simple CV syllabic structures (consonant + vowel) were included in this sample. Children who did not demonstrate these skills were referred to (and later supported by) other school services. Regarding gender, 36 (44.4%) were boys and 45 (55.6%) were girls. Students were aged between 6 and 8 years old ($M = 6.95$; $SD = 0.391$) and were not engaged in any other intervention or additional support for learning in the school.

Measures

Test of Word Reading (TLP; *Teste de Leitura de Palavras*)

The TLP (Chaves-Sousa et al., 2017a,b) is a standardized test comprising four vertically scaled test forms for students in grades one to four to evaluate word reading. The test forms TLP-1 and TLP-2 were used in this study. Each test version includes 30 single words that are displayed consecutively, in a randomized order, via a computer application. The test administration is untimed. During the test application, word reading accuracy (correct/incorrect) is recorded in the platform by the evaluator. The raw scores (total number of words read correctly) are then converted to a standardized (scaled) score. The standardized scores are in a scale with a mean of 100 and standard deviation of 10. The expected mean standardized score is 100 ($SD = 10$) at the end of the first grade and 109 ($SD = 10$) at the end of the second grade. The test has adequate indicators of reliability and validity (Chaves-Sousa et al., 2017a).

Test of Listening Comprehension of Narrative Texts (TCTMO-n; *Teste de Compreensão de Textos na Modalidade Oral-Narrativo*)

The TCTMO-n (Santos et al., 2015; Viana et al., 2015) is composed of four vertical scaled test forms to assess students’ listening comprehension from first to fourth grades. The test forms TCTMO-n-1 and TCTMO-n-2 were used in this study. Each test form includes four texts. Students heard the recorded narrative texts followed by 30 multiple-choice listening comprehension questions, presented in the same format. The questions had three alternatives (one correct). Questions assessed literal comprehension, inferential comprehension, reorganization, or critical comprehension (Català et al., 2001). The test administration is untimed, and the total number of correct answers is computed and converted to a standardized score. The standardized scores are in a scale with a mean of 100 and standard deviation of 10. The expected mean standardized score is 100 ($SD = 10$) at the end of the first grade and 106 ($SD = 10$) at the end of the second grade. High reliability and evidence of construct and criterion validity has been provided for the test (Santos et al., 2015; Viana et al., 2015).

Test of Reading Fluency (TFL; *Teste de Fluência de Leitura*)

The TFL (Ribeiro et al., 2014) assesses oral reading fluency and consists of an unpublished narrative text that students are required to read aloud. The test administration is individual and has a time limit of 3 min. The number of reading errors is registered, and the mean number of words read correctly per minute is calculated.

Parents' Questionnaire

A questionnaire was designed to assess parents' perceptions of the remote intervention. This self-report questionnaire included 17 items that were answered using a 5-point Likert scale from 1 (Totally disagree) to 5 (Totally agree). The items were developed following the social validity dimensions proposed by Wolf (1978): (a) significance of the goals; (b) appropriateness of the procedures; and (c) importance of the effects. Therefore, the items assessed not only the perceived effects of the intervention, but also the suitability of the intervention and materials, the appropriateness of the methodology, the suitability of the schedule and equipment and the interventionists' performance. One open response question was also presented so that parents could provide additional comments. The questionnaire can be consulted in the **Supplementary Material**.

Procedures

This study was approved by, and conducted according to, the ethical recommendations of the Ethics Committee of the University of Minho. Authorization from the municipality and the school boards was also obtained. Additionally, before participating in data collection and intervention delivery we obtained informed consent forms for each student, signed by their parents/tutors. After the universal screening and before the intervention, students who scored below the cutoff scores were administered standardized measures of word reading, oral reading fluency, and listening comprehension (October–November 2020). These measures were also administered after the intervention (May–June 2021) by the same researchers who delivered the intervention. The first two measures were administered individually. The test of listening comprehension was administered in small groups. At the end of the remote intervention, parents were asked to respond to a questionnaire about their perceptions regarding the intervention carried out during the lockdown. This questionnaire was presented using Google Forms and accessed *via* a link sent *via* email. Although the participation was anonymous and all parents were invited to participate, only 49 parents responded.

Intervention

A Tier 2, small group intervention (3–5 students), was organized to promote fluency and listening comprehension. Groups were organized based on the COVID-19 pandemic sanitary rules at the time, that determined that students from different classes could not mix inside the schools. Therefore, students in each intervention group were from the same class. The intervention incorporated activities from the “I am still learning” online platform. Sessions, each lasting approximately 40 min,

occurred twice a week. The face-to-face intervention started in November 2020 and occurred outside the classroom, in a schedule agreed by the elementary school teacher, during the school day. Remote intervention occurred between January 21st until mid-March of 2021. In these remote sessions, the major divergences from the original intervention were that interventionists delivered the program using Zoom and sessions occurred mostly after school hours in a schedule agreed by parents. A mean number of 26.85 sessions ($SD = 6.62$; Median = 29; Minimum = 8; Maximum = 37). The structure of the intervention was similar for all groups. Parents were present during the remote intervention to help children access the link to participate in the session, and to supervise and observe the children's performance. The intervention was delivered by three of the study authors (henceforth referred to as the interventionists). These interventionists are qualified educational psychologists with experience delivering reading interventions. Each session included an introduction to a new text. Texts had a short length—50 to 200 words—and its content was related to the children's experiences like daily routines, animals, and family. The sessions were structured as follows:

(1) Activation of previous knowledge

Using the text title, students were asked to discuss what they thought the story was about and any personal experiences related to the content of the title. If the students had questions about some words, the meaning was discussed.

(2) Word reading training

A selection of complex words in each text—whether due to their low frequency or due to a complex phonological structure—is available on the platform. Students were trained to read these words using the digital platform: each word appeared individually on the computer screen and students heard the reading of the word while performing silent reading. The platform allowed the students to hear the words as many times as they desired. In a final step, all the words trained were randomly presented on the screen in a list format and students were asked to read them aloud. If the student made an error on any of these tasks, the interventionist provided the correct reading and the meaning of the words.

(3) Oral reading fluency training

First, the digital platform provided a model reading of the full text. Next, each student practiced reading the text or an excerpt from the text. This was first done *via* assisted reading (i.e., a recording playing in the platform and students reading silently; Rasinski, 2003), and later *via* independent reading (i.e., each student read the text out loud and the interventionist assisted if necessary and provided comments on the reading). Students were invited to practice the reading at home and the following session began with a new reading of the same text. The interventionist then gave each student feedback about their speed, accuracy, and prosody.

TABLE 1 | Differences in listening comprehension, word reading, and oral reading fluency between pre-test and post-test.

Outcome variable	T1	T2	<i>t</i> (df)	<i>p</i>	<i>d</i>
	M (SD)	M (SD)			
Listening comprehension	95.24 (9.09)	102.31 (9.08)	−6.789 (74)	<0.001	0.778
Word reading	80.97 (27.32)	104.20 (8.17)	−7.913 (74)	<0.001	1.152
Oral reading fluency	9.05 (9.98)	37.02 (18.30)	−17.279 (74)	<0.001	1.898

M, mean; *SD*, standard deviation; *T1*, pre-test; *T2*, post-test.

(4) Listening comprehension training

Students heard the recorded text again followed by multiple-choice listening comprehension questions. The questions, presented orally by the platform, had three alternatives (one correct). Questions considered literal comprehension, inferential comprehension, reorganization, or critical comprehension (Català et al., 2001). Students discussed the alternative they considered correct and the reasons for that choice. This choice was then recorded on the platform by the interventionist. After the questions were answered, the platform gave feedback by showing the list of the chosen alternatives and an indication of whether the answer was correct. The interventionists then discussed the results with the students and provided correct alternatives when the incorrect answer was given.

Data Analysis

Six students were excluded from the analyses because they did not complete at least one of the measures at pre-test or post-test. The values of skewness and kurtosis were analyzed as indicators of a normal distribution: values lower than |3| for skewness and |7| for kurtosis indicated no robust violations to the assumption of normality (Hair et al., 2009). Independent samples *t*-tests were performed to test the differences in the standardized measures administered at pre-test between boys and girls, and the differences among the groups assigned to each interventionist. When the assumption of homogeneity of variances was not met, the Welch correction was applied. Cohen's *d* was used as a measure of effect size: 0.20 suggests a small, 0.50 a medium, and 0.80 a large effect (Cohen, 1988). Three paired-samples *t* tests were computed to assess change students' skills from pre-test to post-test. A standardized gain score was computed using the following equation: (post-test scores–pre-test scores)/standard deviation at pre-test. A Pearson correlation coefficient was estimated to assess whether gains in oral reading fluency were associated with students' initial levels of word reading. Next, three multiple linear regression models were computed to assess the effects of the students' gender, number of intervention sessions attended and interventionist on the gains obtained by the students in the three measures. Multicollinearity, independence and normality of residuals and the presence of severe outliers were checked prior to the analysis. Parents' perceptions of the intervention were analyzed by computing the

frequencies of the scores for each item of the questionnaire. Statistical analyses were performed using IBM® SPSS Statistics 28.

RESULTS

Student Outcomes

Table 1 presents descriptive statistics and the outcomes of the paired-samples *t* tests. The results indicate a significant increase in all skills from pre-test to post-test. The effect was large in all three variables, but oral reading fluency and word reading had the largest effect sizes.

Regarding gender differences in the scores obtained in the pre-test, no significant differences were found in word reading, $t(66) = -1.908$, $p = 0.061$, $d = 0.416$, oral reading fluency $t(73) = -1.154$, $p = 0.252$, $d = 0.268$, or listening comprehension, $t(73) = 1.231$, $p = 0.222$, $d = 0.286$. Additionally, no differences were found between the groups assigned to each interventionist in any of the three variables measured in the pre-test: word reading, $F(2,72) = 1.176$, $p = 0.314$, $p\eta^2 = 0.032$; oral reading fluency, $F(2,72) = 0.287$, $p = 0.751$, $p\eta^2 = 0.008$; listening comprehension, $F(2,72) = 0.781$, $p = 0.462$, $p\eta^2 = 0.021$.

Table 2 shows the results of regression analysis to assess whether gains depended on the children's gender, the interventionist and the number of intervention sessions attended. None of the variables was a significant predictor of the gains in word reading, oral reading fluency or reading comprehension. The association between word reading in pre-test and gains in oral reading fluency was not statistically significant ($r = 0.216$, $p > 0.05$), suggesting that the gain did not depend on the initial level of word reading.

Parents' Perceptions

Table 3 presents the descriptive statistics of the responses to the parents' questionnaire. Parents' perceptions were particularly positive in the items related to the interventionist's performance (answers ranging from 91.8 to 98% of responses agree or totally agree) and in the items related to the suitability of the intervention content, materials, and structure of the sessions (answers ranging from 81.6 to 93.9% of responses agree or totally agree). Nine parents did not consider that the duration of the intervention was adequate. Three of these indicated that the intervention should be longer, but the remaining six did not provide additional comments. Regarding the perceived effects of the intervention, most parents agreed that the intervention improved their children reading skills (83.7%) and learning (87.8%). Additionally, most parents (85.7%) did not find the remote intervention disruptive for the household activities and reported adequate access to the internet and computer availability for the student to perform the intervention tasks (89.8%).

DISCUSSION

The first two goals of this study were to assess the gains in word reading, oral reading fluency, and listening comprehension at the end of a Tier 2 intervention performed using a digital tool,

TABLE 2 | Results of the regression models to predict standardized gains in listening comprehension, word reading, and oral reading fluency.

Independent variables	Model 1: Listening comprehension				Model 2: Word reading				Model 3: Oral reading fluency			
	R (R ²)	β	<i>t</i>	<i>p</i>	R (R ²)	β	<i>t</i>	<i>p</i>	R (R ²)	β	<i>t</i>	<i>p</i>
Gender	0.215 (0.046)	0.014	0.117	0.907	0.251 (0.063)	−0.227	−1.913	0.060	0.117 (0.014)	0.071	0.583	0.562
Number of sessions		−0.157	−1.241	0.219		−0.093	−0.740	0.462		0.062	0.485	0.629
Interventionist (2)		0.149	1.189	0.239		−0.130	−1.047	0.299		0.095	0.750	0.456
Interventionist (3)		0.001	0.007	0.994		−0.122	−0.928	0.356		0.041	0.303	0.763

To test for the effect of the interventionist, students assigned to interventionist 1 were the reference group.

TABLE 3 | Frequencies of responses in each item of the parents' questionnaire.

Items	Totally disagree	Disagree	Do not agree nor disagree	Agree	Totally agree
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)
Suitability of the intervention and materials					
1. The intervention content was adequate	0 (0%)	1 (2.0%)	2 (4.1%)	21 (42.9%)	25 (51.0%)
2. The materials were adequate	0 (0%)	1 (2.0%)	2 (4.1%)	19 (38.8%)	27 (55.1%)
3. The duration of the intervention was adequate	0 (0%)	2 (4.1%)	7 (14.3%)	21 (42.9%)	19 (38.8%)
Appropriateness of the methodology					
4. The structure of the sessions was adequate	0 (0%)	0 (0%)	3 (6.1%)	20 (40.8%)	26 (53.1%)
5. My son/daughter was cheerful during the intervention sessions	1 (2.0%)	0 (0%)	1 (2.0%)	18 (36.7%)	29 (59.2%)
6. My son/daughter talked about the intervention with me and with his/her colleagues	2 (4.1%)	3 (6.1%)	8 (16.3%)	16 (32.7%)	20 (40.8%)
Suitability of the schedule and equipment					
7. The schedule of the intervention was not disruptive for family life	1 (2.0%)	1 (2.0%)	5 (10.2%)	16 (32.7%)	26 (53.1%)
8. It was possible to use adequately a computer and access internet	0 (0%)	2 (4.1%)	3 (6.1%)	18 (36.7%)	26 (53.1%)
Perceived effects of the intervention					
9. The intervention allowed my son/daughter to remind previous knowledge	0 (0%)	1 (2.0%)	9 (18.4%)	16 (32.7%)	23 (46.9%)
10. The intervention allowed my son/daughter to learn more	0 (0%)	1 (2.0%)	5 (10.2%)	21 (42.9%)	22 (44.9%)
11. The intervention allowed my son/daughter to improve his/her reading skills	0 (0%)	3 (6.1%)	5 (10.2%)	22 (44.9%)	19 (38.8%)
12. The intervention made my son gain the habit of studying at home	0 (0%)	3 (6.1%)	13 (26.5%)	17 (34.7%)	16 (32.7%)
Interventionists' performance					
13. The interventionist organized the intervention adequately	0 (0%)	0 (0%)	1 (2.0%)	14 (28.6%)	34 (69.4%)
14. The interventionist was clear about the content to be learned in each session	0 (0%)	0 (0%)	2 (4.1%)	14 (28.6%)	33 (67.3%)
15. The interventionist clarified any doubts	0 (0%)	0 (0%)	1 (2.0%)	15 (30.6%)	33 (67.3%)
16. The interventionist's work raised the interest of the students in the task	0 (0%)	0 (0%)	3 (6.1%)	15 (30.6%)	31 (63.3%)
17. The interventionist's work contributed to the improvement of my son/daughter's reading skills	0 (0%)	0 (0%)	4 (8.2%)	16 (32.7%)	29 (59.2%)

and to investigate whether the gains depended on the students' gender, the number of sessions attended and the interventionist. The intervention included strategies to promote word reading skills and oral reading fluency, as well as vocabulary and listening comprehension. While most Tier 2 interventions in the early years of schooling have focused on the promotion of basic reading skills (Gersten et al., 2020), the inclusion of activities to promote linguistic skills is also vital given that these contribute directly to decoding and to reading comprehension (Cadime et al., 2017; Santos et al., 2020a).

Regarding the first goal, we found significant improvements for all three outcome variables after the intervention. The largest effect size was observed for fluency, followed by word reading. This finding aligns with previous literature that suggests higher

gains in basic reading skills compared to comprehension in Tier 2 interventions in the first years of schooling (Wanzek et al., 2016; Gersten et al., 2020). This finding may reflect the higher focus of the interventions on these skills compared to the shorter time devoted to fostering vocabulary or listening comprehension (Gersten et al., 2020). Moreover, in our study, the intervention to foster listening comprehension encompassed strategies such as activation of previous knowledge, clarification of difficult words and response to questions and feedback, but did not include other strategies that research has shown to be effective, such as the training of cognitive and metacognitive strategies (Goh and Taib, 2006; Baker et al., 2020). Additionally, it is worth noting that the students participating in this study were in an early stage of reading acquisition, and all of them were experiencing difficulties

in the automatization of reading. In the first years of schooling, phonics and fluency instruction are key components in literacy instruction in the classroom (Spear-Swerling et al., 2016) and research has suggested that struggling readers obtain the largest gains with systematic reading interventions (Suggate, 2010; Cruz et al., 2021). Therefore, the characteristics of our sample might have played a role in the large gains observed in word reading and oral reading fluency. Our results also shown that the gains observed in fluency did not appear to depend on initial levels of word reading. Although the absence of a control group prevents us from attributing the observed gains in reading directly to the intervention program, the results of this study are quite encouraging regarding the use of a digital tool to promote reading and linguistic skills in students at-risk for reading disabilities. It is noteworthy that the use of the tool did not totally replace the necessity for an interventionist. Although some of the tasks could be completed independently by the students, our option was to use the tool as part of a more structured intervention supervised and guided by the interventionists. Although there is a large variation in the degree of independency provided by the ICT-based interventions (Alqahtani, 2020), research is suggestive of higher gains when there is some amount of involvement of adults when using digital tools to foster children's reading skills (McTigue et al., 2020). Regarding the second goal, the results of our study showed that the gains did not depend on the children's gender, the number of sessions attended and the interventionist. This finding is similar to the ones reported in meta-analyses that analyzed the effects of interventions with children with reading disabilities (Wanzek et al., 2016, 2018; Gersten et al., 2020).

The third goal of this study was to investigate parents' perceptions regarding the remote intervention with their children. Parents' perceptions were quite positive regarding its content, structure, and materials used, including the digital tool. However, some parents indicated that the duration of the intervention was not sufficient. Parents' perception of the necessity of more hours of intervention may be influenced by the acknowledgment that, although their children experienced significant intraindividual gains, they were still performing below grade level in reading. As an example, in oral reading fluency, the mean number of words read correctly per minute by the students at the end of the intervention was 37.02 ($SD = 18.30$), a number well below the 90 words indicated as a reference in the curricular benchmarks for Portuguese language in the second grade (Buesco et al., 2015). The finding that students with, or at-risk of, learning disabilities can make large gains when they are provided intensive and systematic intervention, yet still lag behind their peers, has been widely reported in the literature (Gilmour et al., 2019), but research has also suggested that early interventions can contribute to reduce this gap (Wanzek et al., 2018). Overall, the findings of our study provide support for the social validity of the intervention. Although evidence of social validity is not in itself evidence of efficacy of an intervention, it involves dimensions that can contribute to efficacy, such as acceptability and viability (Foster and Mash, 1999). In this case, the data collected from parents suggest that the intervention is viable and will be accepted in a particular setting, namely, in a remote modality where students receive the intervention when at home.

The effects of remote interventions with children have been a recent concern in research, mainly due to the disruption induced by the COVID-19 pandemic. Overall, the research results suggest not only positive effects in the targeted skills (e.g., Duijnen, 2021), but also high acceptability and positive perceptions of feasibility and effectiveness of remote interventions with children with reading disabilities or other neurodevelopmental disorders (Beach et al., 2021; Su et al., 2021). Our study adds to this body of research showing also positive perceptions of a remote intervention supported by a digital tool to promote reading skills in students facing reading difficulties.

The main limitation of this study was the absence of a control group, which precludes us from concluding that the observed gains are directly attributable to the intervention. Also, the intervention groups were not randomly assembled. Therefore, the generalization of results should be made with caution. Another limitation is related to the procedures used to assess the intervention fidelity. Although the use of a standardized online platform potentiates the likelihood that the intervention was administered as intended, other procedures to assure intervention fidelity, such as observations and ratings (King-Sears et al., 2018), were not used. Future studies should consider the implementation of these techniques to assess fidelity. A fourth limitation is related to the limited information gathered on the social validity of the intervention. Only parents' perceptions were collected, using a questionnaire. Future studies should also address students' and teachers' perceptions and use complementary methods such as interviews or focus groups, that allow a more in-depth exploration of how these groups perceive the relevance of the goals, the feasibility of the procedures and the effects of the intervention. A final limitation was that no information regarding families, such as socioeconomic status, was collected. Future studies should address whether the intervention effects or the parents' perceptions vary as a function of the families' characteristics.

Nonetheless, the results provide some important insights on the use of ICT and digital tools in reading interventions. Firstly, our tool was easily integrated both in the face-to-face and in the remote intervention phases, following the guidelines for best practices in the use of ICT in intervention in reading disabilities (Lorusso et al., 2021). Specifically, the parents were involved in the intervention sessions and granted access to the digital tool. Only a few parents reported that the intervention was disruptive for the household. Therefore, our results suggest that the use of digital tools in a remote intervention modality can be a feasible alternative to address the needs of students with reading disabilities during a lockdown or prophylactic isolation in a context of a pandemic, if granted access to a computer and an internet connection. Finally, this study suggests that the intervention presented was effective with struggling readers, combining a RTI framework and the use of digital technologies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the University of Minho. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

IC, IR, FV, and SS: conception and design of the study. JC, MC, and DM: data collection and intervention. IC and SS: statistical data analyses. All authors were involved in interpreting and discussing the results and in drafting the manuscript and revising it critically for important intellectual content.

FUNDING

This work was financially supported by Portuguese national funds through the Foundation for Science and Technology, within the frameworks of the CiPsi–Psychology Research Centre (UIDB/PSI/01662/2020), CIEC–Research Centre on Child Studies (UIDB/CED/00317/2020), and CIPD–Psychology for

Positive Development Research Center (UIDB/04375/2020), as well as by the Northern Regional Operational Program (NORTH 2020), through Portugal 2020 and the European Social Fund (ESF), within the framework of the Integrated and Innovative Plan to Combat School Failure of the Municipality of Valongo, and by Investigare–Association for Research in Reading, Writing and Neurosciences. IC was also supported by the Foundation for Science and Technology (CEECIND/00408/2018).

ACKNOWLEDGMENTS

We would like to thank the support of the Municipality of Valongo and to acknowledge and thank the collaboration of the teachers who performed the screening in the schools. We would also like to thank all the students and parents who participated in this study for their time and effort.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.862383/full#supplementary-material>

REFERENCES

- Ahmed, H., Wilson, A., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., et al. (2020). An evaluation of the efficacy of GraphoGame Rime for promoting English phonics knowledge in poor readers. *Front. Educ.* 5:132. doi: 10.3389/educ.2020.00132
- Alqahtani, S. S. (2020). Technology-based interventions for children with reading difficulties: a literature review from 2010 to 2020. *Educ. Technol. Res. Dev.* 68, 3495–3525. doi: 10.1007/s11423-020-09859-1
- Alves, K. D., and Romig, J. E. (2021). Virtual reading lessons for upper elementary students with learning disabilities. *Interv. Sch. Clin.* 57, 95–102. doi: 10.1177/10534512211001865
- Arias-Gundín, O., and Llamazares, A. G. (2021). Efficacy of the RTI model in the treatment of reading learning disabilities. *Educ. Sci.* 11:209. doi: 10.3390/educsci11050209
- Baker, D. L., Santoro, L., Biancarosa, G., Baker, S. K., Fien, H., and Otterstedt, J. (2020). Effects of a read aloud intervention on first grade student vocabulary, listening comprehension, and language proficiency. *Read. Writ.* 33, 2697–2724. doi: 10.1007/s11145-020-10060-2
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., and Gersten, R. (2015). *Evaluation of Response to Intervention Practices for Elementary School Reading (NCEE 2016-4000). National Center for Education Evaluation and Regional Assistance, U.S. Department of Education.* Available online at: <https://files.eric.ed.gov/fulltext/ED560820.pdf> (accessed December 7, 2021).
- Beach, K. D., Washburn, E. K., Gesel, S. A., and Williams, P. (2021). Pivoting an elementary summer reading intervention to a virtual context in response to COVID-19: an examination of program transformation and outcomes. *J. Educ. Students Placed Risk* 26, 112–134. doi: 10.1080/10824669.2021.1906250
- Buesco, H. C., Morais, J., Rocha, M. R., and Magalhães, V. F. (2015). *Programa e metas curriculares de Português do ensino básico [Curricular benchmarks for Portuguese language in basic education]. Lisboa: Ministério da Educação e Ciência.* Available online at: https://www.dge.mec.pt/sites/default/files/Basico/Metas/Portugues/pmcpeb_julho_2015.pdf (accessed November 12, 2021).
- Cadime, I., Rodrigues, B., Santos, S., Viana, F. L., Chaves-Sousa, S., do Céu, et al. (2017). The role of word recognition, oral reading fluency and listening comprehension in the simple view of reading: a study in an intermediate depth orthography. *Read. Writ.* 30, 591–611. doi: 10.1007/s11145-016-9691-3
- Català, G., Català, M., Molina, E., and Monclús, R. (2001). *Evaluación de la comprensión lectora: pruebas ACL (1o - 6o de primaria).* Barcelona: Editorial Graó.
- Chai, J. T., and Chen, C. J. (2017). A research review: how technology helps to improve the learning process of learners with dyslexia. *J. Cogn. Sci. Hum. Dev.* 2, 26–43.
- Chaves-Sousa, S., Ribeiro, I., Viana, F. L., Vale, A. P., Santos, S., and Cadime, I. (2017a). Validity evidence of the Test of Word Reading for Portuguese elementary students. *Eur. J. Psychol. Assess.* 33, 460–466. doi: 10.1027/1015-5759/a000307
- Chaves-Sousa, S., Santos, S., Viana, F. L., Vale, A., Cadime, I., Prieto, G., et al. (2017b). Development of a word reading test: identifying students at-risk for reading problems. *Learn. Individ. Differ.* 56, 159–166. doi: 10.1016/j.lindif.2016.11.008
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences. 2nd ed.* Mahwah: Lawrence Erlbaum Associates.
- Cruz, J., Mendes, S., Marques, S., Alves, D. R., and Cadime, I. (2021). Face-to-face versus remote: effects of an intervention in reading fluency during COVID-19 pandemic. *Front. Educ.* 6:817711. doi: 10.3389/educ.2021.817711
- Dean, J., Pascoe, M., and le Roux, J. (2021). Information and communication technology reading interventions: a scoping review. *Read. Writ. J. Read. Assoc. South Afr.* 12:a294. doi: 10.4102/rw.v12i1.294
- Dogan, S., and Delialioğlu, O. (2020). A systematic review on the use of technology in learning disabilities. *Ankara Univ. Fac. Educ. Sci. J. Spec. Educ.* 21, 611–638. doi: 10.21565/ozelegitimdergisi.563763
- Duijnen, T. V. (2021). *The Efficacy of a Synchronous Online Reading Fluency Intervention with Struggling Readers.* Victoria: University of Victoria. [Ph.D. thesis].
- Foster, S. L., and Mash, E. J. (1999). Assessing social validity in clinical treatment research: issues and procedures. *J. Consult. Clin. Psychol.* 67, 308–319. doi: 10.1037/0022-006X.67.3.308
- Fuchs, D., and Fuchs, L. S. (2006). Introduction to response to intervention: what, why, and how valid is it? *Read. Res. Q.* 41, 93–99. doi: 10.1598/rrq.41.1.4
- Gartland, D., and Strosnider, R. (2020). The use of response to intervention to inform special education eligibility decisions for students with specific learning disabilities. *Learn. Disabil. Q.* 43, 195–200. doi: 10.1177/0731948720949964

- Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., and Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *J. Res. Educ. Eff.* 13, 401–427. doi: 10.1080/19345747.2019.1689591
- Gerzel-Short, L. (2017). “We Conquered This Together”: tier 2 collaboration with families. *Sch. Community J.* 28, 85–112.
- Gilmour, A. F., Fuchs, D., and Wehby, J. H. (2019). Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Except. Child* 85, 329–346. doi: 10.1177/0014402918795830
- Goh, C., and Taib, Y. (2006). Metacognitive instruction in listening for young learners. *ELT J.* 60, 222–232. doi: 10.1093/elt/ccl002
- Grindler, C., Tyler, E., Murray, C., Hastings, R. P., and Lovell, M. (2019). Parent-mediated online reading intervention for children with Down Syndrome. *Supp. Learn.* 34, 211–230. doi: 10.1111/1467-9604.12249
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2009). *Multivariate Data Analysis*. 7th ed. New Jersey: Prentice Hall.
- Jenkins, J. R., Schiller, E., Blackorby, J., Thayer, S. K., and Tilly, W. D. (2013). Responsiveness to intervention in reading: architecture and practices. *Learn. Disabil. Q.* 36, 36–46. doi: 10.1177/0731948712464963
- King-Sears, M. E., Walker, J. D., and Barry, C. (2018). Measuring teachers’ intervention fidelity. *Interv. Sch. Clin.* 54, 89–96. doi: 10.1177/1053451218765229
- Kozleski, E. B., Hunt, P., Mortier, K., Stepaniuk, I., Fleming, D., Balasubramanian, L., et al. (2021). What peers, educators, and principals say: the social validity of inclusive, comprehensive literacy instruction. *Except. Child* 87, 289–306. doi: 10.1177/0014402920969184
- Lindo, E. J., and Elleman, A. M. (2010). Social validity’s presence in field-based reading intervention research. *Remedial Spec. Educ.* 31, 489–499. doi: 10.1177/0741932510361249
- Lorusso, M. L., Borasio, F., Rold, M., Da, and Martinuzzi, A. (2021). Towards consensus on good practices for the use of new technologies for intervention and support in developmental dyslexia: a Delphi study conducted among Italian specialized professionals. *Children* 8:1126. doi: 10.3390/children8121126
- Lyytinen, H., Semrud-Clikeman, M., Li, H., Pugh, K., and Richardson, U. (2021). Supporting acquisition of spelling skills in different orthographies using an empirically validated digital learning environment. *Front. Psychol.* 12:566220. doi: 10.3389/fpsyg.2021.566220
- McAlenney, A. L., and Coyne, M. D. (2011). Identifying at-risk students for early reading intervention: challenges and possible solutions. *Read. Writ. Q.* 27, 306–323. doi: 10.1080/10573569.2011.596100
- McTigue, E. M., Solheim, O. J., Zimmer, W. K., and Uppstad, P. H. (2020). Critically reviewing GraphoGame across the world: recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Read. Res. Q.* 55, 45–73. doi: 10.1002/rrq.256
- Ojanen, E., Ronimus, M., Ahonen, T., Chansa-Kabali, T., February, P., Jere-Folotiya, J., et al. (2015). GraphoGame-A catalyst for multi-level promotion of literacy in diverse contexts. *Front. Psychol.* 6:671. doi: 10.3389/fpsyg.2015.00671
- Pfost, M., Hattie, J., Dörfler, T., and Artelt, C. (2014). Individual differences in reading development: a review of 25 years of empirical research on Matthew effects in reading. *Rev. Educ. Res.* 84, 203–244. doi: 10.3102/0034654313509492
- Rasinski, T. (2003). Oral reading in the school literacy curriculum. *Read. Res. Q.* 38, 510–522. doi: 10.1598/rrq.38.4.5
- Ribeiro, I., Santos, S., Baptista, A., Choupina, C., Brandão, S., Cadime, I., et al. (2016). A utilização da plataforma “Ainda estou a aprender” na avaliação e na intervenção nas dificuldades na aprendizagem da leitura: um estudo de caso [The platform “I’m still learning” in the assessment and intervention in reading disabilities: a case study]. *Calidoscópio* 15, 30–44. doi: 10.4013/cld.2017.151.03
- Ribeiro, I., Viana, F. L., Cadime, I., Chaves-Sousa, S., Santos, S., Silva, C., et al. (2014). *Teste de Fluência da Leitura. Manual não publicado [The Test of Reading Fluency. Unpublished manual]*. Braga: School of Psychology, University of Minho.
- Santos, S., Ribeiro, I., Viana, F. L., Cadime, I., Morgado, C., and Baptista, A. (2020b). *Manual de aplicação e de cotação das provas de rastreio - 2o ano, versão 3 [Manual for the administration and scoring of the screening test - Grade 2, version 3]*. Maia: Lusoinfo.
- Santos, S., Cadime, I., Viana, F. L., and Ribeiro, I. (2020a). Cross-Lagged relations among linguistic skills in European Portuguese: a longitudinal study. *Read. Res. Q.* 55, 177–192. doi: 10.1002/rrq.261
- Santos, S., Viana, F. L., Ribeiro, I., Prieto, G., Brandão, S., and Cadime, I. (2015). Development of listening comprehension tests with narrative and expository texts for Portuguese students. *Span. J. Psychol.* 18, 1–7. doi: 10.1017/sjp.2015.7
- Spear-Swerling, L., Lopes, J., Oliveira, C., and Zibulsky, J. (2016). How Portuguese and American teachers plan for literacy instruction. *Ann. Dyslexia* 66, 71–90. doi: 10.1007/s11881-015-0107-x
- Su, W.-C., Srinivasan, S., Cleffi, C., and Bhat, A. (2021). Short report on research trends during the COVID-19 pandemic and use of telehealth interventions and remote brain research in children with autism spectrum disorder. *Autism* 25, 1816–1822. doi: 10.1177/13623613211004795
- Suggate, S. P. (2010). Why what we teach depends on when: grade and reading intervention modality moderate effect size. *Dev. Psychol.* 46, 1556–1579. doi: 10.1037/a0020612
- Truckenmiller, A. J., and Brehmer, J. S. (2021). Making the most of Tier 2 intervention: what decisions are made in successful studies? *Read. Writ. Q.* 37, 240–259. doi: 10.1080/10573569.2020.1768612
- van Norman, E. R., Nelson, P. M., and Klingbeil, D. A. (2020). Profiles of reading performance after exiting Tier 2 intervention. *Psychol. Sch.* 57, 757–767. doi: 10.1002/pits.22354
- Vaughn, S., and Swanson, E. A. (2015). Special education research advances knowledge in education. *Except. Child* 82, 11–24. doi: 10.1177/0014402915598781
- Viana, F. L., Santos, S., Ribeiro, I., Chaves-Sousa, S., Brandão, S., Cadime, I., et al. (2015). Listening comprehension assessment: validity studies of two vertically scaled tests for Portuguese students. *Univ. Psychol.* 14, 345–354. doi: 10.11144/Javeriana.upsy14-1.lcav
- Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., and Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *J. Learn. Disabil.* 51, 612–624. doi: 10.1177/0022219418775110
- Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., and Capin, P. (2016). Meta-analyses of the effects of Tier 2 type reading interventions in grades K-3. *Educ. Psychol. Rev.* 28, 551–576. doi: 10.1007/s10648-015-9321-7
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *J. Appl. Behav. Anal.* 11, 203–214. doi: 10.1901/jaba.1978.11-203
- Wood, S. G., Moxley, J. H., Tighe, E. L., and Wagner, R. K. (2018). Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *J. Learn. Disabil.* 51, 73–84. doi: 10.1177/0022219416688170

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cadime, Ribeiro, Cruz, Cosme, Meira, Viana and Santos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Investigating Effects of Small-Group Student Talk on the Quality of Argument in Chinese Tertiary English as a Foreign Language Learners' Argumentative Writing

Hui Helen Li^{1,2†} and Lawrence Jun Zhang^{2*†}

¹ School of Foreign Languages, Wuhan University of Technology, Wuhan, China, ² Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

OPEN ACCESS

Edited by:

Pedro García Guirao,
WSB Universities, Poland

Reviewed by:

Paula López,
Universidad de León, Spain
Mostafa Janebi Enayat,
University of Maragheh, Iran

*Correspondence:

Lawrence Jun Zhang
lj.zhang@auckland.ac.nz

†ORCID:

Hui Helen Li
orcid.org/0000-0002-4118-6396
Lawrence Jun Zhang
orcid.org/0000-0003-1025-1746

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 02 February 2022

Accepted: 24 May 2022

Published: 14 June 2022

Citation:

Li HH and Zhang LJ (2022)
Investigating Effects of Small-Group
Student Talk on the Quality
of Argument in Chinese Tertiary
English as a Foreign Language
Learners' Argumentative Writing.
Front. Psychol. 13:868045.
doi: 10.3389/fpsyg.2022.868045

Previous studies have offered a rationale for engaging students in small-group student talk for the planning of L2 individual writing. To further investigate whether such talk effectively promotes the quality of argument in the context of Chinese tertiary EFL learners' argumentative writing and whether such effects could be retained, the current study adopted a quasi-experimental design with a pretest, a posttest, and a delayed posttest in two intact EFL classes. The performance of the intervention group and the comparison group were scrutinized to examine the effects of the intervention. The analytic scores on six components of the writing task (claim, data, counterargument claim, counterargument data, rebuttal claim, and rebuttal data) and the holistic writing scores cumulated of all these components were measured to see the immediate and sustained effects. Significant changes of the holistic scores in both the immediate posttest and the delayed posttest indicated that such small-group student talk enabled students in the treatment class to achieve better performance in the overall quality of argumentation compared with those in the comparison class. Statistical analyses revealed immediate and sustained effects of small-group student talk on the quality of counterargument claim, counterargument data, and rebuttal claim. Counterargument claim was the only element in which students in both classes made significant improvement, but the treatment class demonstrated a larger effect size. No discernible differences were found either between or within the treatment class and the comparison class with respect to the quality of claim, data, and rebuttal data across tests. Possible explanations concerning the findings and limitations of the study were discussed.

Keywords: quality of argument, small-group student talk, sociocultural theory, Chinese tertiary EFL learners, argumentative writing

INTRODUCTION

Theoretical Framework

Writing as the most challenging skill for learners in the process of learning a foreign language (Zhang, 2013; Zhao and Zhang, 2022) has been viewed as a meaning-making activity that takes place as part of a social and cultural scene (Parr et al., 2009; Parr and Wilkinson, 2016) to accomplish inherently social goals (Bazerman, 2016). From the perspective of Sociocultural Theory

(Vygotsky, 1978), the interactions and collaborations during the socially situated writing processes can scaffold the internalization of cognitive and linguistic skills and thus lead to improved writing performance (Lockhart and Ng, 1995). In this vein, the social discussions and collaborations with others scaffold the individual to construct knowledge first externally and then internally within the individual's zone of proximal development (ZPD)—the gap between the individual's actual and potential developmental levels (Vygotsky, 1978; see also Ortega, 2009, for a discussion). Besides ZPD, “collective scaffolding” (Donato, 1994) enables an expert to scaffold a novice in unequal conditions and supports a peer to scaffold one another in equal conditions (Walqui, 2006; Kibler, 2017). In this way, supportive collaborations can be established by means of speech in which the novice or the peer can “participate, and extend current skills and knowledge to higher levels of competence” (p. 40). To be specific, students in the writing classroom can collectively scaffold each other to reach a higher level of writing development in the processes of negotiating writing tasks by means of small-group interactions rather than think alone by themselves (Neumann and McDonough, 2015; Li et al., 2020). The constructs of ZPD and collective scaffolding in Sociocultural Theory solidly theorize the employment of small-group student talk for developing students' individual writing.

Small-Group Student Talk

Strijbos et al. (2004) classified three frequently used types of groups for collaborative interactions: Dyads (two members), small groups (three to six members), and large groups (seven or more). Based on the fact that there are large classes in the Chinese tertiary EFL context, it is thus both difficult for the teacher to manage too many dyads in the writing classroom and also challenging for the teacher to guarantee opportunities for each student to talk in groups of seven or more members within the limited class time. More importantly, the classroom desks and chairs in most of the Chinese universities are not movable, so groups must sit close to each other, which raises noise issues if small groups with fewer students in each group are formed and all groups discuss at the same time. Therefore, in this study, small-group student talk refers to the meaningful interactions among small groups with six students formed by the self-selection method during which students talk about writing tasks before they proceed to their individual writing (Li et al., 2020; Li and Zhang, 2021a). According to Li et al. (2020), students who are engaged in such talk shared the responsibility for managing the group as well as the ongoing process of the writing task instead of directing someone as a fixed group leader throughout the whole procedure. To this end, six students interacting in a small group offers students opportunities to assume various roles to well manage the group and the writing task.

Aligning with this definition, small-group student talk occurs during planning, the preparation for the action of writing, which is believed to be crucial and indispensable to produce written texts (Ellis and Yuan, 2004) because a planned language output will most possibly push learners to reach their potential developmental levels of language use (Mirazi and Mahmoudi, 2016). Within the Sociocultural Theory framework,

meaningful verbal interactions among small groups that occur during planning can foster writing development, knowledge co-construction, and social communication (Fernández Dobao, 2012, 2014; Li and Zhu, 2013; Li and Zhang, 2021a). Specifically, small-group student talk enables all participants to agree or disagree with one another by sharing opinions and evidence on the writing topic, which offers all participants a platform to generate and evaluate ideas and evidence and helps them select and organize those into a writing plan (Neumann and McDonough, 2015; Li and Zhang, 2021b). Additionally, small-group student talk also facilitates students to collectively scaffold one another to negotiate meaning (Storch, 2019), produce writing ideas (Watanabe and Swain, 2007; Shehadeh, 2011), solve linguistic problems in writing (Fernández Dobao, 2012, 2014), promote peer collaboration, and practice for writing development (Storch and Wigglesworth, 2007; Wigglesworth and Storch, 2009), among other things.

Considering that verbal interaction is one of the most extensively-used ways, also the most effective form for knowledge construction (Palincsar and Brown, 1989; Fernández Dobao, 2007), students in the writing classroom should be encouraged to actively talk with their peers so that those who are scaffolded by peers and engaged in small-group student talk have chances to discuss argumentative writing tasks collaboratively, bridge the gap between what each student knows and what they can know together, and thus co-construct a greater knowledge for writing development as a group than any of the group members would do on their own. In other words, when offered opportunities to engage in small-group student talk for the planning of their subsequent individual argumentative writing tasks, students can collectively scaffold each other for their ZPDs to be triggered and kept as active as possible until they accomplish the argumentative writing tasks which are comparatively more difficult when they do them alone (Antón and Dıcamilla, 1999; Ohta, 2000). In the process of doing so, collective scaffolding can be gradually reduced as the student becomes more competent in performing the task. When the student is eventually able to accomplish the argumentative task individually, the effect of the collective scaffolding is considered retained. Underpinned by such a rationale, it is important to exploit the benefits of small-group student talk while maintaining an emphasis on individual argumentative writing development.

Argumentative Writing and Quality of Argument

Argumentative writing, having been viewed as the most demanding writing task (Grabe and Kaplan, 1996; Siregar et al., 2021; Zhang, 2021; Zhang and Cheng, 2021) and a broadly recognized assessment for L2 learners' writing proficiency (Teng and Zhang, 2020), aims to convince readers to accept the writer's point of view by clearly stating a claim, selecting evidence to support such a claim, recognizing counterarguments, responding to opposing claims, and finally reaching a conclusion in a logically organized way (Nippold and Ward-Loneragan, 2010). Since to comprehend, evaluate, and construct written arguments is an essential skill for academic learning at various levels

(Wolfe et al., 2009), thus it is crucial to introduce L2 learners to argumentative writing and familiarize them with the basic terms and a comprehension of the elements of argument together with the processes through which examination of reasoning becomes the burgeoning of a claim (Heidari, 2019).

In the context of the current study, an “argument” is defined, according to Toulmin (1958), as a set of claims, one of which (the primary claim or conclusion) is meant to be supported by the rest (the reasons or premises) (Zainuddin and Rafik-Galea, 2016). Initially, the Toulmin model of argument (Toulmin, 1958) only included three elements: claim, data, and warrant. Later in Toulmin, 2003 extended his original model by adding some second-level elements: qualifier, backing, and rebuttal. Since then, a standard Toulmin model of argument has been viewed as the inclusion of six elements: claim (an arguable assertion), data (the evidence to support and justify the claim), warrant (the reasoning that connects data and claim), backing (the assumptions that reinforce the warrant), qualifier (the strength conferred by the warrant), and rebuttal (the conditions in which the claim does not hold true). The overall quality of argumentative writing is more related to the quality of argument than the mere presence of the argument elements (Paek and Kang, 2017; Özdemir, 2018; Hamam, 2020; Osman and Januin, 2021). Specifically, a good argumentative writing is not the one that only focuses on “the surface structure, or the shell of the argument” (Stapleton and Wu, 2015, p. 12) or that overemphasizes on structural elements of argument at the expense of quality of logic and evidence (Macagno and Konstantinidou, 2013; Abdollahzadeh et al., 2017). Instead, it is the one with a clear claim backed up by relevant and sufficient data as well as counterarguments challenged by effective rebuttals (Qin, 2020). In other words, a good argumentative writing integrates both the surface structure (i.e., essential elements including claims, counterclaims, rebuttals, and their accompanied data) and the substance (i.e., the quality of reasoning) (Ho, 2011; Uysal, 2012; Stapleton and Wu, 2015; Qin, 2020; Siregar et al., 2021; Sundari and Febriyanti, 2021; Zhang et al., 2021). As an essential marker of critical thinking and a significant feature of argumentative writing (Wolfe, 2011; Hirvela, 2017), quality of argument has been measured with an increasing use of the Toulmin model that has gained broad acknowledgment in accounting for the various elements marking the progress of an argument (Qin, 2020) and has been widely used to assess the quality of argumentative writing (Yeh, 1998; Qin and Karabacak, 2010; Liu and Stapleton, 2014; Stapleton and Wu, 2015; Qin, 2020; Zhang and Zhang, 2021a,b).

Effects of Small-Group Student Talk on L2 Argumentative Writing

In the recent past, research evidence in L2 documenting small-group student talk has reported mixed findings on its effects on L2 argumentative writing. One related line of research claimed positive effects by measuring linguistic features (i.e., complexity, accuracy, fluency, subordination) and/or analytic ratings (e.g., content, organization, language use, grammar, mechanics, etc.) (Shin, 2008; Pu, 2010; Neumann and McDonough, 2015; McDonough et al., 2019; Li and Zhang, 2021b). For example,

Shin (2008) investigated such effects through the comparison of individual and pair-group collaborative planning of English learners in a Korean university which claimed that the planning with pair-group discussions achieved significantly higher scores than those without it in the individual planning group on all five analytic measures (content, organization, language in use, grammar, and mechanics). In a similar vein, Pu (2010) investigated the effects of prewriting discussions among groups with three students by taking the factor of different languages into consideration. He assigned 24 first-year Chinese English major students into four groups, namely, the Chinese L1 group, English L2 group, Chinese L1 and English L2 group, and the individual planning group, and measured the language quality of students' argumentative texts written under these four conditions in terms of CAF (complexity, accuracy, and fluency). His study corroborated the effectiveness of small-group student talk given that the argumentative essays written by the English L2 group were much better in language quality with fewer errors and higher syntactical complexity than those of the other three groups. Likewise, McDonough et al. (2019), after exploring the effects of pair-group prewriting discussions in a Thai university, pointed out that although there was no significant difference for complexity measures (coordination and subordination), the texts written by students engaged in such discussions were more accurate and received higher ratings (content, organization, grammar, and vocabulary) than those written during individual planning. The findings regarding analytic ratings in these two studies were further supported by Li and Zhang (2021b), who probed the effects of the structured small-group student talk on Chinese university EFL students' individual writing by measuring the holistic and analytic ratings of content, organization, vocabulary, language use, and mechanics in students' written texts.

However, another line of related research has presented inconsistent findings. For instance, Shi (1998) examined adult international students' opinion essays written independently under the conditions of peer-led pair-group talk, teacher-led pair-group talk, and no talk prior to writing and found that students wrote longer texts in the condition of no talk, shorter texts after teacher-led pair-group talk, and texts with a greater variety of verbs after peer-led pair-group talk. Her study adopted the Hamp-Lyons 9-point band scale (Hamp-Lyons, 1991), which consists of a global scale for general scores and a profile scale to measure communicative quality, organization, argumentation, linguistic accuracy, and linguistic appropriacy. Using this two-part scale for assessment, she concluded that talking versus no talking prior to writing had no noticeable effect on the scores of students' individual writing. Another more recent study (Mirazi and Mahmoudi, 2016) adopted a quasi-experimental design to explore the effects of planning type (pair-group collaborative planning vs. individual planning) and gender on EFL learners' writing quality by measuring the five components of content, organization, vocabulary, language use, and mechanics. Their study revealed that gender had no impact on planning type in relation to learners' writing ability. Moreover, conflicting with Shi's (1998) results, this study advocated that the individual planning groups had outperformed the collaborative

planning groups in regard to students' overall writing ability. Similarly, McDonough et al. (2018) examined Thai students' English texts written under conditions of collaborative writing (writing together with a partner), collaborative prewriting (discuss with a partner during planning but write individually), and no collaboration. The study concluded that collaborative prewriting did not lead to any differences in accuracy as compared to no collaboration texts and no significant differences were found in students' texts written under these three conditions regarding the analytic ratings (content, organization, and language). In addition, collaborative texts were more accurate than the collaborative prewriting and no collaboration texts, while the latter two contained more subordination. In another study, McDonough and De Vleeschauwer (2019), who compared the effects of collaborative planning (discuss with a partner during planning but write individually) and individual prewriting (plan alone by oneself) on Thai EFL learners' writing development, revealed that individual planning resulted in higher analytic ratings (content, organization, and vocabulary), while collaborative planning led to improved accuracy, and no significant differences in coordination or subordination was found between these two types of planning. Their findings further evidenced that using small-group student talk for planning may not have any advantages for overall writing quality as compared to individual planning.

In summary, recent related studies addressing the effects of small-group student talk used during the planning of L2 argumentative writing have reported inconsistent or mixed findings. Besides, no studies have specifically investigated whether such talk has sustained effects on L2 individual argumentative writing. In addition, except Qin (2020) who particularly assessed the quality of argument of United Arab Emirates university students' individual argumentative writing using Stapleton and Wu's (2015) rubric, the majority of prior studies mainly dwelled on the measurement of linguistic features (e.g., complexity, accuracy, fluency, subordination) and/or analytic ratings (e.g., content, organization, language use, grammar, mechanics, etc.) of L2 argumentative texts, which neglected to evaluate the quality of argument—an essential marker of critical thinking and a powerful predictor of good argumentative writing (Wolfe, 2011; Ong and Zhang, 2013; Hirvela, 2017; Paek and Kang, 2017; Huang and Zhang, 2020). According to Stapleton and Wu (2015) and Qin (2020), the quality of argument is meant to be measured from both the presence of structural elements (i.e., essential elements including claims, counterclaims, rebuttals, and their accompanied data) and the quality of substance (i.e., the quality of reasoning). Evaluating the quality of argument in L2 argumentative writing is important not only because good quality of argument is considered significant for effective communication (Nussbaum and Schraw, 2007), but also because, as Qin and Karabacak (2010) argued, doing so can inform the design of instructional materials and the planning of classroom activities for L2 argumentative writing instruction. More importantly, Qin (2020) has further pointed out that it is insufficient to just instruct students to include these structural elements in their writing. In fact, more focus should be put on facilitating students to understand the

different quality of these elements so that they can obtain the ability to evaluate various propositions and thus to develop their own arguments.

To resolve these uncovered issues and better understand using small-group student talk for the planning of L2 argumentative writing, this study aimed to adopt a quasi-experimental design (Creswell, 2014) implementing a pretest, a posttest, and a delayed posttest to measure the effects of such talk on the quality of argument in Chinese tertiary EFL students' individual argumentative writing. As such, this study set out to answer the following two questions:

- (1) Does small-group student talk enhance the quality of argument in Chinese tertiary EFL students' argumentative writing regarding analytic (i.e., claim, data, counterargument claim, counterargument data, rebuttal claim, rebuttal data) and holistic scores?
- (2) Is there any difference in the effect of planning with small-group student talk and that without it on students' claim, data, counterargument claim, counterargument data, rebuttal claim, rebuttal data, and overall scores?

MATERIALS AND METHODS

Participants and Context

The participants were 48 undergraduate students in their second year of study who were conveniently sampled (Creswell, 2014) from a School of Foreign Languages at a large public comprehensive university in Central China. All these students were admitted into the English Language and Literature Program as intermediate-level language learners by the School of Foreign Languages of the sampled university and had finished the same university coursework before their admissions into two intact classes of the same required *English Writing* course. These two classes were randomly assigned into a treatment class and a comparison class with 24 students in each. The treatment class had 19 female students and 5 male students between the ages of 18 and 20 ($M = 19.3$, $SD = 0.76$). The comparison class had 16 female students and 8 male students between the ages of 18 and 21 ($M = 19.5$, $SD = 0.93$).

The compulsory *English Writing* course spanned two semesters in the sampled university, which was designed to develop students' competence in English writing of narrative, expository and argumentative essays and enhance their critical thinking ability (e.g., Huang and Zhang, 2020; Rahimi and Zhang, 2021). The course was instructed by an associate professor of English language and literature, who had 5 years of experience in EFL writing instruction. In the first semester, the course targeted to instruct students to write paragraphs as well as narrative and expository essays. In the second semester, its focus moved to teach students to write argumentative essays and train them to think logically and rationally about what they would be discussing and writing. The current study was conducted in the second semester for its purpose was to measure the quality of argumentation in Chinese tertiary EFL students' individual argumentative writing.

Materials and Design

Writing Tasks

The current study consisted of six structured writing task handouts (see Appendix A). All these writing tasks were chosen from the battery of China's National English As a Foreign Language Test—Test for English Majors – Band 4 (TEM-4). Specifically, the writing tasks for pre-, post-, and delayed post-tests remained the same throughout the quasi-experiment, which is about whether the development of intelligent machines will make human brains lazy (2017 TEM-4). The writing tasks for interventions 1–5 are respectively about whether college students should hire helpers to clean their dormitories (2010 TEM-4), whether private car owners should be taxed for environmental pollution (2011 TEM-4), whether English major students should study Mathematics (2014 TEM-4), whether tourism will bring harm to the environment (2009 TEM-4), and whether it is wise to make friends online (2007 TEM-4). TEM-4 is a nationally standardized test that is taken annually by Chinese tertiary English major students at the second semester of their second year, which guaranteed the high validity and reliability of the writing tasks (Li and Zhang, 2021b). This test is designed particularly to measure whether Chinese tertiary English major students meet the required proficiency levels of English language as specified in the National College English Teaching Syllabus for English Majors (NACFLT, 2004a; cited in Jin and Fan, 2011). It contains six parts, including dictation, listening comprehension, language usage, cloze, reading comprehension, and writing. The writing section is vital for it comprises 20% of the total score of 100. Selecting the argumentative writing tasks from the TEM-4 battery provided students with opportunities to prepare for the test, which is good for arousing students' interest and boosting their enthusiasm to participate in this study. Previous studies (Neumann and McDonough, 2015; McDonough et al., 2019; Li et al., 2020; Li and Zhang, 2021a) have proved that, compared with the naturally occurring peer group talk, structured writing tasks were more beneficial and effective for engaging L2 learners in critical evaluation of the ideas they generated and the organization they planned to make and thus helped develop L2 learners' argumentative writing. In this sense, an extra part, modified from Neumann and McDonough (2015), was incorporated in the writing task handouts. It included three parts, respectively, stating a viewpoint to show agreement or disagreement with the writing topic, producing and evaluating ideas and evidence that support and oppose the stated viewpoint, and selecting and organizing ideas and evidence into a writing plan.

Writing Rubric

To measure the quality of argument in students' argumentative writing, Stapleton and Wu's (2015) *Analytic Scoring Rubric for Argumentative Writing* (see Appendix B) was used. This rubric was modified from the original Toulmin model of argument structure and several other researchers' work (Nussbaum et al., 2005; Nussbaum and Schraw, 2007; Qin and Karabacak, 2010; Zhang and Zhang, 2021b) in an attempt to integrate "the assessment of both argumentative structural elements and the quality of reasoning, one of the few in the field of

L2 argumentative writing assessment" (Qin, 2020, p. 230). It was specifically constructed to measure the quality of argument from both structural and substance levels because "for an argumentative essay to be persuasive, not only must it follow surface structure by including alternative viewpoints and showing their weaknesses, but it must also support claims with good quality reasons that convince others" (Stapleton and Wu, 2015, p. 22).

The rubric uses the surface structure based on Toulmin-like elements as the organizing principle for the rows. Meanwhile, it also includes the quality of the supporting reasons or evidence as the organizing principle with embedded descriptors in the columns. Altogether, there are six elements in the rows, respectively, claim (assertion in response to a writing topic), data (evidence to support claims), counterargument claim (opposing assertion that opposes the writer's main claim), counterargument data (evidence to support the counterargument claim), rebuttal claim (assertion to refute the counterargument), and rebuttal data (evidence to support rebuttal claim). Scoring for these six elements breaks down from a 100-point scale. To assess the higher level of critical thinking and argumentation skills that enable the generation of counterargument claim and rebuttal claim, increased scores are given to these elements. Specifically, they are differentially weighted from a scale of 0–5 for claim and a scale of 0–10 for categories of counterargument claim and rebuttal claim, together with a scale of 0, 10, 15, 20, and 25 for categories of data, counterargument data, and rebuttal data (see Appendix B).

There were two reasons for employing this rubric to measure quality of argument. To start with, besides the assessment of argumentative structural elements, this rubric also enables the evaluation of substance (i.e., quality of reasoning), which is considered paramount because the analysis of structural elements without considering the strength of evidence used is not enough to assess the quality of argument (Paek and Kang, 2017; Zhang, 2018; Qin, 2020). Furthermore, since such a rubric allows graded evaluation and provides descriptors that are as inclusive as possible of the main features of good argumentative writing (Stapleton and Wu, 2015), it is thus convenient for researchers to facilitate the process of scoring.

Research Design

As frequently used in educational contexts which is constructed from situations that already exist in the real world (Campbell and Stanley, 1963), a quasi-experimental design (Creswell, 2014) with a pretest, a posttest, and a delayed posttest (each in 40 min) was adopted to compare the texts written independently by Chinese tertiary EFL students that were already enrolled in two parallel classes which were later randomly assigned by the first researcher into a treatment class and a comparison class.

Prior research (Shi, 1998; Neumann and McDonough, 2015; McDonough and De Vleeschauwer, 2019; Li et al., 2020; Li and Zhang, 2021a) which administered 20 min for the small-group student talk and 40 min for individual writing claimed that students could produce sufficient talk in group discussions in 20 min and drafted texts of reasonable length within 40 min. Therefore, the current study decided to administer 20 min for

the prewriting small-group student talk and 40 min for the subsequent individual writing. Specifically, before proceeding to their 40 min of individual writing, students in the treatment class followed collaborative planning with 20 min of small-group student talk, while those in the comparison class planned individually for 20 min without such talk. The independent variable of the study was small-group student talk, which was operationalized in terms of the 20 min discussion of structured writing tasks in small groups of six students in the treatment class. Due to the benefits of the self-selection method which effectively facilitates participants in group collaboration with a higher sense of goal commitment and group accomplishment (Chapman et al., 2006), self-selected small groups of six students were formed in the treatment class. The dependent variable was quality of argument, which was operationalized in terms of students' argumentative texts and measured using Stapleton and Wu's (2015) rubric.

Intervention

During the quasi-experimental study, both the treatment class and the comparison class met the aforementioned instructor who mainly employed a genre approach for the *English Writing* course in a regular university classroom setting. Both classes had two 45-min class periods in each week with a total of 32 periods in a 16-week semester. Following the same teaching syllabus and plan required by the School of Foreign Languages of the selected university, both classes used a theme-based textbook¹ that aimed to teach students how to write argumentative essays and develop their critical thinking abilities.

Students in both classes participated in the pretest, posttest, and delayed posttest respectively before, at the end of, and 4 weeks after the intervention sessions. Following a practice session in Week 2, five intervention sessions were successively carried out in the treatment class every other week from Week 3 to Week 11. Each intervention was administered with 20 min prewriting small-group student talk and 40 min subsequent individual writing of the tasks chosen from the TEM-4 battery and structured by Neumann and McDonough's (2015) additional section that intends to promote a production of arguments and evidence. Meanwhile, in the comparison class, students were asked to plan alone for 20 min and then proceed to their 40 minutes' individual writing of the same tasks mentioned above. All these three tests and intervention sessions were mainly conducted by the first researcher of this study. In order to obtain students' real EFL writing performance, no external resources or help were allowed in each test. Apart from that, the writing task handouts and time and procedures of each test were kept constant in both classes.

In both the treatment and comparison classes, the instructor taught the same contents and maintained the same schedules throughout the semester. She did not intervene in students'

discussions during the intervention sessions of the treatment class. Nor did she interrupt students' individual planning in the comparison class. Instead, she offered her assistance only when students specifically asked for it.

Data Collection Procedures

A pretest was administered to both classes in the first week. In the following week, students in the treatment class self-selected group members and formed into four groups of six students. To avoid the impact caused by changing group members, each group was told to remain unchanged during the data collection period. After that, a practice session was carried out to help participants familiarize with the processes. Altogether, five rounds of small-group student talk were administered as intervention sessions and were recorded by the first researcher with the help of the course instructor (see **Table 1**). The writing task for the practice session was chosen from the course textbook. The other five writing tasks for the intervention sessions were selected at random from the TEM-4 test battery. Each writing task was structured with an added section of the three requirements mentioned above.

During each round of intervention sessions, students in the treatment class first talked for 20 min about the structured writing task handout and then separated to write a drafted text individually for 40 min. However, students in the comparison class first conducted individual planning for 20 min following the same handout and then proceeded to their individual writing of the task for 40 min. No external sources or help were allowed in either class during each round of intervention. A posttest was given after the fifth round of intervention sessions and a delayed posttest was conducted four weeks after the posttest.

Data Analysis

In sum, 144 ($24 \times 2 \times 3$) drafted texts written by students in both classes were collected from the pretest, posttest and delayed posttest in order to determine what effects small-group student talk had on the quality of argumentation in Chinese tertiary EFL students' individual argumentative writing.

All the draft texts were rated by two Chinese tertiary EFL instructors with Ph. D. degrees in second language acquisition/applied linguistics, who had no direct involvement in this study. A blind scoring was administered so that both raters had no idea which class texts they were scoring, nor did they know if they were scoring a pre-, post-, or a delayed post-test text. To examine rating consistency and reliability, about 33% of the total texts (48/144) was randomly chosen and scored by the raters. The final score of each written text was the aggregated average value of the scores given by the two raters. Independent rating of the texts resulted in satisfactory reliability with the intraclass correlation being 0.953 for the holistic scores, which could be considered acceptable because it was larger than 0.70 (Multon, 2010). As for the analytic scores in terms of the six elements, the interrater reliability for each element was also adequate (claims, $r = 0.817$; data, $r = 0.876$; counterargument claims, $r = 0.964$; counterargument data, $r = 0.985$; rebuttal claims, $r = 0.804$; rebuttal data, $r = 0.975$).

¹The textbook is named *Writing Critically III—Argumentative Writing*, which has been included in the list of officially approved textbooks for use in China's National Standard Textbooks for English Major Students in Tertiary Institutions. The theme-based textbook is specifically designed for English-major undergraduates with orientations in English language and literature, translation, and business English.

TABLE 1 | Procedures of the study.

Week	Treatment class	Comparison class	Writing topic
1	Pretest (40 min)	Pretest (40 min)	Whether human brains will get lazy with intelligent machines to do the thinking (2017 TEM-4)
2	Practice session: Small-group student talk for planning (20 min) + individual writing (40 min)	Individual planning (20 min) + individual writing (40 min)	a) Whether college students should be allowed to take cell phones into classrooms (from the course textbook)
3, 5, 7, 9, 11	Intervention sessions: Small-group student talk for planning (20 min) + individual writing (40 min) Posttest	Individual planning (20 min) + individual writing (40 min)	b) Whether college students should hire helpers to clean their dormitories (2010 TEM-4) c) Whether private car owners should be taxed for environmental pollution (2011 TEM-4) d) Whether English major students should study Mathematics (2014 TEM-4) e) Whether tourism will bring harm to the environment (2009 TEM-4)
12		Posttest	Whether it is wise to make friends online (2007 TEM-4)
16	Delayed posttest	Delayed posttest	Whether human brains will get lazy with intelligent machines to do the thinking (2017 TEM-4)

Statistical analyses using SPSS 26.0 were carried out to address the two research questions. The Shapiro–Wilk tests of normality were run before the analysis to check normality, missing values, and outliers. Results of the Shapiro–Wilk tests revealed that all the data of the current study was normally distributed since the *z*-scores of skewness and kurtosis did not exceed 1.96 (Field, 2009). After that, independent-samples *t*-tests were administered to explore the between-subject differences and see whether there existed any effects of small-group student talk on the quality of argument in Chinese tertiary EFL students' individual argumentative writing between the treatment class and the comparison class. In the following, one-way repeated measures ANOVAs were carried out to further examine within-subjects differences in each class. Finally, paired samples *t*-tests with a Bonferroni correction which was used to avoid Type I errors would be run if significant changes were perceived from the one-way repeated measures ANOVAs to investigate whether the treatment class significantly showed a larger effect size than the comparison class. During such comparisons, the effect sizes were interpreted using the Cohen's (1992) criteria which deem that *d* values of 0.20, 0.50, and 0.80 and partial η^2 values of 0.01, 0.06, and 0.14 are respectively considered as small, medium, and large effect sizes.

RESULTS

Effects on Overall Quality of Argument

The following Table 2 demonstrates the descriptive data for the subscores of the quality of each argumentative element together with the overall scores of the quality of these elements in Chinese tertiary EFL students' individual argumentative writing between the treatment class and the comparison class across the three tests. To make sure the baseline conditions of the two classes at the outset of the study, independent samples *t*-tests were applied.

The between-subjects results (see Table 3) suggested that students' performance in holistic and analytic scores of the quality of argument were similar at the time of the pretest (overall, $p = 0.953$; claim, $p = 0.331$; data, $p = 0.607$; counterargument claim, $p = 0.848$; counterargument data, $p = 0.820$; rebuttal claim, $p = 0.306$; rebuttal data, $p = 0.882$). However, significant differences with large effect sizes were found between the treatment class and the comparison class concerning the overall quality of argument in the immediate posttest ($t = -4.096$, $p < 0.001$, $d = -1.18$) and the delayed posttest ($t = -3.800$, $p < 0.001$, $d = -1.10$). Such results indicated that small-group student talk used as collaborative prewriting discussions enabled students in the treatment class to gain higher scores of the quality of argument compared with those in the comparison class.

The application of one-way repeated measures ANOVAs showed that the scores of the overall quality of argumentation changed differently over time in both the treatment class [$F(2,46) = 71.147$, $p < 0.001$, $\eta_p^2 = 0.756$] and the comparison class [$F(2,46) = 5.972$, $p = 0.005$, $\eta_p^2 = 0.206$]. Paired samples *t*-tests and Bonferroni correction ($p = 0.017$) were further employed to better examine the within-subjects differences in each class. Discernible improvement with large effect sizes in the treatment class was observed across the tests (pretest vs. posttest, $p < 0.001$, $d = -2.23$; pretest vs. delayed posttest, $p < 0.001$, $d = -1.71$) and the effect was retained in the delayed posttest (post vs. delayed posttest, $p = 0.043$).

In contrast, statistically significant differences in the comparison class only appeared from the pretest to the immediate posttest ($p = 0.001$, $d = -0.76$). Such differences did not manifest neither from the immediate posttest to the delayed posttest ($p = 0.148$) nor from the pretest to the delayed posttest ($p = 0.086$). The results of these comparisons suggested that small-group student talk was significantly effective with respect to the overall quality of argument in Chinese tertiary EFL students' individual argumentative writing.

TABLE 2 | Descriptive statistics for holistic and analytic scores of the quality of argument across tests.

Measures	Class	Pretest		Posttest		Delayed posttest	
		Mean	SD	Mean	SD	Mean	SD
Overall	CC	61.25	5.33	63.54	3.70	62.71	3.67
	TC	61.17	4.33	68.00	3.85	66.88	3.93
Claim	CC	4.29	0.690	4.38	0.647	4.29	0.690
	TC	4.08	0.776	4.25	0.676	4.33	0.702
Data	CC	22.04	2.27	22.42	2.02	22.17	1.93
	TC	22.38	2.18	22.67	2.32	22.46	2.21
Counterargument claim	CC	7.33	1.55	8.33	1.05	7.79	1.50
	TC	7.42	1.44	8.96	0.91	8.71	1.33
Counterargument data	CC	16.63	2.45	16.88	2.03	17.08	2.08
	TC	16.79	2.60	18.96	3.03	18.63	2.80
Rebuttal claim	CC	5.88	2.58	6.04	1.49	5.92	2.39
	TC	5.29	1.00	6.88	1.19	5.67	1.34
Rebuttal data	CC	5.08	3.11	5.54	2.21	5.42	1.77
	TC	5.21	2.69	6.21	1.53	6.04	1.37

CC, comparison class; TC, treatment class; SD, standard deviation.

TABLE 3 | Between-subject comparisons of holistic and analytic scores on quality of argument across tests.

Measures	Pretest		Posttest		Delayed posttest	
	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
Overall	0.059	0.953	−4.096	0.000**	−3.800	0.000**
Claim	0.983	0.331	0.655	0.516	−0.207	0.837
Data	−0.518	0.607	−0.399	0.692	−0.488	0.628
Counterargument claim	−0.193	0.848	−2.206	0.032*	−2.234	0.030*
Counterargument data	−0.229	0.820	−2.800	0.007*	−2.167	0.035*
Rebuttal claim	1.034	0.306	−2.142	0.038*	−1.339	0.187
Rebuttal data	−0.149	0.882	−1.216	0.230	−1.371	0.177

* $p < 0.05$; ** $p < 0.001$.

Effects on Claim

No statistically significant differences concerning the quality of claim were found between the treatment class and the comparison class, including the pretest ($t = 0.983$, $p = 0.331$), the immediate posttest ($t = 0.655$, $p = 0.516$), and the delayed posttest ($t = -0.207$, $p = 0.837$) (see **Table 3**). The running of one-way repeated measures ANOVAs revealed that students in each class made no significant improvement across time [$F(2,46) = 1.856$, $p = 0.168$, $\eta_p^2 = 0.075$], and [$F(2,46) = 0.069$, $p = 0.934$, $\eta_p^2 = 0.003$ respectively]. Such results indicated that small-group student talk had no significant effects on the quality of claim in students' individual writing.

Effects on Data

Between-subjects comparisons suggested that students in the treatment class achieved a similar performance to those in the comparison class across the three tests ($t = -0.518$, $p = 0.607$; $t = 0.655$, $p = 0.516$; $t = -0.207$, $p = 0.837$ respectively) (see **Table 3**). The results of one-way repeated measures ANOVAs

indicated that the quality of data did not vary significantly across the tests, neither in the treatment class [$F(2,46) = 1.417$, $p = 0.253$, $\eta_p^2 = 0.058$], nor in the comparison class [$F(2,46) = 1.274$, $p = 0.289$, $\eta_p^2 = 0.053$]. In other words, students in both classes did not perform significantly differently over time concerning the quality of data.

Effects on Counterargument Claim

Concerning the quality of counterargument claim, students in both classes performed similarly in the pretest ($t = -0.193$, $p = 0.848$). However, the results of between-subjects comparisons demonstrated that students in the treatment class achieved better performance than those in the comparison class in the posttest immediately after the treatment ($t = -2.206$, $p = 0.032$, $d = -0.64$) and in the delayed posttest 4 weeks after the treatment ($t = -2.234$, $p = 0.030$, $d = -0.64$) (see **Table 3**).

A further analysis using one-way repeated measures ANOVAs showed that the scores of the quality of counterargument claim changed significantly over time in both the treatment class [$F(2,46) = 21.109$, $p < 0.001$, $\eta_p^2 = 0.479$] and the comparison class [$F(2,46) = 5.536$, $p = 0.007$, $\eta_p^2 = 0.194$]. Within-subjects analysis using the paired samples *t*-tests with Bonferroni correction ($p = 0.017$) indicated a significant effect with a larger size from the pretest to the immediate posttest in the treatment class ($p < 0.001$, $d = -1.40$) compared with that in the comparison class ($p < 0.001$, $d = -0.85$). Such a large size effect was maintained in the delayed posttest in the treatment class ($p < 0.001$, $d = -0.90$). However, no significant improvement manifested in the comparison class, neither from the immediate posttest to the delayed posttest ($p = 0.073$), nor from the pretest to the delayed posttest ($p = 0.217$). Such results suggested that small-group student talk promoted students in the treatment class to produce better quality of counterargument claim in their individual writing across the tests.

Effects on Counterargument Data

Table 3 demonstrates that no significant differences were found between the treatment class and the comparison class in the quality of counterargument data in the pretest ($t = -0.229$, $p = 0.820$). In contrast, students in the treatment class achieved a significantly better performance than those in the comparison class in the immediate posttest ($t = -2.800$, $p = 0.007$, $d = -0.81$) and the delayed posttest ($t = -2.167$, $p = 0.035$, $d = -0.63$).

To analyze whether the quality of counterargument data changed significantly within each class in the immediate posttest and the delayed posttest, one-way repeated measures ANOVAs were used. Results showed that significant changes across tests were observed in the treatment class [$F(2,46) = 66.294$, $p = 0.000$, $\eta_p^2 = 0.742$], while it was not seen in the comparison class [$F(2,46) = 2.066$, $p = 0.138$, $\eta_p^2 = 0.082$]. Paired samples t -tests with Bonferroni correction ($p = 0.017$) were then applied to further explore the within-subjects differences in the treatment class. The results indicated that small-group student talk enabled students to make progress in the quality of counterargument data with a large effect size from the pretest to the immediate posttest ($p < 0.001$, $d = -2.07$). No significant improvement was found from the immediate posttest to the delayed posttest ($p = 0.069$) and the large size effect of small-group student talk was retained in the delayed posttest ($p < 0.001$, $d = -1.68$). Results of the between-subjects and within-subjects comparisons revealed that small-group student talk was discernibly effective in facilitating Chinese tertiary EFL students to improve the quality of counterargument data in their individual argumentative writing across time.

Effects on Rebuttal Claim

The treatment and comparison classes achieved similar performance regarding the quality of rebuttal claim in the pretest ($t = 1.034$, $p = 0.306$) and the delayed posttest ($t = -1.339$, $p = 0.187$). However, the treatment class outperformed the comparison class in this measure in the immediate posttest ($t = -2.142$, $p = 0.038$, $d = -0.62$) (see **Table 3**). The within-subjects analysis using one-way repeated measures ANOVAs showed that the quality of rebuttal claim varied significantly over time in the treatment class [$F(2,46) = 15.979$, $p < 0.001$, $\eta_p^2 = 0.410$], but it was not a case in the comparison class [$F(2,46) = 0.129$, $p = 0.879$, $\eta_p^2 = 0.006$].

A series of paired samples t -tests with Bonferroni correction ($p = 0.017$) were run to further examine the changes across tests in the treatment class. The results indicated that significant differences with a large effect size were discerned in the treatment class from the pretest to the posttest ($p < 0.001$, $d = -1.12$). Such a large size effect of small-group student talk on the quality of rebuttal claim was also seen from the pretest to the delayed posttest ($p = 0.001$, $d = -0.81$). These results suggested that small-group student talk effectively promoted students in the treatment class to produce better quality of rebuttal claim in students' individual writing and such effects could be sustained in the delayed posttest.

Effects on Rebuttal Data

There were no statistically significant differences with respect to the quality of rebuttal data between the treatment class and the comparison class across tests (pretest: $t = -0.149$, $p = 0.882$; immediate posttest: $t = -1.216$, $p = 0.230$; delayed posttest: $t = -1.371$, $p = 0.177$ respectively) (see **Table 3**). The application of one-way repeated measures ANOVAs indicated that students made significant improvement in the treatment class [$F(2,46) = 3.436$, $p = 0.041$, $\eta_p^2 = 0.130$], but not in the comparison class [$F(2,46) = 1.163$, $p = 0.322$, $\eta_p^2 = 0.048$]. However, a further within-subjects analysis of such effects in the treatment class using the paired samples t -tests with Bonferroni correction ($p = 0.017$) revealed no significant changes across the three tests over time (pretest vs. posttest, $p = 0.032$; posttest vs. delayed posttest, $p = 0.583$; pretest vs. delayed posttest, $p = 0.089$ respectively). In this sense, students in the treatment class did not achieve significantly better performance than those in the comparison class across the three tests. These results suggested that small-group student talk had no significant effects on the quality of rebuttal data in Chinese tertiary EFL students' individual argumentative writing.

DISCUSSION

In answer to the first research question, which inquired into whether small-group student talk helps enhance the quality of argument in Chinese tertiary EFL students' argumentative writing, the overall quality of argument in students' argumentative writing during the pre-, post-, and delayed post-tests was assessed following a holistic scoring rubric developed and validated by Stapleton and Wu (2015). This rubric drew upon Nussbaum et al. (2005); Nussbaum and Schraw (2007), and Qin and Karabacak (2010), which included six elements of argument (claim, data, counterargument claim, counterargument data, rebuttal claim, rebuttal data). The significant distinctions of the holistic scores in both the immediate posttest and the delayed posttest showed that small-group student talk enabled students in the treatment class to gain higher scores of the overall quality of argument compared with those in the comparison class. In other words, small-group student talk did exert positive effects on facilitating the quality of argument through which Chinese tertiary EFL students could improve their argumentative writing performance and enhance their critical thinking skills. Such a finding lends support to previous studies (Nussbaum et al., 2005; Shin, 2008; Pu, 2010; Neumann and McDonough, 2015; McDonough et al., 2019; Li et al., 2020; Li and Zhang, 2021a) in that talking prior to writing has impact on students' written texts and quality of argument in students' writing could be effectively facilitated using talking as a scaffolding tool, because talk created opportunities for students to scaffold within each other's ZPDs (Neumann and McDonough, 2015; McDonough et al., 2019; Li et al., 2020), co-construct their knowledge and experience (Shin, 2008; Pu, 2010; Li and Zhang, 2021b), and draw from ideas and practices they learn with their peers (Nussbaum et al., 2005; Olsen and VanDerHeide, 2020).

Nevertheless, such results have challenged Shi's (1998) study which claimed that prewriting small-group discussions had no immediate influence on the writing scores and no noticeable effects of pair-group (both teacher-led and peer-led) talk on students' individual writing were perceived. The possible reason might be the different writing rubrics the two studies adopted. The current study mainly focused on the measurement of quality of argument (i.e., presence of the Toulmin-based argument elements and quality of reasoning), while Shi's (1998) study conducted a more comprehensive measurement of organization, linguistic, communicative, and argumentative aspects. Using different writing rubrics might potentially lead to inconsistent results when judging the quality of argumentative writing (Plakans and Gebriel, 2017). The influence of such a factor is also found in the discordant results between this study and McDonough and De Vleeschauwer's (2019). Besides, this study has also reported conflicting findings with Mirazi and Mahmoudi's (2016) which confirmed that individual planning outperformed pair-group collaborative planning in terms of Iranian students' overall writing ability measured by content, organization, vocabulary use, language use, and mechanics (Jacobs et al., 1981). In addition to the factor of different writing rubrics, various writing tasks and group sizes might be other factors that lead to the discordant results. The current study selected writing tasks from the writing sections of the standard Chinese TEM-4 battery which emphasizes students' critical thinking and argumentative abilities, while Mirazi and Mahmoudi's (2016) study used the ones from a TOEFL essay preparation book which concentrated more on students' reading and listening comprehension as well as summarizing and rewriting abilities. Meanwhile, group sizes (six-student small group VS. pair group) also enable a direct influence on the quality of group discussions and thus cause different results (Burgoon et al., 2002).

With respect to the second research question about whether there exists any difference in the effect of planning with small-group student talk and that without it, this study found that no discernible differences were perceived in the analytic scores either between or within the treatment class and the comparison class with respect to the quality of claim, data, and rebuttal data across tests. However, the statistical analyses did reveal immediate and sustained effects of small-group student talk on the quality of counterargument claim, counterargument data, and rebuttal claim. Counterargument claim was the only element in which students in both classes made significant improvement, but the treatment class demonstrated a larger effect size. In other words, small-group student talk significantly facilitates the quality of argument in students' argumentative writing by promoting the quality of counterargument claim, counterargument data, and rebuttal claim, which were viewed as parts of the second-level key elements of argument (Zhang, 2018; Zhang and Zhang, 2021a), because a critical thinker ought to reflect different stances and weigh the pros and cons of each stance (Qin, 2020). Also, advanced arguments tended to implement counterarguments and rebuttals (Wolfe et al., 2009; Paek and Kang, 2017) and the presence of opposing views and counterarguments is of central importance to argumentative writing (Rusfandi, 2015).

Concerning the finding that no significant effects were found in terms of claim and data, one possible explanation might include that claim (assertion in response to a writing topic) and data (evidence to support claims) are the most fundamental and preferred elements for learners (Qin and Karabacak, 2010; Qin, 2013, 2020; Liu and Stapleton, 2014; Stapleton and Wu, 2015; Abdollahzadeh et al., 2017; Zhang, 2018). Thus, it does not matter whether students receive the intervention sessions or not. Either way, students would follow the most basic and natural way that they were already quite familiar with to present their claims and data in their individual writing. Another probable reason might be related to the moves of small-group student talk which students follow during their collaborative discussions. It is worth mentioning that students' talking moves mainly consisted of three steps as suggested in the structured writing tasks. They first expressed each other's viewpoints of the argumentative writing topic along with corresponding supporting reasons. After that, they began to argue with one another to defend and justify their viewpoints. Finally, they negotiated to decide which ideas and evidence to select and organize into their writing plan. It is obvious that their claims and data were mainly presented during the first move, in which students did not challenge or argue with each other but took turns to give their claims and data until everyone finished; while the counterargument and rebuttal claims and data were largely produced during the second move. In this sense, without negotiation and arguing with each other, the first move in which students' claims and data were generated only plays a similar role as what individual planning does. This could help explain why students in the treatment class did not outperform those in the comparison class in these two measures.

As for the finding that no significant effects were discerned in terms of rebuttal data between the classes and within each class, one possible reason might be that students in both classes were not familiar with the writing topic. The mean score of rebuttal data in each class is lower than 6 (out of 25), which indicated that due to the lack of relevant topical knowledge, students in both classes barely produced any rebuttal data. As "the interaction between one's prior knowledge and the content of a specific passage" (Alexander et al., 1991, p. 334), topical knowledge affects the writing performance and shapes the texts in impromptu essay writing (He and Shi, 2012; Zhang and Zhang, 2021b). In this sense, L2 writing instructors are suggested to give students sufficient exposure to materials covering different types of topical knowledge so that students will be familiar enough with the writing topics to align counterargument and rebuttal claims with the corresponding supporting evidence and thus be able to make their viewpoints logically acceptable and effectively persuasive. Such a finding discords with previous studies that attributed the insufficient generation of counterarguments and rebuttals to cultural influences (Xu and Cao, 2012; Paek and Kang, 2017; Wei et al., 2020), because statistically significant differences were found concerning the immediate and sustained effects of small-group student talk on the quality of counterargument claim, counterargument data, and rebuttal claim. Such effects indicated that engaging students in meaning-making talk scaffolded by the Toulmin argument structure can effectively facilitate them to produce desired elements (Rusfandi, 2015).

Regarding the findings that small-group student talk had immediate and sustained effects on the quality of counterargument claim, counterargument data, and rebuttal claim, a possible explanation might be that small-group student talk enabled students to produce increasingly argumentative structural elements with more logically relevant and acceptable evidence. As the basis for developing content for writing, talk is generative and supportive for the development as well as the articulation of ideas for writing prior to the act of transforming the ideas into written text (Parr et al., 2009). In the process of such collaborative talk, students generated alternative viewpoints, provided, and evaluated reasons and clarifications, and negotiated to decide which ideas and evidence to select and organize into their writing plan. Such collaboration and negotiation led them to consider opinions in opposition to their own (or others') arguments with corresponding supporting evidence, for talk can influence the construction of knowledge around texts and topics and helps students in exploring ideas, informing the argumentative writing that students do in classrooms (Brady, 2018). Therefore, it is not surprising to see that most texts produced in the immediate posttest and the delayed posttest by students in the treatment class included more counterargument claim, counterargument data, and rebuttal claims as well as a wider variety of supporting evidence, such as expert opinions, statistics, examples, personal experiences, common sense, and logical analysis. Such two-sided argument texts offer writers and audiences better opportunities to carry out deeper negotiation, since good or effective arguments are typically expressed with multiple sides (Zhang and Zhang, 2021b).

On the contrary, although students in the comparison class were also provided with the same structured writing tasks, they failed to achieve so in their individual planning. Because without being challenged and arguing with others as their peers did in small-group student talk, they lacked opportunities to collaboratively facilitate complex understanding and co-construct knowledge through meaningful negotiations with others (Winzenried et al., 2017). Correspondingly, their texts were mainly composed of claims and data together with one counterargument claim and one or two counterargument data, which failed to present a certain number of adequate counterargument and rebuttal claims and data, even though good arguments involve counterargument and rebuttal claims and data to augment writing quality (Wolfe et al., 2009). Moreover, texts of the comparison class demonstrated illogical reasoning which are inaccurate and/or irrelevant in terms of quality (Rapanta et al., 2013; Zhang, 2018), mainly in the form of two types—personal experiences and logical analysis. However, in terms of persuasive power, anecdotal evidence is viewed as less effective than expert, causal, and statistical evidence (Hoeken and Hustinx, 2003). Thus, writers' own personal judgments and experiences are not regarded as strong evidence to support a claim (Zhang, 2018). The lack of alternative or conflicting viewpoints together with the insufficiency of logically adequate evidence indicated that these texts mainly presented one-sided argument or my-side bias argument (Felton et al., 2015). Such texts seemed structurally well-designed, but they were significantly low in terms of

argument quality, because this type of argument is regarded as the least sophisticated form of an argument (Rusfandi, 2015).

CONCLUSION

The current quasi-experimental study investigated whether employing small-group student talk as collaborative discussions for prewriting planning helped facilitate the quality of argument (measured by claim, data, counterargument claim, counterargument data, rebuttal claim, and rebuttal data) in Chinese tertiary ELF students' individual argumentative writing. Statistical results of the holistic scores in both the immediate posttest and the delayed posttest showed that small-group student talk facilitated students in the treatment class to achieve higher scores of the overall quality of argument compared with those in the comparison class, which indicated that small-group student talk was effective for promoting the quality of argument in Chinese tertiary EFL students' individual argumentative writing. Immediate and sustained effects were also found on counterargument claim, counterargument data, and rebuttal claim. Counterargument claim was the only element in which students in both classes made significant improvement, but the treatment class demonstrated a larger effect size. However, no evident effects were perceived regarding claim, data, and rebuttal data across the three tests. Such findings imply that small-group student talk enabled students to achieve collaborative planning for their individual writing, effectively facilitate themselves to produce desired Toulmin-like elements (Rusfandi, 2015), and promote the quality of counterargument claim, counterargument data, and rebuttal claim in their argumentative writing.

These findings imply that the employment of small-group student talk in the Chinese tertiary EFL learners' writing classroom is beneficial for developing students' quality of argument in their individual argumentative writing. Therefore, L2 writing instructors are encouraged to provide their students with enough opportunities to engage in such a talk during which they are able to develop sufficient Toulmin-like elements and present good quality reasoning. Besides, given that no immediate and sustained effects of small-group student talk was identified on rebuttal data due to students' unfamiliar with the writing topics, L2 writing instructors ought to exposure their students to writing materials with a wide range of topical knowledge to mitigate its influence on students' writing performance (He and Shi, 2012; Zhang and Zhang, 2021b) and help students accumulate sufficient supporting evidence and data.

Despite these findings and implications, this study has certain limitations. Firstly, due to a small sampling size of participants ($N = 48$) in this study, such findings might not be ideal for generalization. Further studies in this vein can amplify the size of sampling to magnify the reliability of the results. Secondly, a repeated writing task for all the three tests (pre-, post-, and delayed post-tests) was used, which might lead to memory issues and influence students' writing performance. Therefore, different writing tasks for the three tests can be tried in future research. Finally, this study mainly dwelled on a quantitative measurement and analysis of the argument elements and quality, which lacks

an in-depth evaluation and interpretation of the features of each element and the advancement process of reasoning. Thus, future studies are suggested to combine qualitative and quantitative analyses for a more thorough and comprehensive understanding of the quality of argument from both the aspects of structural elements and quality of reasoning.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Auckland Human Ethics Committee. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individuals for the publication of any potentially identifiable data included in this article.

REFERENCES

- Abdollahzadeh, E., Farsani, M. A., and Beikmohammadi, M. (2017). Argumentative writing behavior of graduate EFL learners. *Argumentation* 31, 641–661. doi: 10.1007/s10503-016-9415-5
- Alexander, P. A., Schallert, D. L., and Hare, V. C. (1991). Coming to terms: how researchers in learning and literacy talk about knowledge. *Rev. Educ. Res.* 61, 315–343. doi: 10.3102/00346543061003315
- Antón, M., and Dicamilla, F. J. (1999). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Modern Lang. J.* 83, 233–247. doi: 10.1111/0026-7902.00018
- Bazerman, C. (2016). “What to sociocultural studies of writing tell us about learning to write?,” in *Handbook of Writing Research*, 2nd Edn, eds C. A. MacArthur, S. Graham, and J. Fitzgerald (New York, NY: Guilford), 11–23.
- Brady, J. (2018). Investigating the relationship between classroom conversation and argumentative writing using writing moves and types of talk. *Acta Paedagogica Vilnensia* 40, 94–110. doi: 10.15388/ActPaed.2018.0.11890
- Burgoon, J. K., Bonito, J. A., Ramirez, A. Jr., Dunbar, N. E., Kam, K., and Fischer, J. (2002). Testing the interactivity principle: effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *J. Commun.* 52, 657–677. doi: 10.1111/j.1460-2466.2002.tb02567.x
- Campbell, D. T., and Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston, MA: Houghton Mifflin Company.
- Chapman, K. J., Meuter, M., Toy, D., and Write, L. (2006). Can't we pick our own group? The influence of group selection method on group dynamics and outcomes. *J. Manage. Educ.* 30, 557–569. doi: 10.1177/1052562905284872
- Cohen, J. (1992). Quantitative methods in psychology: a power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th Edn. Thousand Oaks, CA: Sage Publications.
- Donato, R. (1994). “Collective scaffolding in a second language,” in *Vygotskian Approaches to Second Language Research*, eds J. P. Lantolf and G. Appel (New Jersey, USA: Ablex Publishing Corporation), 33–56.
- Ellis, R., and Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Stud. Second Lang. Acquisition* 26, 59–84. doi: 10.1017/S0272263104261034
- Felton, M., Garcia-Mila, M., Villarreal, C., and Gilabert, S. (2015). Arguing collaboratively: argumentative discourse types and their potential for knowledge building. *Br. J. Educ. Psychol.* 85, 372–386. doi: 10.1111/bjep.12078

AUTHOR CONTRIBUTIONS

HL designed the study, collected and analyzed the data, and drafted the initial manuscript. LZ revised, proofread and finalized the draft for submission as the corresponding author. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by a research grant from the Hubei Provincial Department of Education of China under grant 2018112 and another grant from the Fundamental Research Funds for the Central Universities of China under grant 2021VI007.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.868045/full#supplementary-material>

- Fernández Dobao, A. (2007). “The Use of Circumlocution Strategies in Foreign Language Interaction: A Collaborative Creation of Meaning Process,” in *Proceedings of the 30th International AEDEAN Conference: [electronic resource]*, eds M. Losada Friend, P. Ron Vaz, S. Hernández Santano, and J. Casanova (Huelva: Universidad de Huelva).
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: comparing group, pair, and individual work. *J. Second Lang. Writing* 21, 40–58. doi: 10.1016/j.jslw.2011.12.002
- Fernández Dobao, A. (2014). Attention to form in collaborative writing tasks: comparing pair and small group interaction. *Can. Modern Lang. Rev.* 70, 158–187. doi: 10.3138/cmlr.1768
- Field, A. (2009). *Discovering Statistics using SPSS*. Thousand Oaks, CA: Sage Publications.
- Grabe, W., and Kaplan, R. (1996). *Theory and Practice of Writing*. London: Addison Wesley Longman.
- Hamam, D. (2020). A study of the rhetorical features and the argument structure of EAP essays by L1 and L2 students in the UAE. *J. Asia TEFL* 17, 699–706. doi: 10.18823/asiatell.2020.17.2.28.699
- Hamp-Lyons, L. (1991). *Assessing Second Language Writing in Academic Contexts*. Norwood, NJ: Ablex Publishing Corporation.
- He, L., and Shi, L. (2012). Topical knowledge and ESL writing. *Lang. Test.* 29, 443–464. doi: 10.1177/0265532212436659
- Heidari, F. (2019). The effect of dynamic assessment of Toulmin model through teacher and collective scaffolding on argument structure and argumentative writing achievement of Iranian EFL learners. *Iran. J. Appl. Lang. Stud.* 11, 115–140.
- Hirvela, A. (2017). Argumentation and second language writing: are we missing the boat? *J. Second Lang. Writing* 100, 69–74. doi: 10.1016/j.jslw.2017.05.002
- Ho, V. L. (2011). *Non-Native Argumentative Writing by Vietnamese Learners of English: A Contrastive Study*. (Ph.D.thesis). USA: Georgetown University.
- Hoeken, H., and Hustinx, L. (2003). “The relative persuasiveness of different types of evidence,” in *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*, eds F. H. van Eemeren, J. A. Blair, C. A. Willard, and A. F. Snoeck Henkemans (Amsterdam, The Netherlands: Sic Sat), 497–501.
- Huang, Y., and Zhang, L. J. (2020). Does a process-genre approach help improve students' argumentative writing in English as a foreign language? Findings from an intervention study. *Read. Writing Q.* 36, 339–364. doi: 10.1080/10573569.2019.1649223

- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., and Hughey, J. B. (1981). "Testing ESL Composition: A Practical Approach," in *English Composition Program*. (Rowley, MA: Newbury House Publishers).
- Jin, Y., and Fan, J. (2011). Test for English majors (TEM) in China. *Lang. Test.* 28, 589–596. doi: 10.1177/0265532211414852
- Kibler, A. (2017). Peer interaction and learning in multilingual settings from a sociocultural perspective: theoretical insights. *Int. Multiling. Res. J.* 11, 199–203. doi: 10.1080/19313152.2017.1328970
- Li, H. H., and Zhang, L. J. (2021a). Effects of structured small-group student talk as collaborative prewriting discussions on Chinese university EFL students' individual writing: a quasi-experimental study. *PLoS One* 16:e0251569. doi: 10.1371/journal.pone.0251569
- Li, M., and Zhang, M. (2021b). Collaborative writing in L2 classrooms: a research agenda. *Lang. Teach.* 1–19. doi: 10.1017/S0261444821000318
- Li, H. H., Zhang, L. J., and Parr, J. M. (2020). Small-group student talk before individual writing in tertiary English writing classrooms in China: nature and insights. *Front. Psychol.* 11:570565. doi: 10.3389/fpsyg.2020.570565
- Li, M., and Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Comput. Assist. Lang. Learn.* 26, 61–82. doi: 10.1080/09588221.2011.631142
- Liu, F., and Stapleton, P. (2014). Counterargumentation and the cultivation of critical thinking in argumentative writing: investigating washback from a high-stakes test. *System* 45, 117–128. doi: 10.1016/j.system.2014.05.005
- Lockhart, C., and Ng, P. (1995). Analyzing talk in ESL peer response groups: stances, functions, and content. *Lang. Learn.* 45, 605–651. doi: 10.1111/j.1467-1770.1995.tb00456.x
- Macagno, F., and Konstantinidou, A. (2013). What students' arguments can tell us: using argumentation schemes in science education. *Argumentation* 27, 225–243. doi: 10.1007/s10503-012-9284-5
- McDonough, K., and De Vleschauwer, J. (2019). Comparing the effect of collaborative and individual prewriting on EFL learners' writing development. *J. Second Lang. Writing* 44, 123–130. doi: 10.1016/j.jslw.2019.04.003
- McDonough, K., De Vleschauwer, J., and Crawford, W. (2018). Comparing the quality of collaborative writing, collaborative prewriting, and individual texts in a Thai EFL context. *System* 74, 109–120. doi: 10.1016/j.system.2018.02.010
- McDonough, K., De Vleschauwer, J., and Crawford, W. J. (2019). Exploring the benefits of collaborative prewriting in a Thai EFL context. *Lang. Teach. Res.* 23, 685–701. doi: 10.1177/1362168818773525
- Mirazi, S., and Mahmoudi, A. (2016). The effects of collaborative and individual planning on Iranian male and female EFL learners' writing ability. *Theory Practice Lang. Stud.* 6, 1630–1636. doi: 10.17507/tpls.0608.16
- Multon, K. D. (2010). "Interrater reliability," in *Encyclopedia of Research Design*, ed. J. S. Neil (Thousand Oaks, CA: Sage Publications), 626–628.
- NACFLT (2004a). *Syllabus for TEM4*. Shanghai: Shanghai Foreign Language Education Press.
- Neumann, H., and McDonough, K. (2015). Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing. *J. Second Lang. Writing* 27, 84–104. doi: 10.1016/j.jslw.2014.09.009
- Nippold, M. A., and Ward-Lonerger, J. M. (2010). Argumentative writing in pre-adolescents: the role of verbal reasoning. *Child Lang. Teach. Ther.* 26, 238–248. doi: 10.1177/0265659009349979
- Nussbaum, E. M., Kardash, C. M., and Graham, S. E. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *J. Educ. Psychol.* 97, 157–169. doi: 10.1037/0022-0663.97.2.157
- Nussbaum, E. M., and Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *J. Exp. Educ.* 76, 59–92. doi: 10.3200/JEXE.76.1.59-92
- Ohta, A. S. (2000). Rethinking interaction in SLA: developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. *Sociocult. Theory Second Lang. Learn.* 4, 51–78.
- Olsen, A. W., and VanDerHeide, J. (2020). Representations of students becoming as writers. *Learn. Cult. Soc. Interact.* 24:100285. doi: 10.1016/j.lcsi.2019.02.010
- Ong, J., and Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Q.* 47, 375–398. doi: 10.1002/tesq.55
- Ortega, L. (2009). "Studying writing across EFL contexts: Looking back and moving forward," in *Writing in Foreign Language Contexts*, ed. R. Manchon (Clevedon: Multilingual matters), 232–255. doi: 10.21832/9781847691859-013
- Osman, W. H., and Januin, J. (2021). Analysing ESL Persuasive Essay Writing Using Toulmin's Model of Argument. *Psychol. Educ. J.* 58, 1810–1821. doi: 10.17762/pae.v58i1.1034
- Özdemir, S. (2018). The effect of argumentative text pattern teaching on success of constituting argumentative text elements. *World J. Educ.* 8, 112–122. doi: 10.5430/wje.v8n5p112
- Paek, J. K., and Kang, Y. (2017). Investigation of content features that determine Korean EFL learners' argumentative writing qualities. *Eng. Teach.* 72, 101–122. doi: 10.15858/engtea.72.2.201706.101
- Palincsar, A. S., and Brown, A. L. (1989). Classroom dialogues to promote self-regulated comprehension. *Adv. Res. Teach.* 1, 35–71.
- Parr, J. M., Jesson, R., and McNaughton, S. (2009). "Agency and platform: The relationships between talk and writing," in *The SAGE Handbook of Writing Development*, eds R. Beard, D. Myhill, J. Riley, and M. Nystrand (London, England: Sage), 246–256. doi: 10.4135/9780857021069.n17
- Parr, J. M., and Wilkinson, I. (2016). Widening the theoretical lens on talk and writing pedagogy. *Int. J. Educ. Res.* 80, 217–225. doi: 10.1016/j.ijer.2016.08.011
- Plakans, L., and Gebril, A. (2017). An assessment perspective on argumentation in writing. *J. Second Lang. Writing* 36, 85–86. doi: 10.1016/j.jslw.2017.05.008
- Pu, X. (2010). The effects of prewriting discussions on the content of argumentative writing of freshmen. *J. Tangshan Teach. College* 32, 150–153.
- Qin, J. (2013). Applying Toulmin model in teaching L2 argumentative writing. *J. Lang. Learn. Teach.* 3, 21–29.
- Qin, J. (2020). "Assessing L2 Argumentation in the UAE Context," in *The Assessment of L2 Written English across the MENA Region*, eds L. McCallum and C. Coombe (Cham: Palgrave Macmillan), 225–240. doi: 10.1007/978-3-030-53254-3_10
- Qin, J., and Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System* 38, 444–456. doi: 10.1016/j.system.2010.06.012
- Rahimi, M., and Zhang, L. J. (2021). Effects of an engaging process-genre approach on student experiences, engagement, and writing achievements. *Read. Writing Q.* 51, 1–18. doi: 10.1080/10573569.2021.1982431
- Rapanta, C., Garcia-Mila, M., and Gilabert, S. (2013). What is meant by argumentative competence? An integrative review of methods of analysis and assessment in education. *Rev. Educ. Res.* 83, 483–520. doi: 10.3102/0034654313487606
- Rusfandi. (2015). Argument-counterargument structure in Indonesian EFL learners' English argumentative essays: a dialogic concept of writing. *REL C J.* 46, 181–197. doi: 10.1177/0033688215587607
- Shegade, A. (2011). Effects and student perceptions of collaborative writing in L2. *J. Sec. Lang. Writing* 20, 286–305. doi: 10.1016/j.jslw.2011.05.010
- Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *J. Sec. Lang. Writing* 7, 319–345. doi: 10.1016/S1060-3743(98)90020-0
- Shin, Y. (2008). *The Effects of Planning on L2 Writing: A Study of Korean Learners of English as a Foreign Language*. Ph.D. thesis. Iowa: The University of Iowa.
- Siregar, R. N., Syarif, H., and Amri, Z. (2021). "An analysis of argument of students' argumentative essay by using Toulmin Model," in *Advances in Social Science, Education and Humanities Research Series: Eighth international conference on English language and teaching (ICOELT-8 2020)*, ed. W. Strielkowski (Dordrecht, The Netherlands: Atlantis Press), 98–102. doi: 10.2991/assehr.k.210914.018
- Stapleton, P., and Wu, Y. A. (2015). Assessing the quality of arguments in students' persuasive writing: a case study analyzing the relationship between surface structure and substance. *J. Eng. Acad. Purposes* 17, 12–23. doi: 10.1016/j.jeap.2014.11.006
- Storch, N. (2019). Collaborative writing. *Lang. Teach.* 52, 40–59. doi: 10.1017/S0261444818000320
- Storch, N., and Wigglesworth, G. (2007). "Writing tasks: The effects of collaboration," in *Investigating Tasks in Formal Language Learning*, ed. M. D. P. García Mayo (New York, NY: Multilingual Matters), 157–177. doi: 10.21832/9781853599286-011
- Strijbos, J. W., Martens, R. L., and Jochems, W. M. (2004). Designing for interaction: six steps to designing computer-supported group-based learning. *Comput. Educ.* 42, 403–424. doi: 10.1016/j.compedu.2003.10.004

- Sundari, H., and Febriyanti, R. H. (2021). The Analysis of Indonesian EFL Argumentative Writing Using Toulmin's Model: The Structure and Struggles from the Learners. *Scope* 5, 67–78. doi: 10.30998/scope.v5i2.8544
- Teng, L. S., and Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: can self-regulated learning strategies-based writing instruction make a difference? *J. Sec. Lang. Writing* 48:100701. doi: 10.1016/j.jslw.2019.100701
- Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Toulmin, S. E. (2003). *The Uses of Argument*, updated Edn. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511840005
- Uysal, H. H. (2012). Argumentation across L1 and L2 Writing: Exploring Cultural Influences and Transfer Issues. *Vigo Int. J. Appl. Linguistics*, 9, 133–159.
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Walqui, A. (2006). Scaffolding instruction for English language learners: a conceptual framework. *Int. J. Biling. Educ. Biling.* 9, 159–180. doi: 10.1080/13670050608668639
- Watanabe, Y., and Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners. *Lang. Teach. Res.* 11, 121–142. doi: 10.1177/136216880607074599
- Wei, X., Zhang, L. J., and Zhang, W. (2020). Associations of L1-to-L2 rhetorical transfer with L2 writers' perception of L2 writing difficulty and L2 writing proficiency. *J. Eng. Acad. Purp.* 47:100907. doi: 10.1016/j.jeap.2020.100907
- Wigglesworth, G., and Storch, N. (2009). Pair versus individual writing: effects on fluency, complexity and accuracy. *Lang. Test.* 26, 445–466. doi: 10.1177/0265532209104670
- Winzenried, M. A., Campbell, L., Chao, R., and Cardinal, A. (2017). *Co-constructing writing knowledge: Students' collaborative talk across contexts*. US: Association of Teachers of Advanced Composition.
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Commun.* 28, 193–219. doi: 10.1177/0741088311399236
- Wolfe, C. R., Britt, M. A., and Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Commun.* 26, 183–209. doi: 10.1177/0741088309333019
- Xu, J., and Cao, Z. (2012). The influence of different pairing patterns on the student-student interaction in college English class—An empirical study. *For. Lang. China* 9, 67–77.
- Yeh, S. S. (1998). Validation of a scheme for assessing argumentative writing of middle school students. *Assess. Writing* 5, 123–150. doi: 10.1016/S1075-2935(99)80009-9
- Zainuddin, S. Z., and Rafik-Galea, S. (2016). Effect of training in the use of Toulmin's model on ESL students' argumentative writing and critical thinking ability. *Malaysian J. Lang. Linguistics* 5, 114–133. doi: 10.24200/mjll.vol5iss2pp114-133
- Zhang, L. J. (2013). Second language writing as and for second language learning. *J. Sec. Lang. Writing* 22, 446–447. doi: 10.1016/j.jslw.2013.08.010
- Zhang, L. J. (2021). "Crossing literacy borders through writing: Transformational apprenticeship and repositioning of EAL learners," in *Crossing Borders, Writing Texts, Being Evaluated: Cultural and Interdisciplinary Norms in Academic Writing*, eds A. Golden, L. A. Kulbrandstad, and L. J. Zhang (Bristol: Multilingual Matters), 147–163. doi: 10.21832/9781788928571-011
- Zhang, L. J., and Cheng, X. L. (2021). Examining the effects of comprehensive written corrective feedback on L2 EAP students' linguistic performance: a mixed-methods study. *J. Eng. Acad. Purp.* 49:101043. doi: 10.1016/j.jeap.2021.101043
- Zhang, S. J., Yu, H., and Zhang, L. J. (2021). Understanding the sustainable Growth of EFL students' writing skills: differences between novice and expert writers in their use of lexical bundles in academic writing. *Sustainability* 13:5553. doi: 10.3390/su13105553
- Zhang, L., and Zhang, L. J. (2021a). Development and validation of the Questionnaire on EFL Students' Perceptions of Authorial Stance in Academic Writing. *Appl. Linguistics Rev.* 12, 1–33. doi: 10.1515/applirev-2021-0003
- Zhang, T. T., and Zhang, L. J. (2021b). Taking stock of a genre-based pedagogy: sustaining the development of EFL students' knowledge of the elements in argumentation and writing improvement. *Sustainability* 13:11616. doi: 10.3390/su132111616
- Zhang, Y. (2018). An investigation into the development of structure and evidence use in argumentative writing. *Theory Pract. Lang. Stud.* 8, 1441–1448. doi: 10.17507/tpls.0811.08
- Zhao, H. H., and Zhang, L. J. (2022). *Teaching Writing in English as a Foreign Language: Teachers' Cognition Formation and Reformation*. Chem: Springer Nature.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



From Text on Paper to Digital Poetry: Creativity and Digital Literary Reading Practices in Initial Teacher Education

Moisés Selfa Sastre* and Enric Falguera Garcia

Department of Specific Didactics, Universitat de Lleida, Lleida, Spain

OPEN ACCESS

Edited by:

María Isabel de Vicente-Yagüe Jara,
University of Murcia, Spain

Reviewed by:

Inmaculada Clotilde Santos Díaz,
University of Malaga, Spain
Proshan Mahbub Ibna Seraj,
American International
University-Bangladesh,
Bangladesh
Anastasio Garcia-Roca,
University of Almeria, Spain

*Correspondence:

Moisés Selfa Sastre
moises.selfa@udl.cat

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 24 February 2022

Accepted: 11 April 2022

Published: 17 June 2022

Citation:

Selfa Sastre M and Falguera
Garcia E (2022) From Text on Paper
to Digital Poetry: Creativity and Digital
Literary Reading Practices in Initial
Teacher Education.
Front. Psychol. 13:882898.
doi: 10.3389/fpsyg.2022.882898

The new contexts of literary education allow for the creation of digital reading and writing practices related to what specialised literature calls digital literature. Among these practices and with an eminently theoretical content and with an example of this content, in this paper, we want to focus our gaze on cyberpoetry, conceived as an exercise in literary creativity that firstly involves use of technology and specific software for the digital creation of poetic texts and, last but not least, knowledge and mastery of poetic language and the literary conventions linked thereto. From this point of view, in initial teacher training, we work with future teachers to create cyberpoems with a dual purpose: on the one hand, to reflect on what literary reading in digital format entails and to rehearse reading mediation processes that can be carried out with this type of literature, and, on the other, to begin in the digital creation of cyberpoems that, later, may be presented and worked in a real context of the school classroom. In this paper, we present digital practices of literary reading that have been created by student teachers in initial training. The creation of these practices has been carried out with the *Genially* tool. For data collection, the URL of each creation is accessed in order to analyse them. The analysis of the data follows the parameters of the qualitative methodology, specifically based on three categories of analysis for each digital creation: multimodality, hypertextuality and the interaction of each digital reading proposal. The conclusions of all this allow us to affirm that the creation of cyberpoems is an exercise in literary creativity that has to take into account the digital dimension of the literary text and its reading comprehension in a multimodal environment. Thus, the teacher in initial training carries out a digital literary mediation exercise, of a creative type, which he will later carry out in his pedagogical practice.

Keywords: digital literature, digital practices, cyberpoetry, creativity, digital writing, teachers

INTRODUCTION

Literature, i.e., the art of written or spoken expression, is an artistic manifestation that every cultural community has cultivated throughout ages. As such, it is no stranger to advances in civilisations that have been imposed due to many factors. Two can be generalised among them: the development of human thought and the technological progress associated with it (Solbes

and Traver, 2015). From this point of view, literature has posed a challenge to the reader, especially since the discovery of the printing press and, therefore, of the transmission of literary text from a printed medium, which can be summarised as follows: the author expects a response from his potential receiver, which is the object of study of literary criticism (Chiappe, 2015).

Today we are immersed in a markedly digital age, where technological advances have expanded the author's relationship with the reader of his text (Handayani et al., 2020: 65–74). Therefore, the questions we ask ourselves are twofold: how the reader, from a markedly technological environment, understands literature and how he sees it, and whether there are substantial changes, inherited from written culture, in the way of relating to the literary text. In order to approach these issues, which have a conceptual and markedly epistemological basis, we will focus on the field of literary reading and, more specifically, on digital literature, understood as those literary works created specifically for the digital format—mainly, but not only, for the Internet—, and which could not exist outside of it (Bootz, 2021: 7–22). It is not digital literature what is thus found in a virtual library, nor, strictly speaking, the linear narrative that is created for the Internet, since that same literature could be presented in book format. The interaction of the reader with the literary text based on the possibilities enabled by technology is what defines digital literature.

To this end, we establish three questions that we want to develop in their most fundamental features, namely, approaching the context of digital culture in which reading and literature are clearly inserted; identifying the parameters that allow us to underline creativity as the basis and sustenance of the reading of digital texts; and, finally, defining the variables of the creation of digital texts that imply a very particular reading of these texts. All the foregoing aims at advancing in the strategies for understanding and training readers and writers in the context of digital literature.

LITERATURE IN THE DIGITAL CULTURE SPHERE

Literature, understood as the art of the word, generates discourses of all kinds. And not only does it generate them, but it is also connected to such discourses, as they are all interrelated to each other through relations of continuity, opposition and provocation that link and oppose them (Morales-Sánchez and Martin-Villarreal, 2019). Digital literature has also changed the traditional relationship between the author of a text and its reader to the point of almost eliminating the role of the publisher at a stroke. This traditional triangle formed by author, reader and publisher is modified in a literate context dominated by digital culture, due to the near disappearance of the publisher, a fact that allows for simultaneous, reciprocal and not at all deferred communication between the author and the reader of the literary text (Cordón García, 2016).

The problem is no longer whether it will be possible to read literary texts on paper in 100 years' time, but whether it will make sense to read them on paper at all. Therefore, when

thinking about digital literature, it would be logic to focus on the dialogical strategies involved in thinking digital (Engen, 2019: 9–18). When all the foregoing is approached from literary discourse, an author's approach to writing a literary text or the process of reading it takes on a truly interesting magnitude. Therefore, both reading and generating literature from this new digital space are essential to understanding digital literature and the concept of reading and creation that contextualises it (Pajares, 2015).

Focusing on the aspects that define the habits of writing and reading digital literary texts, it is necessary to highlight how digital literature has undeniably evolved at the vertiginous pace at which technology is also changing.

In the light of this, digital literary discourse demands an active response from the reader and an interaction with the literary work in which the classical hierarchical relations between author and reader are broken. Consequently, when we ask ourselves how, the way of constructing and receiving the literary text—as an aesthetic discourse—has changed in terms of literary communication, we see significant changes not only in the ways of conveying information, but also in articulating the mode of persuasion, as the focus is no longer only on the pure content of the literary message (Alghadeer, 2014: 87–96). It is also in the form it takes. We can then logically think that each historical-cultural stage has its own way of creating and persuading according to its own technological domain and the dimensions of the reading masses, who are the recipients of literary discourse.

Nowadays, digital culture, which includes literature as a digital form of expression, brings infinite communicative possibilities demanding new ways of reading and creating in line with the digital and technological divide (Kangasharju et al., 2021: 52–91). All this is reflected in the most current digital literature in the form of artistic proposals, which convey the hybridity, versatility and complexity of new ways of understanding and interpreting literary discourse.

DIGITAL LITERATURE, LITERARY READING AND CREATIVITY: AN ESSENTIAL TRINOMIAL

Reading a given literary text involves the effort to construct the meaning of a message with one or more aesthetic intentions. Hence, literary reading is specifically understood as an intellectual process whose essential principle is that of literary construction. By literary construction, we do not only understand the act of writing for aesthetic purposes, but we must also understand how reading requires a similar, and sometimes greater, effort than creating.

Digital literature presents the reader with a set of texts in which the word is not only the ultimate goal of reading, but also broadens its horizons towards other analogue, visual and sound codes, thus allowing experimentation linked to the technique, the software and the digital possibilities surrounding the literary message (Torres and Córtes, 2021: 3–26). All the

foregoing means that the process of constructing the digital text escapes the variables usually handled by who is reading the text. One of the singularities of digital authorship is its markedly collective nature, as writers, designers, computer scientists, audiovisual technicians and among others often join hands in a choral exercise of harmony in which the common denominator is creativity (Saum-Pascual, 2018: 34). Hence, the confluence of different creation nodes that transcend literary creativity as it has been conceived until a few days ago. We can therefore speak of a polyphonic discourse where creativity symbiosis allows the reader to grasp different perspectives and simultaneous aesthetic approaches.

From this point of view, it is coherent, following the approaches of Bakhtin's concept of aesthetic polyphony, to speak of polyacrosis as the basis and foundation of digital literary texts, given that readers belong to different cultural and digital spaces and, therefore, with divergent reading habits and intertexts (Navarro-Romero, 2021: 75–81). The range of readings that the same digital text can suggest allows us to speak of the concept of reading-writing (Llosa, 2019: 105–112), i.e., the possibility of handling the text not only in the reader's intellect, but the reader takes on the role of creator by being able to intervene in the literary message directly. We are thus talking about someone who produces his own texts almost simultaneously with their reception.

This approach leads us to affirm that the authorship of the digital text can also lie with the reader. Thus, three more factors in this writing process should be considered: firstly, the vision of the reader as a player who develops strategies and techniques derived from the multimedia world; secondly, the identification of the reader as an actor who simultaneously constructs and deconstructs a pattern of reading and writing that must be seen while selecting the tools to do so and accessing the meaning of the digital text; and thirdly, digital literature has a markedly artistic vein in which the writer resorts in parallel to a double code, that of literary language and that of computer software, in such a way that the presence of technology in the parameters of creation is indisputable: That of literary language and that of computer software, in such a way that the presence of technology in the parameters of creation is indisputable. In this vein, digital literature turns readers into experimental creators (Escandell, 2020: 91–103).

All this means that the vision of the digital work can become an apparent obstacle for the reader, who may give up trying to set the digital work in motion. This occurs when the digital base is dissociated from the text by approaching it exclusively with conventional keys. It is therefore a technological dynamism where the literary game also includes a technical game in terms of mastering software and the multiple resources linked thereto. The reader needs to understand digital communicative strategies and the effects they can have on the digital text itself.

Barely a decade ago, Romero-López (2012) proposed a first classification of digital literature insofar as it was handled by the reader. He distinguishes three types of this literature: hypertextual literature, in which the link plays a key role in the creative process; ecphrastic literature, in which text and image form a harmonious whole; and serendipitous literature

(serendipity), marked by chance or unpredictability through which the reader can achieve his creative purpose.

In any case, there is no doubt that digital literature, with its markedly creative character, offers the reader many creative possibilities that he will have to experiment with. From the mastery and knowledge of the essence, intention and meaning of the literary message, and from the competence in knowing how to master the techniques of digital creation, a creative experience will emerge, in which reading and digital writing act in parallel and at the same level.

POETRY IN THE CONTEXT OF DIGITAL LITERATURE

The developments in technology have always had a huge impact on every field of human activity. It is therefore necessary to recognise that technology has been and is a key factor to be taken into account in the progress of the history of literature. If the technology of the printing press changed the medium of transmission of literature from oral to written, the Internet allows the poet to approach the reader through digital poetry, whereby the latter can get to know the poem's setting of movements and sounds beyond its mere representation in a digital environment.

The Internet has provided a new scenario for poetry in two ways: as a means of dissemination and promotion of the poetic text, as well as a creative tool for the design of the literary message. As Torres (2013) points out in this regard, the poetic genre has found in technology 'a new transtemporal and trans-spatial dimension, expanding its frontiers with surprising and open forms'.

It should be noted that one of the literary genres with the least presence in formal literacy settings has been chosen for this change of formal discourse. The reason for this—or one of them at least—is that the potential reader of digital poetry is an active and immediate recipient, a true performer of the literary work in digital environments. It should be made clear that not everything published on the Internet is digital poetry, since digitisation implies, as we have already said, interaction with the text.

Towards a Definition of Digital Poetry or Cyberpoetry: Identifying Characteristics and Classification Proposal

Digital poetry is the latest challenge of digital literature in the so-called paraliterary era. In this sense, poetry spontaneously shifts to digitalisation, visuality, fragmentation and, at times, absurdity. As we will see in our study, all this has led to the emergence of poetic practices enabled by technology and for digital media, since today printed texts lose their importance due to the technological development of digital media (Soccavo, 2015). However, there is no doubt that this new poetry, of a markedly electronic nature, whose natural space is the web and which makes the most of the capabilities that the digital medium offers (Saum-Pascual, 2019), must maintain the aesthetics

and creativity that are typical of this genre. The reason for this lies in the main characteristic of poetic texts: recipients thereof continue to attach greater importance to the aesthetic value of the linguistic message and to the genre's ability to move the listener's mood and its special versatility in renewing language. However, it must be clear that cyberpoetry is not writing a poem and posting it, for example on Twitter.

As a mirror of the most recent human activities and philosophies of society, poetry moves in step with modern technology, which believes in the globalisation of art and literature. Therefore, digital poetry not only breaks all the traditions and limitations of the written text, but also frees itself from the constraints of a certain culture or a certain language. In other words, digital poetry can be described as an electronic, multicultural and multilingual global movement, which can merge the textual and visual elements of poetry with technology (Kovalik and Curwood, 2019: 185–195). It will also be necessary to mention that the ability of digital poetry to consolidate itself as a human activity born within digital culture is its flexibility and lack of concrete form, its willingness to show a strong symbiotic relationship with the technology that hosts it and its ability to adapt to a wide variety of environments for its transmission.

Digital poetry is heir to the aesthetics of the historical *avant-gardes* and, to a greater extent, to the *avant-gardes* of the second half of the 20th century, such as Lettrism, concrete poetry and conceptual poetry (Saum-Pascual, 2019). Lettrism, founded by Isidore Isou in Paris, understood that poetry should be based on the constituent elements of language such as letters and sounds, beyond meaning or semantics, used in a new way that took into account their materiality and not their representational capacity. In the context of lyricism, the notion of *détournement* also arose, in which social messages were appropriated, altered, intervened, subverted or 'hacked' (Bentivegna, 2019: 135–156).

Digital poetry has received plenty of names, but one of them that seems very appropriate to the specialised bibliography is cyberpoetry (Escandell, 2019), that is, a form of cyberliterature in which the aesthetic function of language dominates the literary message. Technically, cyberpoetry is characterised by the use of various electronic resources, including the well-known hypertext, two- and three-dimensional animation and more advanced realities such as virtual and multimodal. However, it cannot be forgotten that cyberpoetry, in addition to being associated with the emerging technology of the moment, is above all associated with the biological component of literature, that is, it is created to be enjoyed aloud insofar as it constitutes a full aesthetic experience (Escandell, 2019).

There are three basic characteristics of cyberpoetry: firstly, experimentation, since it seeks the limits of the art of literature and its ultimate aim is to break the rules of this art (Morales-Sánchez, 2019); secondly, collectivisation in digital environments that can be accessed by an unlimited public; and finally, co-creation, given that it is not an individual work but can involve artists from different backgrounds and tendencies who converge in the same text: from the writer himself to the author of the visual animation that surrounds the work (Campos-F-Figares, 2021: 7–10; Falguera-Garcia and Selfa-Sastre, 2021: 4–5).

There is a possible classification of the cyber-possibility if we take into account the technology used for its construction. Such classification would be as follows:

Hypertextual poetry is poetry that is linked to another location in the same document or to different documents on the Internet. Although there are fewer examples, it is also the one that uses hypertext for the creation of a poetic work (Morales-Sánchez, 2019).

Animated or moving poetry, also called kinetic poetry (Campos-F-Figares, 2021: 9): poetic works where the text's words are gradually moved or modified, either by the user's interaction with the text or automatically.

Holopoetry or holographic poetry: This is the name given to texts with poetic content developed through the use of the holographic technique. Thus, the poem is organised in a non-linear way in a three-dimensional immaterial space, changing and conveying different meanings even as the reader looks at it (Morales-Sánchez, 2019).

OBJECTIVES

In order to analyse digital literary reading practices and, specifically, the creative editing of cyberpoems in initial teacher training, our study has two main objectives, listed and explained below.

The first objective of our study concerns the digital literary reading practices to be carried out by student teachers in their initial training. Specifically, this objective would read as follows:

- To select literary texts, specifically poems, to which a digitalisation process will be applied and which, after this process, can be handled through digital tools without losing the original meaning of the poem.

In this vein, the content and nature of the poem upon digitalisation cannot lose its original meaning. Therefore, there is a first phase of selection of texts, which have generally been published on paper.

The selection of poems should take two criteria into account: (a) the authors of the texts selected should be recognised by literary tradition; (b) that they can be worked on in Catalan schools throughout the First Cycle of Primary Education, that is, with students between 6 and 8 years of age.

The second objective concerns the analysis of the digitised poem. In particular, the following objective is pursued as:

- To relate digital literary reading practices to one of the types of cyberpoetry described in the previous section and to analyse them in relation to the categories and analysis processes related to the most important characteristics of digital poetry. These characteristics have to do with the multimodality, hypertextuality and interactivity that are characteristic of digital poetry.

RESEARCH DESIGN AND METHOD

Research design takes into account the data collection techniques, method and analysis categories and processes chosen by a researcher to combine them in a reasonably logical way so that the research problem is handled efficiently.

Research Design, Participants, and Data Collection Techniques

The aforementioned digital literary reading practices were designed within the subject Didactics of Language and Literature II, 3rd year of the Double Degree in Early Childhood Education and Primary Education at the Faculty of Education of Universitat de Lleida (UdL). In this subject, we work on content and skills related to children's poetry, its authors in Catalan and the characteristics of digital poetry to be worked on in the classroom in order to develop the reading and literary skills of students aged between 6 and 8.

The research design was carried out as follows. The participants are 30 students who work in groups of 5 people and have received theoretical and practical training in the classroom on what digital literature is, its most important characteristics, the definition of what digital literary reading practices are with examples of this and the types of cyberpoetry that the specialised bibliography contemplates.

They are digital natives, which means that they are familiar with digital tools for the creation of virtual literary reading environments. From this point of view, the participants in this research designed their digital literary reading practices based on the *Genially* application,¹ an online tool suitable for creating creative, interactive and animated content. This tool allows work based on three basic principles that can be used to create digital literary reading practices. These are the following:

- Animation. *Genially* allows to animate images and graphics through motion. Thus, the creator of the digital content can easily configure entry or exit animations, to indicate just one example. In this sense, it enriches the content created with visual effects that convey the idea of movement. In the specific case of digital literary reading practices, with this application, we create moving images associated with poetic text.
- Interactivity. *Genially* allows also, without the need to have specific programming knowledge, to interact with the poem that it is going to be digitized. The creator of the digital proposal decides where to intervene to create more visual content, which has often movement and, therefore, transmits less saturated information.
- Integration. *Genially* facilitates the integration of information that it is on the Internet or that the user has saved on different platforms. Thus, it can place content from different sources within *Genially* that will continue to work through hyperlinks. We are referring to platforms such as YouTube, Videos, Documents, Gadgets and 3D Images.

Once the proposals have been created in this *Genially* application, the application generates a URL as an electronic address, which we have used to collect the data. The collection of these electronic addresses helps that researcher to access each proposal created for subsequent analysis.

¹<https://genial.ly/es/>

Method and Categories and Analysis Processes

The research method is defined as the set of techniques that, consistent with the orientation of an investigation and the use of certain tools, will allow obtaining a particular product or result. In our case, we opted for qualitative research, which tends to seek the causes of phenomena in the depth of the interpretations that the subjects make of them. In this way, we have worked with a sample of materials, which has constituted our sample. The qualitative methodology allows us to obtain data from the meanings that each part of the sample wants to convey. This obtaining allows us to establish categories of analysis for each digital practice of literary reading.

The analysis categories to study the digital literary reading practices in question, which is the result of a previous process of digitisation applied to a literary text that was conceived to be read as printed, are the following:

- Multimodality: the resulting digital text presents audio, voice and image modes that refer to the main content(s) of this literary text. These poems contain sensory words easily associated with sounds or images. The potential for digitisation of poems is greater where verses include words that can be quickly and almost involuntarily related to specific images, colours or sounds. These concepts are called sensory words, as they trigger memories related to the senses and evoke and express emotions.
- Hypertextuality: the digital text establishes a relationship that a text b (hypertext) maintains with a previous text a (hypotext) in which it is inserted without being a commentary. One of the characteristics of digital literature is the possibility of including hyperlinks to connect the different intertexts that the poem may include. This way a universe of connection between readings is created.
- Interactivity: the digital text is focused as a game with the explicit participation of the reader. Therefore, it allows for the narrative creation of the reader, as the reader is positioned as a collaborator in the storey, either as an avatar or as an outsider, with a narrative responsibility. They are poems that convey a sense of dynamism/action. In reference to this criterion, we find that certain poems quickly evoke a sense of movement, either because of the rhythm of the poem or because of what the content of the poem expresses. This feature is accentuated when the text contains action verbs.

Therefore, our study, as we have said before, is both qualitative and descriptive, as it aims at showing very specific digital literary reading practices whose main purpose is the creation of virtual environments for working with poetry. It should be noted that the qualitative study is one of the most operational in educational research as it allows for a direct relationship between the researcher and the object of study.

RESULTS AND DISCUSSION

In order for the results and their analysis to be presented, we will consider the two objectives of our research and the three analysis

categories and process defined and explained in the ‘method’ section above. As we have already said, this is a qualitative, descriptive study which, it must be said, selects those digital practices that allow us to draw conclusions, as we will discuss later.

Text Selection

Text selection is related to the level of teaching and learning they are aimed at. In the case of poetry, a genre generally considered eclectic in the school sphere with respect to other genres more commonly worked on in the classroom, such as the narrative genre, it has to present a set of characteristics for students from 6 to 8 years. These characteristics are the following: it allows students to approach their immediate environment through simple verses in their first years of compulsory schooling. The practice of this literary genre should awaken a special motivation among our schoolchildren to learn it, as long as this activity is approached in a playful and attractive way. Secondly, this taste for verses will give rise to poetic texts that will allow the education of the aesthetic sensibility and taste of students who will begin to be motivated by the creation of their first poems based on the correct interpretation of those read in the classroom. This creation can take the form of oral and written productions. Finally, the selected poetry should serve to express one’s own inner world in words that constitute a privileged instrument for the oral and written expression of feelings, emotions and univocal experiences (Selfa and Azevedo, 2013: 58).

The selected poems, the primary basis for applying a digitisation proposal, have very appropriate characteristics that make them especially suitable to be worked on with children between the ages of 6 and 8. From a formal point of view, they are short poems, of no more than two stanzas and with lines of verse generally of less than 8 syllables. They are therefore very suitable for memorisation and subsequent recitation aloud. From a thematic point of view, they deal with realities very close to children of the ages outlined above. As we will see later when we analyse the digitisation proposals applied to these, they refer to realities such as letters, the alphabet, food (tomatoes and bread, for example), animals (frogs, for example) and elements of nature that form part of children’s imaginary (the sun, rain and spring, for example). These are just a few examples, which can be accessed through the hyperlinks below.

Types of Digital Reading Practices

Generally speaking, in the poems that form the basis of our analysis, kinetic poetry predominates, that is, poetry that produces some kind of movement from the presentation of the digital text to the reader and, above all, in the interaction that the latter has with the poetic text. The poem *Sopa de Lletres (Alphabet Soup)* would be a first example of this.² The way in which it is initially presented to the reader is in movement. The letters of the alphabet, which are the main motif of the poem, are presented in twists and turns until they adopt a static position. The same applies

to the set of letters below the poem with which the reader can interact to create new words.

A similar case would be that of the digital poem *Cançoneta de les granotes (Frog song)*,³ in which the reader perceives the movement of the numbers 1, 2 and 3 linked to the numerical concept they present.

The kinetics of the digital text can also be represented by the movement offered by some of the images that refer to the content of the literary text. In the digital text dedicated to letter *F*⁴ and entitled *F*, the pendular movement of the capital letter, adorned with flower attributes, is reminiscent of one of the words that are central to the understanding of the literary message: flower. This movement is the one produced when the flower is shaken by the wind.

This type of kinetic poetry sometimes presents a more static movement, involving a visual transit through the poem and then interacting with it. The poem entitled *Pa amb tomàquet (Bread and tomato)*⁵—a typical breakfast in Catalonia—in the first picture thereof, the reader is informed that in order to play with it, he must pick the picture of one of the three cooks on the title page of the poetic text. Subsequently, a second picture appears in which the reader must fill in missing words in each of the lines of the poem with images that appear visible or hidden in some parts of the kitchen.

There are also some cases of hypertextual poetry, in which the digital poem itself can be used to access other documents posted on the Internet that are related to the content of the poem. This is the case of *Cançoneta de les granotes (Frog song)*.⁶ In this one, through the selection of a few words, the reader can travel through the Internet through hyperlinks created to broaden and deepen the meaning of the poem.

The same can be said for the hypertext-based digital text *Aficions*,⁷ in whose verses highlighted words appear and which refer to sound and visual archives that broaden the meaning of the concepts expressed.

We have not found example of holopoem in the group of cyperpoems that form the basis of our study. This may be because the poetry is three-dimensional, very digitally creative and involves a spatial vision that is more suited to adult readers. It should be remembered that, as we have already pointed out, the cyberpoems created here are aimed at students aged between 6 and 8.

Multimodality

Multimodality aims to facilitate the reader’s digital interaction with the poem. Normally, from a very early age, students are immersed in the digital world, which allows them new ways of reading that modify traditional ways of learning. In the digital

²<https://view.genial.ly/61bf5b75e6cbe90e3841ff7c/interactive-image-interactive-image>

³<https://view.genial.ly/61ba10d1295ff00de456c1e2/interactive-image-canconeta-de-les-granotes>

⁴<https://view.genial.ly/61c345f84d9c8c0df24d07e0/interactive-content-f>

⁵<https://view.genial.ly/61c2fee64d9c8c0df24cfc8f/interactive-content-poemadigitalcostaforonsgallegosole>

⁶<https://view.genial.ly/61c3057a4d9c8c0df24cfddd/interactive-image-canconetadelesgranotes>

⁷<https://view.genial.ly/61bc52ad5ea1040d618602c6/interactive-content-poema-digital>

proposal *Sopa de Lletres* (see footnote 2), multimodality is provided by the link between the digital text and a sound file located at the bottom right of the text. To access this file, the reader selects the voice icon which takes him/her to a sound video, in which you can discover all the letters of the alphabet and see them appear in movement in full colour.

Similarly, in initial interface of *Pa amb tomàquet* (see footnote 5), a musical icon appears next to the title of the poetic text, which again allows us to highlight the multimodality of the poetic text. In this case, by selecting the aforementioned icon, the reader accesses a Youtube video in which the reader mediator addresses the reader of the poem, explaining how this breakfast is prepared through gestures and the use of the narrative voice.

Multimodality is also expressed by moving and sometimes static images representing key words in the digital text. In *F* (see footnote 4), the word *ferida* (wound) is healed by a band-aid and, in the last verse of the text, the word *petó* (kiss) is represented by red lips that prefigure the image of a kiss in movement.

Hypertextuality

The digital text is the hypotext, i.e., the starting point that refers to new texts, called hypertexts, which provide information of all kinds related to the main text. Hypertextuality could be said to be a static function, as one simply shifts from one digital stage to another digital stage. As mentioned above, there is an informational connection between the two. This can be found in *Pa amb tomàquet*,⁸ by the universal Catalan poet Miquel Martí i Pol. In the initial presentation of the poem, next to the author's name, there is an icon that refers to a call for information. When the reader selects it, they are taken to an Internet page where they can zoom in and find out in depth who the author of the poem is and what kind of poetry he has written. In this case, it is thus a matter of extending information through a digital narrative text.

In other cases, hypertextuality, as outlined above, allows the reader to virtually scroll through other links on the Internet to learn more about the words in the digital text. In *Cançoneta de les granotes*,⁹ the digitisation proposal allows, in the third verse, access to an image that presents the word *estiu* (summer) and, in other cases, access to the visualisation of writing in capital letters, typical of the early stages of learning to read and write, of the word *sol* (sun).

Interactivity

Poetry is a game. As such, it allows an interaction that places the reader as a creator of new poems, enabling new learning in accordance with his age and reading and literary competence. In *Sopa de Lletres* (see footnote 2) above, we can see how the reader can create new words without losing the original meaning with which the poem was created. To do this, the digital editor

of the poem warns the reader that he can move the letters and lines and create new words, as well as compose his own poem. Letters and verses that are not needed can be left out of the yellow box. This is the poetic and digital game that is proposed. We see how the initial word *mitjó* (sock) which appears in the initial proposal of the author of the poem, when the reader plays with it, he can modify it and obtain a new word, as in this case is *águila* (eagle).

At other times, it will not be a matter of creating new words, but inviting the reader to create a new poem by moving the words of the poem. As can be seen in the poem entitled *F* (see footnote 4), the words written in capital letters can be moved and combined to obtain a new poem and thus give the text a new poetic meaning without losing the spirit of the text.

Another example of the creation of a new poem from the initial stage offered by the digital text is also the poem *F*,¹⁰ to which a different digitisation proposal than the one presented above is applied. In this case, the reader can compose new verses, different from the original version created by the author of the poem, by moving the words and placing them where they can make sense.

As mentioned above, interactivity requires manipulating the digital text. This manipulation can also be associated with images that help to better understand the content of the manipulation, as we can see, for example, in *Pa amb tomàquet* (see footnote 11). The reader of the text, once he has accessed the stage with which he can play with the text, will be able to complete the missing words in the poem by moving images that refer to these words. These images are visible at first glance, as in the case of the ham on the table, or they may be hidden in the kitchen cupboards or fridge. The reader of the text will have to digitally open the doors of these cupboards and fridges to find the images that will later be placed in the digital poem.

The evaluation of the digital practices of literary reading was carried out with an evaluation rubric, which is a tool that helps to evaluate the students' learning, making the students themselves also know their successes and errors through self-assessment. This rubric was given to the students before starting the creation project and has two elements: a vertical column, which contemplates the evaluation criteria of said creation and which, in our case, have to do with the categories and analysis processes; a horizontal column with the quality grades of those criteria. The rubric used was the following (Table 1).

The results that we have obtained and that we have discussed in the theoretical framework of this research allow us to affirm that, in general terms, there are studies related to cyberpoetry, as a branch of digital literature, that describe and analyse the digital environments in which cyberpoems are published. They are studies that research the work of young authors and that the literature calls *Millennials*. They are called as 'childrens of Instagram' (Sánchez García and Aparicio Durán, 2020: 41–53). Many of them have a place in the digital channels of literary communication that are

⁸<https://view.genial.ly/61c2fee64d9c8c0df24cfc8f/interactive-content-poemadigitalcostaforonsgallegosole>

⁹<https://view.genial.ly/61c3057a4d9c8c0df24cfd4d/interactive-image-canconetadelesgranotes>

¹⁰<https://view.genial.ly/61c20f6461aa590e006ef6a0/interactive-image-f-salvador-comelles>

TABLE 1 | Evaluation rubric.

Indicators	Very good	Good	Regular	Insufficient
Analysis categories and processes				
Multimodality				
Hypertextuality				
Interaction				

Source. Own creation.

widely distributed on the Internet. An example of what was said is *Wattpad* (Falguera-Garcia and Selfa-Sastre, 2021). Some have even been recognised with national and international literary awards. However, the implication of our study goes beyond analysing creators of digital literature. In our case, we are talking about literary reading mediators who create from digital tools, as is the case with *Genially*. Starting from a text by another author, the teacher in initial training applies a digitization process to the literary text in order to be able to play, interact and experiment with it. Only in this way is the interaction effective, practical and the transition between the paper text and the digital text are completely assured without altering its message.

CONCLUSION

This article aims at exploring the role of digital literature in initial teacher education. From an eminently theoretical perspective, we identify the concept of digital literature in current literacy contexts and, specifically, what we understand by digital poetry from the point of view of creativity. The final objective of this work is precisely this to show models of digital literary reading practices that can be carried out in the classroom, starting from the premise that these are texts selected for students aged 6 to 8 years old and that they are still in the initial stages in the formation of their literary competence.

From our study we conclude that, in the first place, the selection of texts published in printed anthologies follows the criteria of children's poetry: poems with short verses and few stanzas, close to the oral language and therefore easy to memorise and later to declaim. These poems thematically refer to realities that are well known to students aged 6 to 8: nature, animals, food and childhood experiences. From this point of view, the digitisation proposal applied to these poems will take these elements into account and, therefore, digital play with this poetry will be the common denominator of the digital practices analysed in our study.

On the other hand, if we look at the classification of the digital practices presented, which are the result of a previous digitisation of poetry that has been initially conceived to be declaimed, the vast majority of these are kinetic poems or poetry in movement. These are digital proposals that allow experimentation through the movement of letters, words and images that result in new poems with a different

meaning to the original poem, although the initial essence of the poem is not lost, i.e., its structure is practically the same.

Secondly, in terms of the categories of analysis of digital poetry practices, interactivity, understood as a game, as experimentation and as a new way of learning concepts, stands out in the first place. Through digital tools, the digital reader can move, change and transform the reality of the poem they are working with.

Multimodality is also present in the digital proposals under analysis. In these cases, when we speak of multimodality, we are referring to sound files that reproduce sounds related to the key words of the poem. These are characteristic and well-defined sounds that facilitate the understanding of the poem and are usually hosted on well-known platforms such as Youtube. Similarly, multimodality is also related to moving images, which again, link to key words in the poetic text.

Finally, there are also digital practices of literary reading of poems in which hypertextuality prevails in the digitisation proposal. These are less elaborate, more primitive practices that basically establish a relationship between a hypertext and a previous text, the initial poem, which we call hypotext. In the examples shown in this text, hypertextuality is often used to expand info on the author of the initial poem. Its function is basically informational.

Finally, our work has its limitations, which are inherent to the selection of a very specific sample of texts. Among these limitations, we highlight the following: (a) one of them has to do with the selection of texts in a particular language: Catalan. We have already said that student teachers in initial teacher training prepare these digital practices for a specific school context, schools in Catalonia. While this must be the case because it is a reality that determines the choice of literature to work with, future studies can compare the creation of digital practices in Catalan with other languages. Secondly (b), the age groups targeted by digital practices also condition the preparation of these practices. These are early ages, where literary and digital competence are being constructed. These are therefore sometimes flat proposals which, with a greater degree of elaboration, can be applied at older ages. Finally (c), the digital practices analysed have an impact on reading comprehension of the text and very little on the written elaboration of new messages. From this point of view, some kind of interaction with the text could have been thought of, involving the rewriting of the text not only by moving images or letters, but also by creating more elaborate messages.

All of this do not detract from the digitalisation proposals presented, understood within the context of the digital literacy that takes place in schools that cater for children of an early age, as are the targets of the digital practices we have analysed. Digital literature is emerging strongly from creative, original perspectives that facilitate access to the text. This way, digital literature and, specifically, digital literary reading practices, as particular manifestations of the former, are conceived a way of presenting and working

with the initial reader in a literary genre very close to oral language, which offers many possibilities for interaction, play and learning. The objective of the digital texts under analysis in this study is to learn from poetry, through playful environments previously created from the premise of digital creativity.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

REFERENCES

- Alghadeer, H. A. (2014). Digital landscapes: rethinking poetry interpretation in multimodal texts. *J. Humanities* 3, 87–96.
- Bentivegna, A. (2019). Concepto y empleo del “Détournement” en el ámbito del proyecto de renovación social de la Internationales Situationniste (1958–1972). *Fedro Revista de Estética y Teoría de las Artes* 19, 135–156.
- Boetz, P. (2021). The ephemerae of digital literature. *Neohelicon* 48, 7–22. doi: 10.1007/s11059-021-00583-0
- Campos-Figares, M. (2021). Creación poética en nuevos contextos: poesía en red y ciberpoesía. *Ocnos. Revista de estudios sobre lectura* 20, 1–14. doi: 10.18239/ocnos_2021.20.3.2547
- Chiappe, D. (2015). La autoridad de la narrativa multimedia recae sobre el lector. *Madridlnea.com*. Recuperado de. Available at: <http://madrilnea.com/2015/03/12/domenico-chiappe-la-autoridad-de-la-narrativa-multimedia-recae-sobre-el-lector/> (Accessed February 18, 2022).
- Cordón García, J. A. (2016). La lectura en el entorno digital: nuevas Materialidades y prácticas discursivas. *Rev. Chil. Lit.* 94, 15–38. doi: 10.4067/S0718-22952016000300002
- Engen, B.-K. (2019). Understanding social and cultural aspects of teachers' digital competencies. *Comunicar* 27, 9–19. doi: 10.3916/C61-2019-01
- Escandell, D. (2019). Ciberpoesía y literatura en red. eds. E. M. Campos Fernández-Figares and D. E. Montiel Poesía en red y ciberpoesía.
- Escandell, D. (2020). Lecturas juveniles mediadas por las TIC: pantallas, mundo digital y transmedialidad. Presencias y percepciones entre estudiantes. *Contextos educativos: Revista de educación* 25, 91–103. doi: 10.18172/con.4270
- Falguera-García, E., and Selfa-Sastre, M. (2021). Poesía en red: lectura y escritura en Wattpad. *Ocnos. Revista de estudios sobre lectura* 20, 1–13. doi: 10.18239/ocnos_2021.20.3.2431
- Handayani, S., Youlia, L., Bunga Febriani, R., and Syafrudin, S. (2020). The use of digital literature in teaching reading narrative text. *J. Engl. Teach. Appl. Ling. Literatures* 3, 65–74. doi: 10.20527/jetall.v3i2.8445
- Kangasharju, A. I., Ilomäki, L., Toom, A., Lakkala, M., Kantosalo, A., and Toivonen, H. (2021). The digital poetry machine supporting lower secondary students' poetry writing. *J. Literacy Technol.* 22, 52–91.
- Kovalik, K., and Curwood, J. (2019). Poetryisnotdead: understanding Instagram poetry within a transliteracies framework. *Literacy* 53, 185–195. doi: 10.1111/lit.12186
- Llora, Á. (2019). “Escrituras punto cero: los espacios digitales para la escritura y su difusión en red,” in *Lectoescritura digital*. ed. D. Escandell (Madrid: Ministry of Education and Vocational Training CNIIE), 105–112.
- Morales-Sánchez, I. (2019). “La estética del lenguaje en el mundo digital,” in *Poesía en red y ciberpoesía*. eds. M. C. Fernández-Figares and D. E. Montiel.
- Morales-Sánchez, I., and Martín-Villarreal, J. P. (2019). Double-click rhetoric: rhetorical strategies of communication in the digital context. *Res. Rhetorica* 6, 1–16. doi: 10.29107/rr2019.1.1
- Navarro-Romero, R. M. (2021). *Una imagen, dos mil historias: reto creativo de microrrelato hipermedial*. *Escrituras enredadas*, 71–85. Iberoamericana Editorial Vervuert: Madrid.
- Pajares, S. (2015). Autoría y pedagogía digital. *Narrar en la era digital*, Especulo, 54. Available at: http://pendientedemigracion.ucm.es/info/especulo/Narrar_en_la_era_digital_Especulo_54_UCM_2015.pdf (Accessed February 18, 2022).
- Romero-López, D. (2012). De lo analógico a lo hipermedi. *Vías de entrada a la literatura digital*. (eds.) E. R. Alemany Ferrer and F. Chico Rico *Ciberliteratura i comparatisme*. Alicante: Universidad-SELGYC.
- Sánchez García, R., and Aparicio Durán, P. (2020). Los hijos de Instagram. Marketing editorial. Poesía y construcción de nuevos lectores en la era digital. *Contextos Educativos. Revista de Educación* 25, 41–53.
- Saum-Pascual, A. (2018). #Postweb!: crear con la máquina y en la red. Editorial Iberoamericana/Vervuert. Available at: <https://elibro.net/es/lc/udl/titulos/105242> (Accessed February 18, 2022).
- Saum-Pascual, A. (2019). “¿Qué es ciberpoesía?” in *Poesía en red y ciberpoesía*. eds. M. C. Fernández-Figares and D. E. Montiel.
- Selfa, M., and Azevedo, F. (2013). Poesía en castellano para la Educación Primaria: Algunas secuencias prácticas de trabajo con textos poéticos. *Ocnos* 10, 55–69. doi: 10.18239/ocnos_2013.10.03
- Soccavo, L. (2015). La disparition du livre est-elle inevitable? Viavooks. Retrieved from <http://www.viavooks.fr/article/la-disparition-du-livre-inevitable-lorenzo-occavo-52029>
- Solbes, J., and Traver, M. (2015). Science, scientists and literature. The role of literature in promoting science and technology. *Métode Sci. Stud. J.* 5, 89–97. doi: 10.7203/metode.82.3428
- Torres, R. (2013). Criação digital. Paper presented at the II Simpósio Internacional/VI Simpósio Nacional de Literatura e Informática. Universidade Federal de Santa Catarina/Núcleo de Pesquisas em Informática, Literatura e Linguística (NuPILL), Brasil. Available at: <http://www.telepoesis.net/confes.html> (Accessed February 18, 2022).
- Torres, R., and Côrtes, D. (2021). “Wreadings: digital poetry and collaborative practice,” in *The Community and the Algorithm: A Digital Interactive Poetics*. ed. A. Klobucar (Wilmington: Vernon Press), 3–26.

AUTHOR CONTRIBUTIONS

MS is responsible for the theoretical framework and the design of the research that we present in this research. EF is responsible for the practical implementation of digital literary reading practices in initial teacher training. All authors contributed to the article and approved the submitted version.

FUNDING

This research has been funded by the AGAUR (Agència de Gestió d'Ajuts Universitaris i de Recerca) under Grant 2020 ARMIF00020.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Selfa Sastre and Falguera García. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Image as Language: The Creation and the Use of the Visual Message by Young University Students in Their Communicative Social Activity

Alfredo Ramón-Verdú* and José Víctor Villalba-Gómez

Department of Plastic, Musical, and Dynamic Expression, University of Murcia, Murcia, Spain

OPEN ACCESS

Edited by:

Pedro García Guirao,
WSB Universities, Poland

Reviewed by:

Pilar Manuela Soto Solier,
Universidad de Granada, Spain
Pedro Victorio Salido López,
University of Castilla-La
Mancha, Spain

*Correspondence:

Alfredo Ramón-Verdú
alfredoramon@um.es

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 14 May 2022

Accepted: 10 June 2022

Published: 07 July 2022

Citation:

Ramón-Verdú A and
Villalba-Gómez JV (2022) The Image
as Language: The Creation and the
Use of the Visual Message by Young
University Students in Their
Communicative Social Activity.
Front. Psychol. 13:944187.
doi: 10.3389/fpsyg.2022.944187

Visual language, taken from the creation and reception of image perspectives, is ever-present in mediatized societies. With an interest in knowing what the experience of this is, a study is carried out in a university context, in the Faculty of Education at the University of Murcia, with the participation of 321 young students. The main objective of this study is to delve into the visual message, as a universal language in its productive and interpretative aspects, within a context marked by technology and the large-scale creation and use of images. The investigation is carried out by starting with a typical target performance, an individual, objective questionnaire which included three dimensions: the use of photographs, the use of graphic images, and the level of reflection and veracity of the image. A Principal Components Analysis (PCA) is carried out, which gives rise to a new conceptual organization after the Oblimin rotation based on the underlying variables. The data shows significant differences depending on the educational interests, a steady decrease in the use of images as ages increase, as well as less use of the image as a language among participants who are men. Men also make fewer requests for permission and transfer of permission, which are for the use of their own image by third parties. There is also evidence of a greater social value given to the image by women, as well as greater reflection and questioning of the message over the thirties, also by women.

Keywords: visual language, visual communication, image, interpretation of images, photography, visual narrative

INTRODUCTION

Photographic, analogical, and digital techniques are the result of obtaining lasting images and capturing the action of light on sensitive mediums. The very approximate representation of the reasons for, or the types of scenarios or contexts photographed, ultimately depends on the interpretation of the observer, conditioned by their own experience and culture. The images captured represent visual ideas that make up a language, a way of communicating, through creations with dialogic intentionality, on the basis of plastic resources similar to those of graphic design, as regards the compositive elements and the message, making a particular visual message act as a visual text, a transmitter of concepts and ideas.

The image, the visual culture, and the study of this in education is something which has been delved deeper into, on a large scale since the end of the XX and beginning of the XXI centuries,

with great emphasis on the artistic educational aspect for the comprehension of visual culture (Hernández, 1996). In this sense, the possibilities images and their communicative aims provide, depend on whether they are from a creative perspective or if they take on a receptive role in which the message is interpreted. In both cases, an idea is constructed from the particularities of visual language (Berger, 2005, p. 41). There has always been much interest in studying visual semantics and the mechanisms which encourage the use of images, which can be seen in studies such as those of Domínguez-Toscano (1996) and Ferradini and Tedesco-Prieto (1997), on the use of advertising images for teaching and didactic purposes. One of the areas of interest is based on the education of what is seen (Bustamante-Bohórquez et al., 2008), the critical reading of the image (Revuelta-Blanco, 2008; Ubreva-Amor, 2008), and studies of the analysis of the representations of reality (Aparici et al., 2009). With respect to this, we have seen that the image has become a tool in a pedagogy that transforms (Subirana, 2015; De-Andrés-del-Campo et al., 2016; Huerta, 2021a) when approached from a consumerist and creative point of view. From a didactic point of view, during the process of creating an image, narratives are created in cognitive development in processes such as those of language, creativity, reasoning, or organization, in different areas such as the denoted and the connoted, which have learning effects on the creating person and on the context in which it takes place. In this way, visual language takes up its position at a level of transporter of ideas and concepts which need a dialogue, reflection, and interpretation. This tendency has been completely renewed nowadays with the staggering advance in technology.

When looking at the creation of images as a language and as a detonating action in the communicative process, there are investigations in which the field of arts education has been studied more in-depth and in human capabilities (Dondis, 2002; Manifold, 2009; Marcellán-Baraze et al., 2013; González et al., 2015), when investigating visual literacy in order to improve students' competences (Mesquita-Romero et al., 2022). Using images is very much related to the expectations and interests of the individual, which limits their reading and interpretation, as these would have been previously conditioned, having a certain influence on the daily and professional tasks of those who use the images (Marcellán-Baraze and Agirre-Arriaga, 2008; Marzal-Felici and Soler-Campillo, 2011). Besides, there is another set of limitations that affects the interpretation and visual reading, as demonstrated in studies by Casper et al. (2018) and Tallon-Miles et al. (2021).

In the context in which these communicative competencies develop, both formal and non-formal spaces come together in the daily space (Marcellán et al., 2014). This means that part of the creative activity is going to take place outside the educational scope, being heavily influenced by the social context of development, where varied languages of expression are interrelated, which are assiduously reflected in the different artistic manifestations. The creation of visual experiences predominates here with a visual language that can be interpreted, based on fundamentally allegoric images.

Due to the surge in these phenomena and the overflow in the use and creation of images, new studies and didactic proposals

arise which try to give an answer to a subjectivity caused by a society that is completely mediatized, as can be seen in several papers such as those of Marcos-Dipaola (2019), Corrales (2018), Aguado and Villalba (2020), Ortega Caballero (2014), Estrada-Gonzalez et al. (2020), or Ramón-Verdú et al. (2022), where the creation of photographic images, in situations of sanitary confinement is studied in order to transmit visual messages about feelings and moods. This trend is in full progression, encouraged by communication of the masses, supported mainly by the use of the Internet and *a la carte* contents to be used, *via* platforms of audiovisual dissemination and those of videogames (Fanjul-Peyró et al., 2019), all of which exceed the individual's capacity of assimilation (Andrade-Vargas et al., 2021).

At this moment in time, in the themed context of visual education, with multiple studies and theories, on which this investigation centers, with the interest focused on young university students studying a teaching degree in primary education, young people who are in very close contact with the image, as both users and creators. Work by authors, such as Marín-Viadel (2005, 2011), Acaso López (2009); Acaso López et al. (2011), Hernández (2010), and Huerta (2021a,b), have been taken as reference, papers in which studies have been carried out for the development of language and visual culture in educational contexts, with the photographic and artistic image as a connecting thread. It is worth pointing out, at this stage, that in the future profession of the teacher training students who are the object of study in this paper, the image and language will be a first-level resource for them, and they will be able to use it to educate their students and identify a range of situations in them. With respect to this, the use of the image to discover aspects such as its use and creation is also investigated, placing emphasis on the level of reflection that takes place, and the grade of veracity that is given to the message in a system of dialogue and transmission, of both explicit and implicit ideas and messages. In this sense, the main objective is to learn what the relationship that students establish with visual language is—the use of photographic creation, examining this relationship from the students' facet of potential creators and also as a consumer of these images.

A social context that is so mediatized, as is the current one, encourages the dynamic exchange of personal stories being spread *via* technological devices. This means that the user has to be in contact with others in a flexible and dynamic way too, faced with a huge variety of graphic and visual dissemination means and formats, of highly manipulative messages created by the users themselves or by professionals with objectives that may not always be very clear. The creation and reading of visual messages are interpreted here as a need to confront visual language, as it is a transporter of ideas and concepts which affect people.

Investigation Questions

With the aim of delimitating the different elements of interest in the investigation, the study concentrated on: (i) the educational specialty studied, the gender and age of the surveyed people; (ii) the visual dialogue through the creation and use of images, with the follow-up interpretation and questioning about the veracity

of the message. With these nuclear questions as a base, the following investigation questions were drawn up:

- From the premise of the use of visual language by university students, and taking into account the educational specialty being studied, are there significant differences as regards the creation and use of visual messages? The aim is to delve deeper into the possible differences, depending on whether one has an interest in one particular type of study being undertaken or another.
- Given the gender of the students, is there a difference, with respect to the use and creation of related images which are given a visual language, and their social participation? The objective is to see if there is a difference in attitude between genders with respect to the social function of visual language through images.
- Due to the different ages in which studies related to education tend to be carried out, is there a difference between the age of the students and their interests when it comes to visual communication, whether this is for the creation, use, or interpretation of the messages? It is of interest to know if there are particular variations per age range when it comes to the creation or interpretation of the visual message.

MATERIALS AND METHODS

The study was carried out in two phases, the first with a transversal study based on a typical individual objective performance target questionnaire. In the second phase, a factorial Principal Components Analysis (PCA) was used to investigate the possible underlying variables there were in the questionnaire, which would give an answer to the fundamental questions of the investigation, which would then be followed by non-parametric tests, so that the hypothesis test would be validated, or not, which would show significant differences from which conclusions could be drawn.

Description of the Sample

The study group was made up using convenience sampling of a group of students from the subject of “Development of visual and plastic language” in the degree course of Primary school teacher training at the University of Murcia. The selection criteria were defined by non-probability incidental sampling (Vieytes, 2004), as the sample was chosen intentionally, given that there was easy access to the sample for a particular, temporary period of time (Sabariego and Bisquerra, 2009). From an annual number of 393 people, the sample ($N = 321$) represents 81.67% of the people in the fourth year of the previously mentioned degree course, for the academic year 2020–2021, a number considered to be a representative percentage. The sample was distributed among eight different degree specialties of the degree course: Audición y Lenguaje (AL), Educación Física (EF), Educación Intercultural (EI), Educación Musical (EM), Lengua Extranjera Francés (LEF), Lengua Extranjera Inglés (LEI), Pedagogía Terapéutica (PT), and Recursos Educativos para la escuela y el Tiempo Libre (RE).

As regards gender, there were 257 women in the sample group (80.1%), 62 were men (19.3%), and 2 from the gender other

(0.6%). In terms of age, the sample consisted of 297 between the ages of 20 and 25 (92.5%), 10 between 25 and 30 (3.1%), 5 between 30 and 35 (1.6%), and 9 older than 35 (2.8%). The distribution of the sample, depending on the specialty being studied was as follows: AL ($N = 28$; % = 8.7), EF ($N = 49$; % = 15.3), IE ($N = 38$; % = 11.8), EM ($N = 24$; % = 7.5), LEF ($N = 28$; % = 8.7), LEI ($N = 53$; % = 16.5), PT ($N = 52$; % = 16.2), and RE ($N = 49$; % = 15.3).

Instrument

Data was gathered by means of a questionnaire with an ordinal number scale or four levels, attributing values of 1 (not at all), 2 (some), 3 (quite a lot), and 4 (a lot). It was structured using four blocks of questions: In the first block ($N = 3$), sociological-type questions were asked with context variables (specialty, gender, and age). The second block investigated the use of the photographic image ($N = 8$), the third the graphic image ($N = 6$), and the fourth was used to obtain answers related to the interpretation of the images used ($N = 7$). The validation of the instrument was done by putting it before five experts in the area being investigated, using a template of qualitative evaluation with the areas of sufficiency, clarity, coherence, and relevance of the proposed questions (Urrutia et al., 2014). Several questions were modified, and the pilot test was carried out ($N = 35$; NC = 90%; Q = 5%), obtaining a uniform, coherent set of questions being dealt with (See instrument in **Supplementary Material**).

Procedure

The instrument was used on students in their fourth year of the degree, at a time considered ideal, given the intellectual maturity that could be attributed to them. The instrument was given to them *via* a link, before the teaching of the subject actually began, avoiding thus previous conditioning of the questions. Tacit consent was obtained from the participants once they had been informed of the procedure as well as the aim of the investigation, which was filled in freely within the allocated time. Once this had concluded, the data was downloaded to be later analyzed.

Data Analysis

Analysis was conducted with the program used for statistical analysis IBM SPSS Statistics 28 for Windows. The Kolmogorov-Smirnov test acted as a normality test, showing that the data do not follow a normal distribution, achieving a result of $p < 0.05$ in all variables. The reliability of the construct was checked with an analysis of the internal consistency with the Alfa de Cronbach test, giving coefficient reliability of $\alpha = 0.807$; $n = 20$, once questions of sociological background and a question of a qualitative nature have been excluded (P8).

The validity of the construct was performed using the factorial analysis of principal components (PCA), to check the feasibility of factorial analysis. By using the sample suitability test Kaiser-Meyer-Olkin (KMO = 0.801) and the Bartlett ($p < 0.005$) sphericity test, it was shown that existing correlations between items did not constitute a matrix identity, showing statistically significant data and therefore a null hypothesis could be ruled out (H0).

For component rotation, the oblique rotation method Direct Oblimin was used with a Kaiser normalizer, as it was considered that there was some correlation between variables. This action tends to reveal a structure that facilitates data interpretation (Lloret-Segura et al., 2014), besides showing evidence of the underlying latent variables, which can give way to a new distribution of the reactive of the instrument, so as to explain the relationships there are between the categorical context variables and the components extracted. The factorial scores were kept as regression variables.

RESULTS

The PCA revealed six components with their own values of > 1 , which explained the 60.053% of the accumulated variance, with a percentage of variance in each of the components as follows: C1 = 23.400; C2 = 11.217; C3 = 7.404; C4 = 6.68; C5 = 5.867; C6 = 5.485. After 10 iterations, the rotated Oblimin matrix component gave a new distribution, becoming conceptually configured depending on the function of the component and the reactive, in the order of the value of its variance in the following way: C1 = P1, P9, P11, P10, P2; C2 = P21, P16, P19, P20; C3 = P6, P5, P3, P7, P4; C4 = P18, P17; C5 = P14, P13, P12; C6 = P15.

The new distribution provided by the PCA shows a new alternative grouping, allowing for a new interpretative focus to be given to the instrument. This organization suggests a change in the initial naming, where C1 becomes: Use of professional and recreational visual messages, for grouping issues related to the use of images for their profession and personal recreational interests; C2: Grade of interpretation of the visual message, for grouping issues related to the interpretation of the visual message; C3: Grade of transmission of the image and one's own feelings, as they are related to the emission of visual messages of a personal nature; C4: Value of privacy of one's own and others' visual image, by grouping those related to giving or requesting permission to use images; C5: Grade of creation of professional and recreational visual messages, for grouping those related to the creation of visual messages related to their profession, RRSS and personal concerns; C6: Veracity of the visual message, related to the reality represented in the visual message.

In **Table 1** the descriptive statistics of the instrument can be observed after the recoding of variables. Component 4 stands out due to the marked dispersion of data ($SD = 1.01$) and particularly marked mesokurtic kurtosis ($g_2 = -0.88$), which could indicate, as a whole, diversity in the answers. Component 2 also stands out somewhat because it is the one with the least dispersion ($SD = 0.583$; $\sigma = -0.155$; $g_2 = 0.218$) with relatively low asymmetry and kurtosis.

Comparative Components-Variables of Context

To check if there were significant differences in each of the areas studied, comparatives were carried out between the components and the context variables, in order to accept or reject the H_0 null hypothesis: the distribution of component data was the same in all categories; H_1 : the distribution of component data was not

the same in all categories. This comparative was done with the non-parametric Kruskal-Wallis H -test for several independent samples ($\alpha < 0.05$; level of reliability = 95%), taking the six extracted components as test variables, and the context variables of Specialty, Gender, and Age as group variables.

As can be observed in **Table 2**, the H statistics of the components and the specialty group variable showed significant differences in Component 1 (p -value = 0.004; $X^2 = 20.921$), in Component 3 (p -value = 0.029; $X^2 = 15.579$), in Component 4 (p -value = 0.006; $X^2 = 19.946$) and in Component 5 (p -value = 0.021; $X^2 = 16.486$), therefore, in these mentioned components the null hypothesis (H_0) for each of the group variables indicated was rejected. The *post hoc* comparatives in Kruskal-Wallis in Component 1 showed significant differences between the specialties PT-EF ($t = -68.181$; sig. adjust. = 0.006), PT-LEI ($t = -63.963$; sig. adjust. = 0.012), and PT-EI ($t = -62.407$; sig. adjust. = 0.046). In Component 3, significant differences were found in the comparatives of the specialties EF-PT ($t = -58.39$; sig. adjust. = 0.044) and between EF and RE ($t = -65.204$; sig. adjust. = 0.014), differences which can be observed in the graph in **Figure 1**. This figure shows only the PT-EF comparisons because they coincide in components C1 and C3. Here, in C1, EF shows a lower dispersion of data and more negative values with respect to PT, as in C3, where PT also shows a greater positive trend and a lower dispersion that does not exceed value 2 in the graph.

As regards Component 4, a significant difference was seen in the comparison between the AL-RE ($t = -68.235$; sig. adjust. = 0.047), and finally, in Component 5 there was a significant difference between the specialties LEI-RE ($t = -60.016$; sig. adjust. = 0.031), all of which had significance values of < 0.05 . These differences can be observed in the graph in **Figure 2**, where C4 and C5 are shown in function of the specialties that these significant differences represent. Here we can see in C4 a clearly negative trend in AL with respect to RE, where positive values predominate, and in C5, a clearly greater grouping of LEI in negative values, with respect to RE, with much more dispersed values.

The statistics of the H -test of **Table 2** between the Components and the group variable Gender, showed significant differences in Component 1 (p -value = 0.016; $X^2 = 8.327$), and in Component 3 (p -value < 0.001 ; $X^2 = 28.875$) and in Component 4 (p -value = 0.035; $X^2 = 6.682$), thus, in the mentioned components, the null hypothesis (H_0) was rejected. In the *post hoc* analysis, in Component 1 significant differences could be seen between men and women ($t = 37.459$; sig. adjust. = 0.013), and in Component 3 of the same duality ($t = 69.403$; sig. adjust. = 0.000), and the same thing occurred in Component 4 ($t = -33.942$; sig. adjust. = 0.029), with the significance values level of all of these at < 0.05 , as can be observed in the diagrams in **Figure 3**, however, in the other gender, no variations are observed, because it is a very small sample, with data that may not be representative.

As for the statistics of the Kruskal-Wallis H -test, between the Components and the group variable Age (**Table 2**), significant differences showed up in both Component 1 (p -value = 0.003; $X^2 = 13.801$) and Component 3 (p -value < 0.001 ; $X^2 = 23.022$), and

TABLE 1 | Descriptive statistics of recoded variables.

Comp.	Name	No. items		Sd	σ	g_2
1	Use of professional and recreational visual messages	5	2.850	0.624	0.117	−0.506
2	Grade of interpretation of the visual message	4	2.841	0.583	−0.155	0.218
3	Grade of transmission of the image and one's own feelings	5	3.112	0.684	−0.205	−0.650
4	Value of privacy of one's own and others' visual image	2	2.866	1.011	−0.477	−0.880
5	Grade of creation of professional and recreational visual messages	3	2.489	0.707	0.199	−0.222
6	Veracity of the visual message	1	2.386	0.612	0.360	−0.029

TABLE 2 | Kruskal-Wallis *H*-test statistics: variables of test-group.

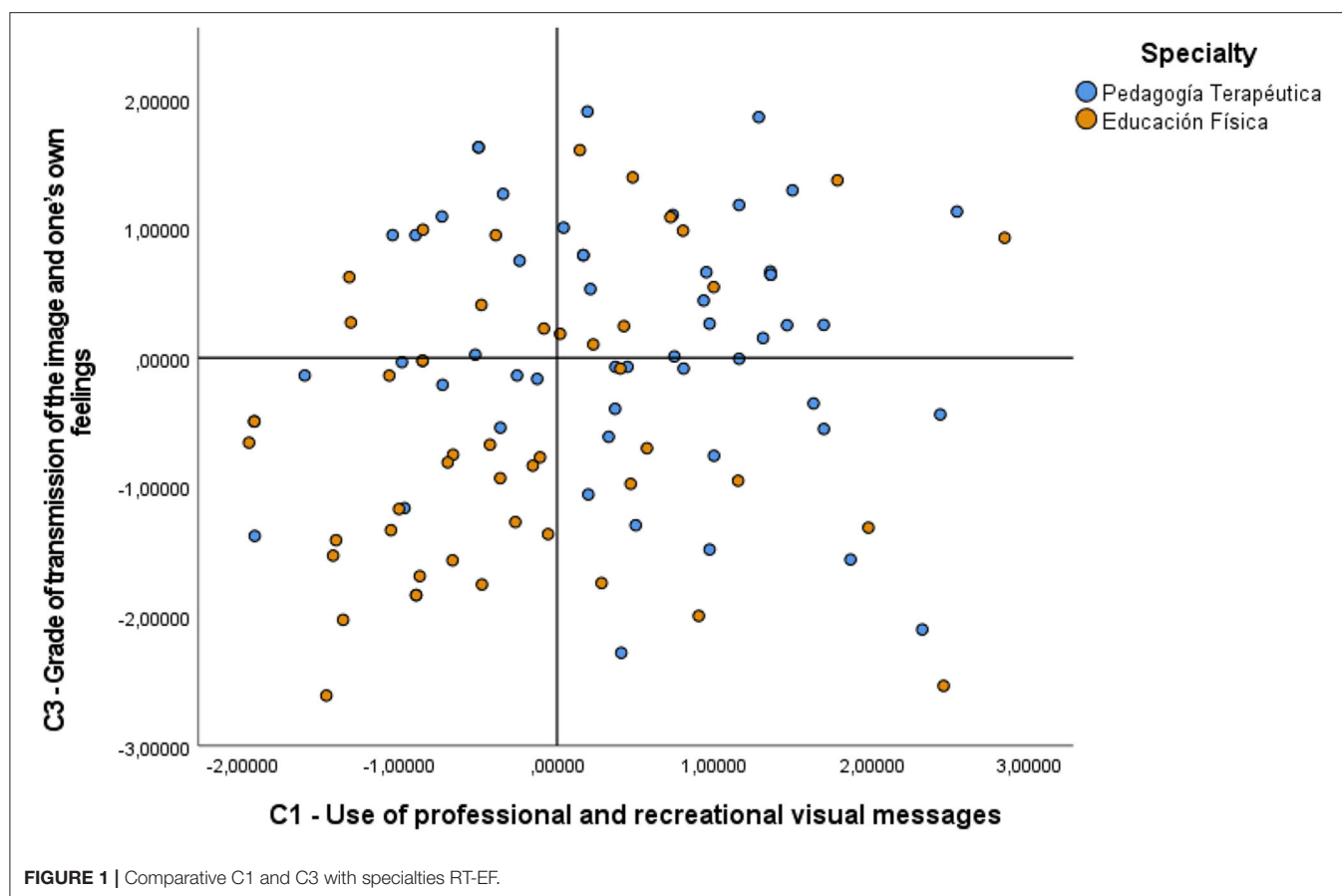
	Component					
	1	2	3	4	5	6
Group variable: speciality						
Chi-squared	20.921	4.047	15.579	19.946	16.486	10.205
gl	7	7	7	7	7	7
Asymptotic sig.	0.004	0.774	0.029	0.006	0.021	0.177
Group variable: gender						
Chi- squared	8.327	0.183	28.875	6.682	0.178	4.529
gl	2	2	2	2	2	2
Asymptotic sig.	0.016	0.912	<0.001	0.035	0.915	0.104
Group variable: age						
Chi- squared	13.801	4.908	23.022	1.864	1.159	3.876
gl	3	3	3	3	3	3
Asymptotic sig.	0.003	0.179	<0.001	0.601	0.763	0.275

as a result, the (H_0) null hypothesis is rejected in the mentioned components. In the *post hoc* analysis, Component 1 showed differences only between the age ranges 30–35 and 25–30 ($t = 186.1$; sig. adjust. = 0.002). In Component 3, the differences were found between two age ranges: Over 35 and 25–30 ($t = 126.022$; sig. adjust. = 0.019), and over 35 and 20–25 ($t = 134.731$; sig. adjust. = 0.000), taking, once again, as a level of signification, values of <0.05 . In **Figure 4** the distribution of the C1 and C3 values can be seen on the graph, in the function of the age ranges and the aforementioned significant differences. Visually, in C1 the negative values in the 30–35 age range clearly stand out, as opposed to the positive ones in the 25–30 age range. In C3, large differences are observed between those over 35, with clearly negative values, with respect to the 25–30 age range, mostly positive, as well as with the 20–25 age range, which, although they cover all the ranges of the graph, positive values with little dispersion predominate.

DISCUSSION AND CONCLUSION

As demonstrated by this and other studies mentioned, the consumption and creation of images in our society is a topic

of concern to numerous researchers, particularly those related to education, as is the case of the study by Ramón-Verdú et al. (2022), where fundamental aspects about their behaviors and communication through images in teachers in training are questioned, or also, the work of Aguado and Villalba (2020), where the consumption of illustrations in the classroom of Early Childhood and Primary Education is studied. However, in this sense, we did not find current studies that directly analyze the consumption and production of images in university students as future teachers in training. We did find other closely related research, such as the study by Andrade-Vargas et al. (2021), which analyzes the behavior and use of social networks in young people, where the image plays a priority role in the current communicative model, as well as the studies by Mesquita-Romero et al. (2022) and Marzal-Felici and Soler-Campillo (2011) where research is conducted on literacy and critical sense about the image coming from digital and media environments, as well as the consumption habits and uses of photography by students in the digital era. These are some of the references in which the research carried out is contextualized, analyzing behaviors and showing relevant and closely related data.

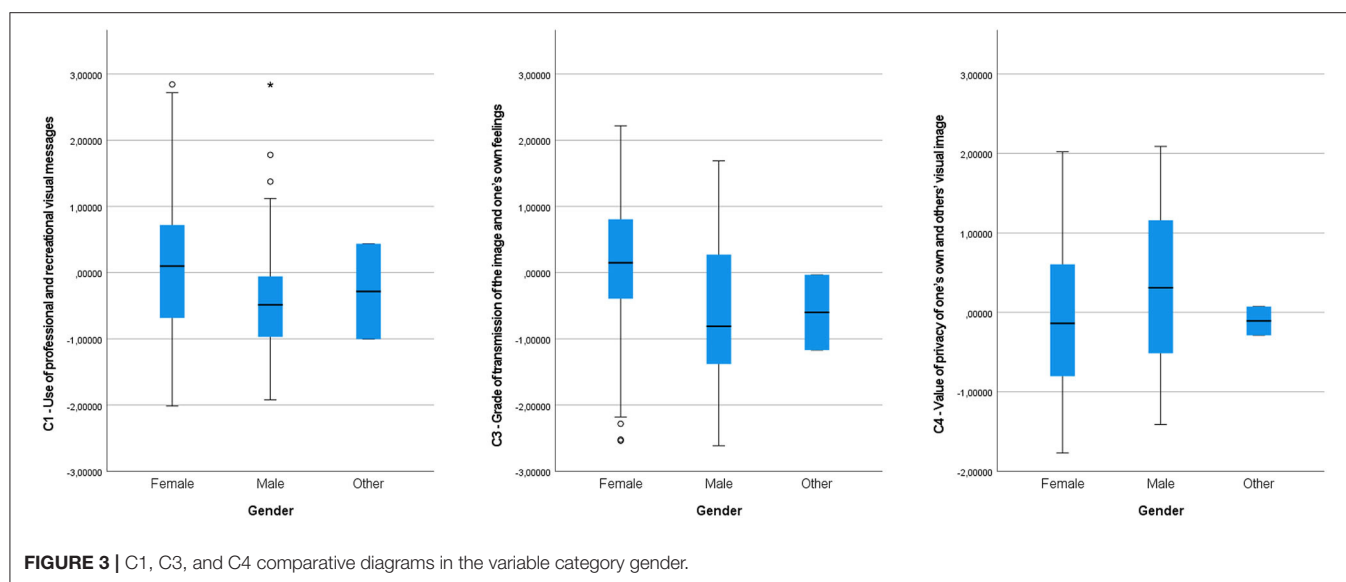
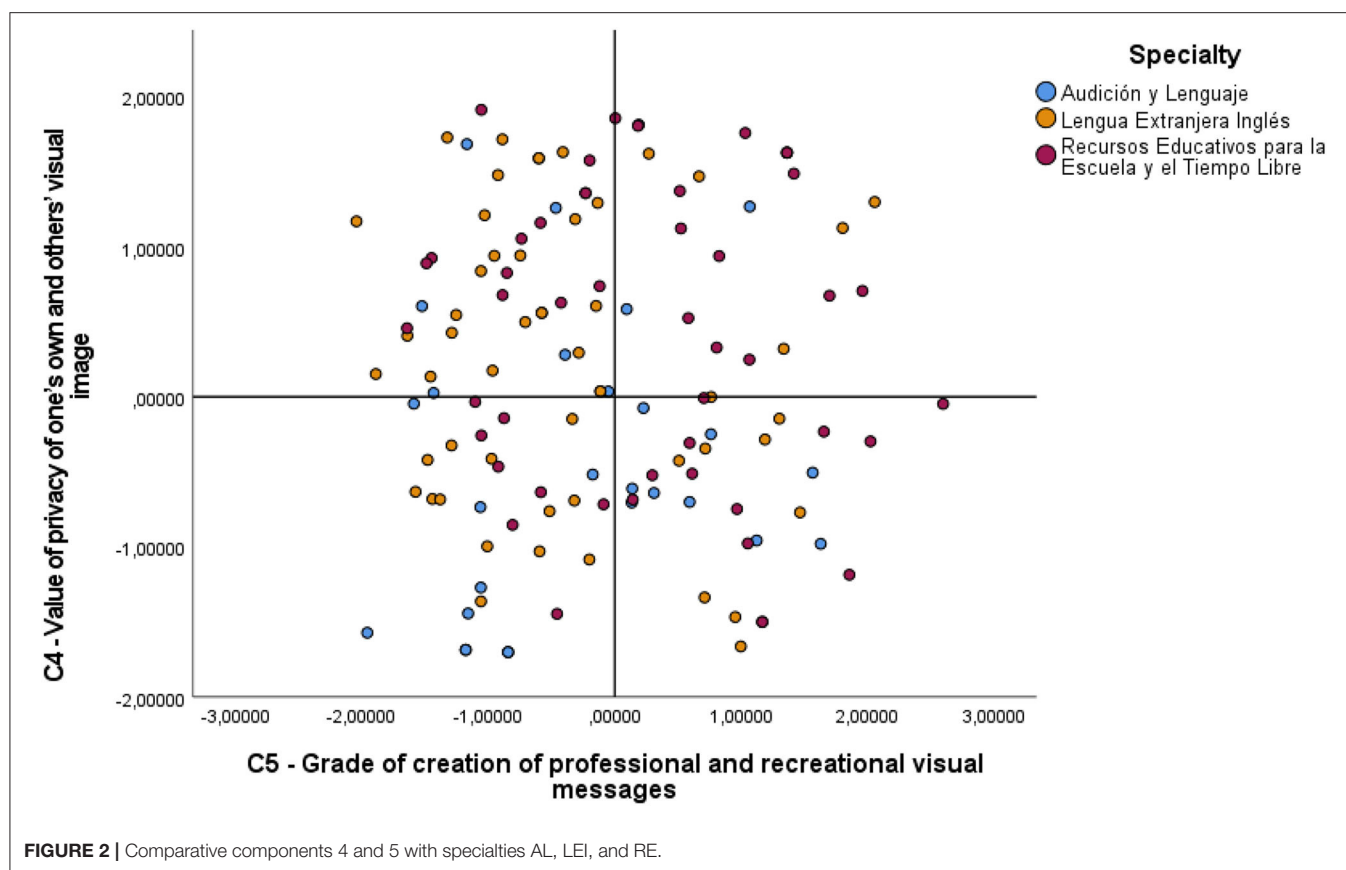


The data results of the analysis reveal certain behavioral differences depending on the premises raised at the beginning of this document. The tests carried out were done so in a special context due to, on the one hand, the social isolation that took place because of the world pandemic, and on the other, the university context in which the investigation was focused. This context has special characteristics as the students form part of a social network made up of educational professionals. This means that the use and limitations of visual language in contexts of development, where that which is visual is predominant, depends on the command one has of the area in question, and the importance lent to visual language. The capabilities as regard the mastery of visual language will very much constitute the basis for the construction of the professional personality of the future teacher, putting creativity and an artistic vision of education into practice. This is the domain in which the ability to dialogue through the use of images in digital spaces of social confluence is built and developed, places that abound with accurate visual information, but where there are also misleading messages, with very different aims.

With respect to the first premise made, the possible differences depending on the creation and use of images in professional profiles of the specialties, there are differences in several of the components extracted. As regards Component 1, and its relationship with the use of professional and recreational visual

messages, the specialty which stands out most is that of PT, as there are clear differences between this and the other specialties of mainly EI, LEI, and EF. As far as we are concerned, this is a reflection of greater interest by PT students in using images that act as support when intervening with children with special needs, given the overriding value that this specialty has to provide corrective solutions. It seems that the use of images on a professional level or for recreational purposes is of great help in teaching which is of a therapeutic nature. This, however, does not happen in the EF specialty, where there is least usage of visual messages on a professional or recreational level, due to, in our opinion, the fact that, in the field of sport, although images are used, these are of a documentary-type, more so than those that can be used as transmitters of messages. When it comes to the specialties that are not mentioned (LEF, EI, AL, and EM), there are no significant differences.

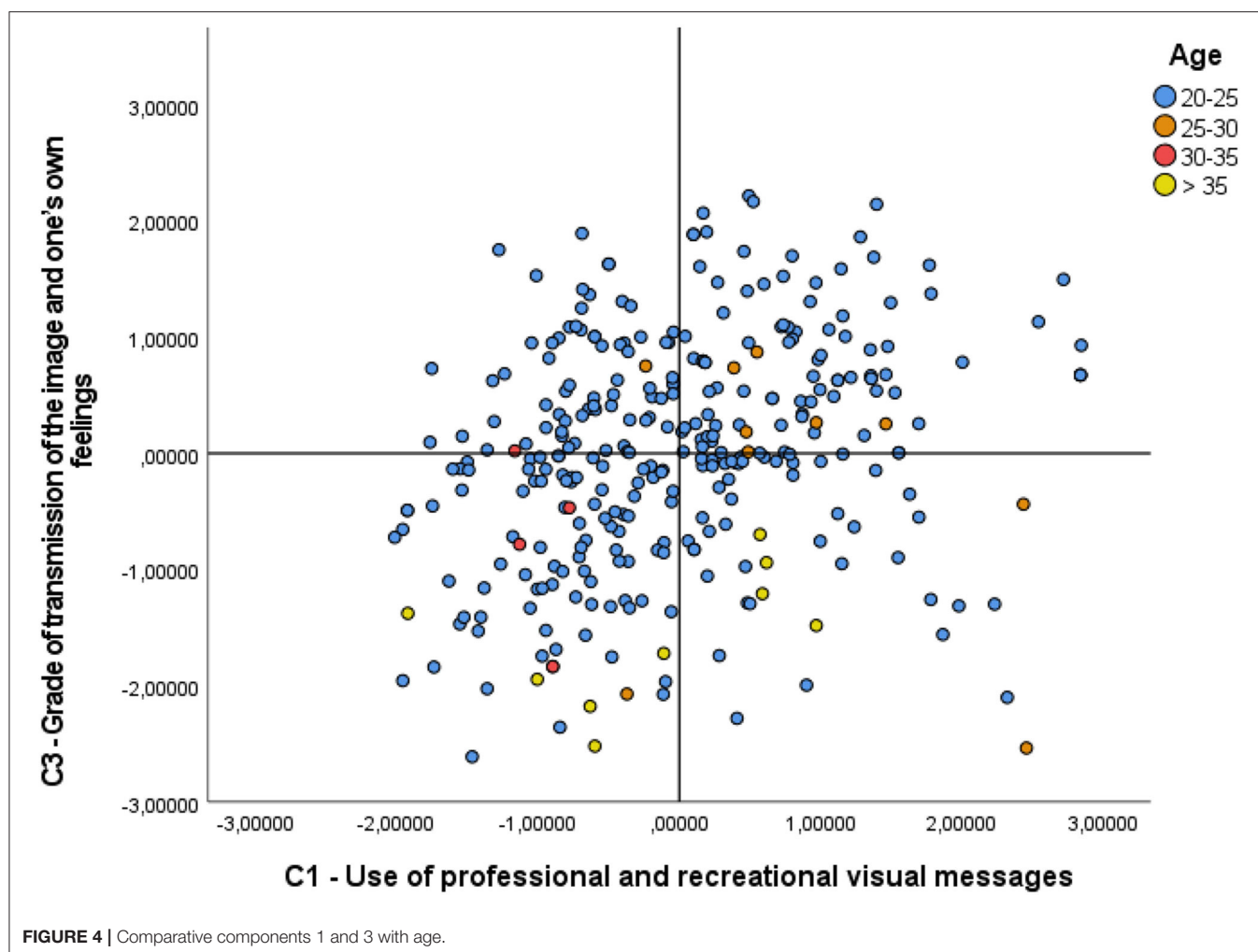
With respect to this first premise and Component 3, which is related to the grade of transmission of one's own feelings and the image, the differences shown stand out in EF as opposed to PT and RE, with those of EF being clearly less. It seems that students of this specialty are less interested in transmitting their own images or images which reflect feelings. Despite being a specialty that transmits values, this could be interpreted as that the students are less interested in transmitting feelings, less so than in other specialties, as the socialized physical improvement



dominates over other educational aspects, which are, nonetheless, also present. In the rest of the specialties, no particular differences worth pointing out were detected. Graphically-speaking, the different values arising from the Kruskal-Wallis H -test in the C3 comparative and the specialties, can be seen in **Figure 5**.

As for Component 4, which concerns the value given to the privacy of one's own and others' image, and the first premise,

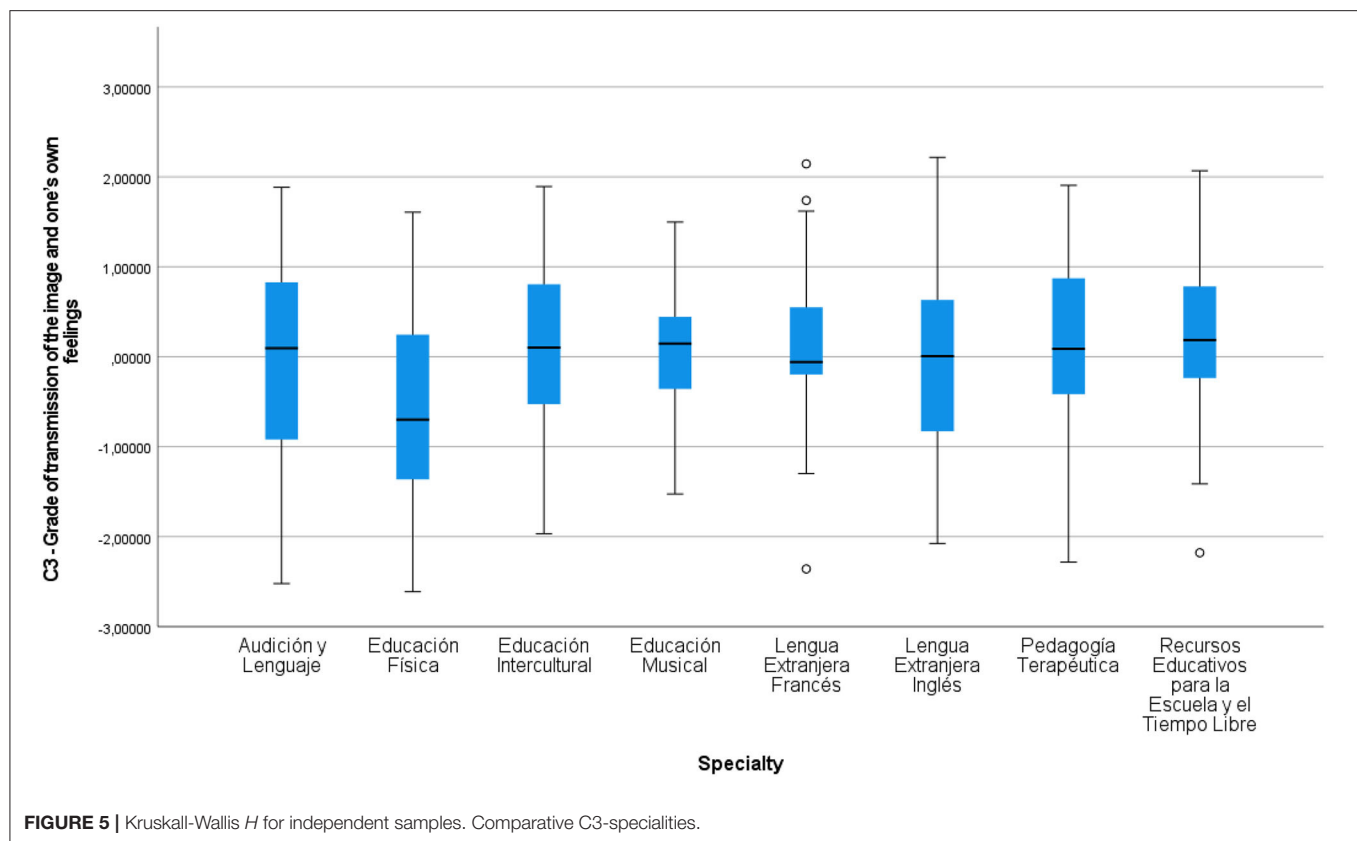
the only difference detected was that in the AL and RE duality. In the first specialty, it could be seen that privacy was valued a lot less than in the second specialty. We believe that this is due to the fact that in the second specialty, education is of a more recreational nature, where the most are made of the educational time available, which, obviously leads to graphic creations in situations that are highly expressive and shared.



Finally, in Component 5 and the level of creation of professional and recreational visual messages, the comparative which stands out most is that of LEI-RE, where, just like Component 4, the second specialty (RE), is where there is most participation, a lot more than the first. Once again, in the RE specialty, in which education takes place in a more recreational way, the creation of images takes a more important role, contrary to that in LEI.

With regards to the second premise, which refers to whether there are differences within genders in the attitude demonstrated toward the social function of visual language through images, differences were seen between Components 1, 3, and 4. The first, which refers to the use of professional and recreational images, is clearly superior in women to men. The level of transmission of one's own image and feelings is still less among men in comparison to women. It is our opinion that this may be due to the fact that the preconceptions of the masculine stereotype may still exist and the refusal to show feelings or to avoid comparisons about their own image, also. This fact may be in relation to the C4-Gender comparative, as in this case, it is the male gender that is more reserved when it comes to showing its own image or using that of others, placing greater value on privacy.

On analysis of the third premise, that of the variability of behaviors regarding the visual message depending on age, significant differences were shown in two of the components. In Component 1, this difference could be seen in the age ranges of 30–35 and 25–30. The first range revealed a lower level of use of visual messages of a professional or recreational level, as opposed to much higher use in the age range of 25–30 (Figure 5). It is our belief that this may be due to the maturity level of the participant and there being less interest in using recreational-type images to share on social networks, which is something that does not occur in the age range 25–30, this being the natural age of full social vibrancy. In Component 3 there are differences between the age ranges above 35 and 25–30, as well as above 35 and 20–25. In this case, the level of transmission of one's own image or feelings is extremely low. At this age, there is a natural loss of interest in achieving social recognition, and also, active participation in social networks has decreased enormously, where the need for recognition of the self in social groups that one forms part of, fades, as personal and intellectual maturity has normally been reached. Images that reflect moods, or “selfie” images in which one



directly shows one's physical aspect are quite marked in the age range 25–30, but even more so in the 20–25 age range, which is an age in which one actively looks for support and recognition, socially.

It is important to highlight the fact that Component 2: Level of interpretation of the visual message did not show significant disparities in any of the analyses carried out. This shows that the interest in the interpretation of the message transported by the images is very alike in all the context variables, showing a certain interest in the search for the meaning of the message. On the other hand, as regards Component 6: Veracity of the visual message, where the level of veracity attributed to the message is measured, even though there were no significant differences, the valuations are clearly of a negative kind. These results reaffirm the studies done by Beneyto (2021) or Marzal-Felici and Soler-Campillo (2011), in relation to the semantics of the image and the reality it shows, and the attribution of the value and lack of credibility of journalistic photography.

The results show, generally speaking, considerable differences when it comes to the use of the image as an instrument for transferring meanings. There are differences as regards ages, in the aspect of both the user and the creator of the images where the older age groups, although they value these aspects more, they use them less. There is also evidence among the genders, as men tend to be more reluctant to lend their image to others and to show feelings which, as far as we are concerned,

could show vulnerability, something that does not happen among women.

What this investigation has come up with constitutes the ratification and the starting point from which future educational topics of interest within the ecosystem of visual culture, something which is more and more present, and to a greater degree in a mediatized society, can be confronted. The differences detected between groups show the need to carry out future training action in youth, who can provide the mastery and control of what is visual, and where, more than ever, certain styles of communication that camouflage the truth by manipulation, are becoming the norm. We believe that visual language makes up one of the universal languages which need to be investigated further. The investigation of the meanings, messages, and formats is a real challenge nowadays, given the dynamicity and versatility that the Internet and social networks promote. In general terms, there must be a command of this language so that, from the point of view of future teachers, it can be established as a habitual, quality system of communication.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

- Acaso López, M. (2009). *La educación artística no son manualidades. Nuevas prácticas en la enseñanza de las artes y la cultura visual*. Madrid: Catarata.
- Acaso López, M., Hernández Belver, M., Nuere Menéndez, S., Moreno Sáez, M. C., Antúnez del Cerro, N., and Ávila Valdés, N. (2011). *Didáctica de las artes y la cultura visual*. Madrid: Akal.
- Aguado, M., and Villalba, M. (2020). La ilustración como recurso didáctico. (2020). *DEDiCA Revista de Educação e Humanidades* 17, 337–359. doi: 10.30827/dreh.v0i17.15158
- Andrade-Vargas, L., Iriarte-Solano, M., Rivera-Rogel, D., and Yunga-Godoy, D. (2021). Young people and social networks: between the democratization of knowledge and digital inequality. [Jóvenes y redes sociales: Entre la democratización del conocimiento y la inequidad digital]. *Comunicar* 69, 85–95. doi: 10.3916/C69-2021-07
- Aparici, R., Fernández-Baena, J., García-Matilla, A., and Osuna, S. (2009). *La imagen. Análisis y representación de la realidad*. Barcelona: Gedisa.
- Beneyto, R. F. (2021). El signo fotográfico. La relación semántica de la fotografía con la realidad. *Revista de Humanidades*. 43, 349–370. Available online at: <https://repositorio.unab.cl/xmlui/handle/ria/19658>
- Berger, J. (2005). *Modos de ver*. Barcelona: Gustavo Gili.
- Bustamante-Bohórquez, B., Aranguren-Díaz, F., and Chacón, M. (2008). Towards a democratic education of the look. [Hacia una educación democrática de la mirada]. *Comunicar* 31, 41–49. doi: 10.3916/c31-2008-01-005
- Casper, M., Nyakatura, J., Pawel, R., Schubert, T., and Lauschke, M. (2018). The movement-image compatibility effect: embodiment theory interpretations of motor resonance with digitized photographs, drawings, and paintings. *Front. Psychol.* 9:991. doi: 10.3389/fpsyg.2018.00991
- Corrales, A. (2018). Una aproximación a la imagen contemporánea: la educación como propuesta educativa interdisciplinar para una ciudadanía crítica. *Aularia*. 1, 21–28. Available online at: <https://www.aularia.org/ContadorArticulo.php?idart=305>
- De-Andrés-del-Campo, S., Nos-Aldas, E., and García-Matilla, A. (2016). La imagen transformadora. El poder de una fotografía para el cambio social: La muerte de Aylan. [La imagen transformadora. El poder de cambio social de una fotografía: La muerte de Aylan]. *Comunicar* 47, 29–37. doi: 10.3916/C47-2016-03
- Dominguez-Toscano, P. (1996). El trasfondo de la imagen. *Comunicar* 6, 123–128. doi: 10.3916/C06-1996-25
- Dondis, D. (2002). *La sintaxis de la imagen. Introducción al alfabeto visual*. Barcelona: Gustavo Gili.
- Estrada-Gonzalez, V., East-Scott, G., and Spehar, B. (2020). Viewing art in different contexts. *Front. Psychol.* 11:569. doi: 10.3389/fpsyg.2020.00569
- Fanjul-Peyró, C., González-Oñate, C., and Peña-Hernández, P. (2019). eGamers' influence in brand advertising strategies. A comparative study between Spain and Korea. [La influencia de los jugadores de videojuegos online en las estrategias publicitarias de las marcas. Comparativa entre España y Corea]. *Comunicar* 58, 105–114. doi: 10.3916/C58-2019-10
- Ferradini, S., and Tedesco-Prieto, R. (1997). Lectura de imagen. *Comunicar* 4, 157–160. doi: 10.3916/C08-1997-21
- González, M. R., Marcellán, I., and Arriaga, A. (2015). “Produciendo cultura visual: ¿qué se aprende dentro y fuera de la escuela?” in *II Congreso Internacional de Investigación en Artes Visuales ANIAV* (Valencia).
- Hernández, F. (1996). Educación artística para la comprensión de la Cultura Visual. *Curriculum. Revista de Teoría, Investigación y Práctica Educativa*. 12, 11–28.
- Hernández, F. (2010). *Educación y cultura visual*. Barcelona: Octaedro.
- Huerta, R. (2021a). *La imagen como experiencia*. Sevilla: Aula Magna.
- Huerta, R. (2021b). *Cementerios para educar*. Sevilla: Aula Magna.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., and Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Ann. Psychol.* 30, 1151–1169. doi: 10.6018/analesps.30.3.199361
- Manifold, M. C. (2009). What art educators can learn from the fan-based artmaking of adolescents and young adults. *Stud. Art Educ.* 50, 257–271. doi: 10.1080/00393541.2009.11518772
- Marcellán, I., Arriaga, A., Aguirre, I., Calvelhe, L., and Olaiz, I. (2014). “El papel de la escuela en la producción de cultura visual de los jóvenes,” in *Atas do I Congresso Internacional. Envolvimento dos Alunos na Escola: Perspetivas da Psicologia e Educação Veiga*, eds F. En Veiga. ISBN: 978-989-98314-7-6.
- Marcellán-Baraze, I., and Agirre-Arriaga, I. (2008). Mass media educative value: social correlations. [El valor educativo de los medios: correlaciones socioculturales]. *Comunicar* 31:46. doi: 10.3916/c31-2008-03-046
- Marcellán-Baraze, I., Calvelhe, L., Agirre, I., and Arriaga, A. (2013). Estudio sobre jóvenes productores de cultura visual: evidencias de la brecha entre la escuela y la juventud. *Arte Individuo y Sociedad* 25, 525–535. doi: 10.5209/rev_ARIS.2013.v25.n3.40752
- Marcos-Dipaola, E. (2019). Producciones imaginales: lazo social y subjetivación en una sociedad entre imágenes. *Arte Individuo y Sociedad* 31, 311–325. doi: 10.5209/ARIS.59483
- Marín-Viadel, R. (2005). “La Investigación Educativa basada en las Artes Visuales o Arte Investigación Educativa,” in *Investigación en educación artística*, ed R. Marín Viadel (Granada: Universidad de Granada), 223–274.
- Marín-Viadel, R. (2011). La Investigación en Educación Artística. *Educatio Siglo XXI* 29, 211–230. Available online at: <https://revistas.um.es/educatio/article/view/119951>
- Marzal-Felici, J., and Soler-Campillo, M. (2011). Consumption patterns and uses of photography in digital era among communication students. [Hábitos de consumo y usos de la fotografía en la era digital entre estudiantes de comunicación]. *Comunicar* 37, 109–116. doi: 10.3916/C37-2011-03-02
- Mesquita-Romero, W., Fernández-Morante, C., and Cebreiro-López, B. (2022). Critical media literacy to improve students' competencies. [Alfabetización mediática crítica para mejorar la competencia del alumnado]. *Comunicar* 70:4. doi: 10.3916/C70-2022-04
- Ortega Caballero, M. (2014). Proyecto de Intervención para mayores: La Imagen Fotográfica como agente psicoeducativo. *Publicaciones* 44, 39–56. Available online at: <https://revistaseug.ugr.es/index.php/publicaciones/article/view/3084>
- Ramón-Verdú, A. J., Villalba-Gómez, J. V., and Mascarell Palau, D. (2022). El selfi metafórico en situación de confinamiento. Análisis de la creación fotográfica de estudiantes universitarios. *ArtsEduca* 31, 148–162. doi: 10.6035/artseduca.6150
- Revueta-Blanco, J. (2008). Image Pedagogy: TV advertisement critical reading. [Pedagogía de la imagen: lectura crítica de publicidad televisiva]. *Comunicar* 31:58. doi: 10.3916/c31-2008-03-058
- Sabariego, M., and Bisquerra, R. (2009). “El proceso de investigación,” in *Metodología de la investigación educativa*, ed R. Bisquerra (Madrid: La Muralla), 89–125.
- Subirana, V. (2015). *La pedagogía transformadora*. Alcorcón: Sanz y Torres.
- Tallon-Miles, G., Wagner, E., and Frick, U. (2021). How do art skills influence visual search? – Eye movements analyzed with hidden markov models. *Front. Psychol.* 12:594248. doi: 10.3389/fpsyg.2021.594248

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.944187/full#supplementary-material>

- Ubrevia-Amor, A. (2008). Image as a daily learning opportunity. [La imagen: una oportunidad diaria para aprender]. *Comunicar* 31:50. doi: 10.3916/c31-2008-03-050
- Urrutia, M., Barrios, S., Gutiérrez, M., and Mayorga, M. (2014). Métodos óptimos para determinar validez de contenido. *Educación Médica Superior* 28, 547–558. Available online at: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412014000300014
- Vieytes, R. (2004). *Metodología de la investigación en organizaciones, mercado y sociedad. Epistemología y Técnicas*. Buenos Aires: Editorial de las Ciencias.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ramón-Verdú and Villalba-Gómez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

EDITED BY

Olivia López Martínez,
University of Murcia, Spain

REVIEWED BY

Javier Corbalán,
Universidad de Murcia, Spain
Cristina Del Moral Barrigüete,
University of Granada, Spain

*CORRESPONDENCE

Bai Li
libai@cpu.edu.cn

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 06 June 2022

ACCEPTED 04 July 2022

PUBLISHED 27 July 2022

CITATION

Li B (2022) Research on correlation
between English writing self-efficacy
and psychological anxiety of college
students.

Front. Psychol. 13:957664.

doi: 10.3389/fpsyg.2022.957664

COPYRIGHT

© 2022 Li. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Research on correlation between English writing self-efficacy and psychological anxiety of college students

Bai Li*

School of Foreign Languages, China Pharmaceutical University, Nanjing, China

English writing is not only a cognitive process for college students, it is also influenced by factors related to the emotional or psychological level. With the continuous improvement of the quality of university education cultivation, the pressure faced by college students has increased significantly, which has also given rise to general anxiety among college student groups. In order to investigate the correlation between college students' English writing self-efficacy and English writing psychological anxiety, we conducted a questionnaire survey on 595 current college students who were not majoring in English at a university in China, and used SPSS 25.0 software to make an empirical analysis of the collected data. The results showed that college students' English writing self-efficacy was at a moderate level, and they lacked self-confidence in their English writing skills and tended to complete writing tasks; college students generally had a moderate level of psychological anxiety in English writing, among which avoidance behavior was the most significant; college students' English writing self-efficacy was significantly negatively related to English writing psychological anxiety. English writing self-efficacy can reduce psychological anxiety level by enhancing college students' English writing motivation. Therefore, we suggest that we should set up English writing courses based on the process teaching method, change college students' English writing psychology in teaching, set reasonable writing goals, mobilize their enthusiasm and motivation in English writing, reduce college students' writing anxiety, and enhance their self-confidence in English writing.

KEYWORDS

English writing self-efficacy, English writing psychological anxiety, college students, English writing motivation, correlation

Introduction

With the development and the progress of society over time, English has gradually become a compulsory course for every Chinese student, which has been generally valued by schools and students at all levels. For non-English majors, they face a series of changes in learning requirements and learning patterns from high school English to

college English, and are especially prone to the problem of psychological anxiety in English writing. As a basic skill, English writing can not only express the values and emotional attitudes of college students, but also measure their comprehensive quality. According to Mcleod (1987), writing is not only a cognitive composition of English, but also an effective activity that is accompanied by an individual's emotional engagement. In recent years, more and more attention has been paid to the influence of emotional and psychological factors of college students on English writing. These studies have mainly focused on the relationship between college students' English writing self-efficacy and their writing level, writing ability, and writing achievement.

According to the latest requirements of Chinese English curriculum standards, one of the goals of English writing teaching is to enhance students' self-confidence and motivation in English writing. In order to achieve this goal, Chinese colleges and universities have set up corresponding writing tasks in English courses for college students. For example, college students are given writing tasks of practical value (such as invitations, notices, etc.) that they can complete independently, thus improving their writing skills and self-efficacy. Bandura (1977) put forward the concept of self-efficacy, which is defined as the individual's confidence and belief in his or her ability to complete a task. In 1986, Bandura (1986) went further and pointed out that self-efficacy is not only a judgment of one's own ability, but also includes expectations of one's own ability. Combining self-efficacy with English writing gives rise to the concept of English writing self-efficacy. English writing self-efficacy refers to the confidence of English writers in their ability to perform a particular English writing task (Mitchell et al., 2017). Research on English writing self-efficacy points out that English writing self-efficacy can improve an individual's English writing ability and writing achievement (Murdoch and Kang, 2019).

However, for Chinese non-English majors, writing in English is a very challenging task, which will be affected not only by positive factors (e.g., self-efficacy), but also by negative anxiety (Cheng, 2004; Woodrow, 2011). In the context of English writing, college students' psychological anxiety is regarded as a kind of trait anxiety, which is a relatively strong fear and avoidance psychology for English Writing (Daly and Wilson, 1983; Woodrow, 2011) and can seriously hinder the improvement of college students' English writing skills. A large amount of literature and research results confirm that Chinese college students have a relatively low level of English writing, have obvious fear of writing in English, generally lack motivation and self-confidence in writing, and have more obvious writing psychological barriers and anxiety (Woodrow, 2011; Zhang and Guo, 2012; Li et al., 2013).

Although there are many discussions on English writing self-efficacy in the existing literature (Sun et al., 2021; Zhang et al., 2021), most studies have focused on the current situation

of English writing among college students and the impact of English writing self-efficacy on improving college students' English writing performance, and there is lack of sufficient attention and empirical research on the relationship between English writing self-efficacy and psychological anxiety and the mediation of English writing motivation among Chinese college students. Based on this, the current study takes college students in a Chinese university as an entry point and focuses on the current situation of English writing psychological anxiety among Chinese college students, the current situation of Chinese college students' English writing self-efficacy, the correlation between Chinese college students' English writing self-efficacy and psychological anxiety, and the effect of English writing motivation in the relationship between English writing self-efficacy and psychological anxiety. By answering these questions, this study hopes to put forward targeted suggestions to alleviate writing psychological anxiety and improving English writing self-efficacy among Chinese college students on the basis of complementing existing research.

Literature review

English writing psychological anxiety

Psychological anxiety is an important component of anxiety, which corresponds to physiological anxiety. Generally, psychological anxiety refers to the emotional experience and behavioral performance of an individual in an anxious state. In terms of emotional experience, individuals under psychological anxiety may perceive fear, worry, nervousness, uneasiness, and annoyance, and may even have a sense of panic and imminent death. In terms of behavioral performance, individuals under psychological anxiety may experience fidgeting, facial tension, and sleepless nights (Li et al., 2021; Wang and Zhang, 2021). Focusing on the aspect of English writing, the term psychological anxiety can be extended to writing anxiety or English writing psychological anxiety. The term writing anxiety was first introduced by Daly and Miller in 1975, who argued that writing anxiety is an anxious behavior exhibited by writers in the aggregate of the writing process, such as avoiding writing task and worrying about the content of their writing being read or commented on by others. Furthermore, Daly and Miller (1975) proposed that writing anxiety is a higher level and more intense state of psychological anxiety or psychological fear, not a general state of psychological tension. When writers experience psychological anxiety, their attention to writing is distracted, and the quality of their writing decreases dramatically. Since then, this concept has attracted the attention of many scholars, and these studies cover the conceptual deepening and scale development of writing anxiety (Cheng, 2004), the current situation of foreign language writing anxiety among students in different countries, their coping strategies for writing anxiety

(Salam and Al Dyar, 2014; Torres et al., 2020; Qadir et al., 2021; Abolhasani et al., 2022), and the consequences of writing anxiety (Zhu, 2021). In this study, we no longer distinguish between psychological anxiety and writing anxiety, but rather in the form of psychological anxiety throughout the text.

English writing self-efficacy

The concept of English writing self-efficacy is derived from self-efficacy, which was introduced by Bandura (1977), who summarized the influencing results of self-efficacy in four aspects, which are determining the individual's choice and persistence in a certain activity, affecting the individual's attitude toward difficulties, affecting the acquisition and subsequent performance of new behaviors, and affecting the individual's emotion when engaging in a certain activity. Existing research has shown that individuals with this overall confidence are able to make good predictions about the problems they face, and can effectively cope with whatever difficulties they face (Jin et al., 2022). The concept of writing self-efficacy or English writing self-efficacy emerged when self-efficacy was developed in English writing. McCarthy et al. (1985) first put forward the concept of writing self-efficacy, defining it as an individual's perception and evaluation of his or her own writing skills. Pajares (2003) also combined self-efficacy with writing in his study and proposed a definition of writing self-efficacy, stating that writing self-efficacy is an individual's self-perceived judgment of the ability to use various writing knowledge and skills to accomplish different writing tasks. Furthermore, he also provided an in-depth analysis of writing self-efficacy, pointing out that writing self-efficacy can significantly promote the improvement of writing achievement, and may also be related to writing anxiety, achievement goals, and expected outcomes. In recent years, writing self-efficacy has received more and more attention, and the research has identified group differences and fluctuations in writing self-efficacy (Mendoza et al., 2022).

English writing motivation

Motivation is a necessary factor of language learning, the resource that language and language use are equally dependent on. Lowie et al. (2009) claimed that motivation is an adaptable interconnected factor but not a fixed one; high levels of it will lead to learning success. For example, Kormos and Csizér (2008) found that the idea that students' attitudes toward the role of English in the globalized world are important in English learning. Statistics indicated that Chinese college students vary in their motivation to learn English with variables, such as regional and urban/rural disparities, as well as educational discrepancy (You and Dörnyei, 2016). Writing is difficult

with motivational challenges and writing motivation plays a critical role in predicting writers' performance, shaping learners' writing experience, and performance (Wright et al., 2019). Previous research showed that more motivated students with past successful experiences in writing are more likely to improve their writing, while college students with low motivation are more likely to procrastinate writing (Hayes, 1996; Fritzsche et al., 2003; Wright et al., 2019).

English writing self-efficacy and psychological anxiety

In terms of the relationship between English writing self-efficacy and psychological anxiety among college students, the present study holds that there is a negative influence on each other. On the one hand, from the perspective of the differential effects of the two on college students' English writing performance and writing improvement, English writing-related studies have found that psychological anxiety was significantly and negatively related to writing performance and writing improvement. Students with higher levels of psychological anxiety performed worse in writing performance than those with lower levels of psychological anxiety (Horwitz, 2001; Li et al., 2013). However, the relevant studies focusing on the relationship between writing self-efficacy and writing achievement have reached the opposite conclusion. These studies showed that there was a significant positive correlation between English writing self-efficacy and writing achievement (Pajares, 2003; Woodrow, 2011). Thus, it can be seen that college students' writing self-efficacy has a negative relationship with psychological anxiety. On the other hand, college students with a high sense of self-efficacy in English writing mean that individuals have high confidence in their ability to successfully complete their writing task. They will be subject to self-reinforcement (observation, feedback, self-reward, etc.) during the English writing process, and this reinforcement effect will directly serve as important motivational factors to enhance college students' English writing, such as effort, commitment, persistence, strategy use, ability attribution, etc., thus promoting them to better complete the writing task. As a result, they will have less writing anxiety and less writing avoidance behavior, and their psychological anxiety level will naturally decrease (Zeng et al., 2020). Accordingly, the present study suggests that there is a negative relationship between college students' English writing self-efficacy and psychological anxiety.

Furthermore, college students' English writing self-efficacy can alleviate their psychological anxiety by enhancing their English writing motivation. Gardner and Lambert (1972) classified learning motivation into instrumental learning motivation and integrative learning motivation. The former refers to learning a language and using it as a tool to achieve a

practical purpose, while the latter refers to learning a language and viewing it as an activity to understand and integrate into the culture. However, more research has divided learning motivation into intrinsic motivation and extrinsic motivation. Intrinsic motivation corresponds to integrative motivation, which is to derive pleasure and satisfaction from language learning activities, while extrinsic motivation corresponds to instrumental motivation, which is to take language learning as a way to achieve a certain goal. Existing studies have linked learning motivation with self-evaluation, learning strategies, and psychological anxiety. In the present study, we believe that college students with high self-efficacy in English writing will enhance their writing activity and enthusiasm to participate in English writing, which significantly stimulates their learning motivation, encourages them to put more efforts into English writing, and thus overcoming difficulties and reducing psychological anxiety. Accordingly, we suggest that college students' English writing motivation plays a mediating role between writing self-efficacy and psychological anxiety.

Research methodology

Research questions

Understanding the developmental level of English writing self-efficacy and the status quo of psychological anxiety among Chinese college students, and presenting the relationship between the two, will be of reference and guidance to both the educational activities of Chinese college teachers and the learning activities of college students. The research questions in this research include:

- (1) What is the status quo of English writing self-efficacy among Chinese college students?
- (2) What is the status quo of English writing psychological anxiety among Chinese college students?
- (3) Is there a correlation between English writing self-efficacy and psychological anxiety among Chinese college students? What is the correlation?
- (4) Can Chinese college students' English writing self-efficacy reduce their psychological anxiety by enhancing their writing motivation?

Participants

In this study, 595 current college students who were not English majors at a Chinese university were randomly selected as the research participants, and four upper and lower grades were also sampled to find out the overall status of English writing self-efficacy and psychological anxiety. There are two reasons for this selection: first, college students in each grade have different

English learning times, different English learning strategies, different abilities to use different strategies to complete English writing tasks, and different writing knowledge and writing skills. Therefore, we randomly select college students from the first to the fourth year of college as the research subjects, which can better enhance the accuracy and comprehensiveness of this study. Second, the selection of college students in four grades as the research subjects can reflect the status quo of English writing self-efficacy and psychological anxiety of Chinese college students more comprehensively, and reflect the overall state and characteristics of English writing of Chinese college students. The research results can better provide a reference for the design of college English writing teaching.

Instruments

English writing self-efficacy questionnaire

This study used the writing self-efficacy scale developed by Jones (2008) to measure the status quo of Chinese college students' English writing self-efficacy, while we made adjustments to some of the textual expressions in the scale to form the Chinese college students' English writing self-efficacy scale. This scale includes English writing task self-efficacy and English writing skill self-efficacy. English writing task self-efficacy refers to college students' judgments about their ability to successfully complete a certain writing task. English writing skill self-efficacy refers to college students' judgments about their ability to successfully complete various English writing skills. The task self-efficacy subscale has 10 items, which mainly measure general writing tasks and application writing tasks. The sample items are "I believe I can write an email in English to introduce my school to my foreign friends" and "I believe I can write an English composition of at least 80 words within the specified time according to the topic requirements." The skill self-efficacy sub-scale has 10 items, which mainly measures expressive writing skills, organizational writing skills, stylistic writing skills, and revision writing skills. The sample items are "I believe I can spell all words in an English composition correctly and use punctuation correctly," "I believe I can finish a writing task within the specified time according to the topic requirements," "I believe I am good at writing different subjects in English, such as narratives essays, argumentative essays, and letters," and "I think I can find out the advantages and disadvantages and correct the shortcomings after finishing my English composition."

English writing psychological anxiety questionnaire

The present study used the foreign language writing anxiety scale revised by Guo and Qin (2010) to measure the status quo of English writing psychological anxiety among Chinese college students based on Cheng (2004) study. We adjusted some of

the expressions in the original scale to form the Chinese college students' English writing psychological anxiety scale. This scale includes project teaching anxiety, ideation anxiety, avoidance anxiety, and confidence anxiety. The project teaching anxiety sub-scale has five items, and a sample item is "When I write an English composition, I will feel nervous and uneasy if I know that the teacher will review it." The ideation anxiety sub-scale has five items, and a sample item is "My mind stops spinning when I am asked to write a time-limited English composition without preparation." The avoidance anxiety sub-scale has five items, and a sample item is "I usually don't write in English unless I have no other choice." The confidence anxiety sub-scale has five items, and a sample item is "I will worry about getting a very low score when my English composition is reviewed."

English writing motivation questionnaire

This study modified and adjusted some of the expressions to form the Chinese college students' English writing motivation scale based on the English learning motivation questionnaire developed by Gao et al. (2003). This scale includes two sub-scales: intrinsic motivation sub-scale and extrinsic motivation sub-scale. The intrinsic writing motivation scale has seven items, and a sample item is "The reason why I finish my English writing is my love for English." The extrinsic writing motivation scale has 15 items, and a sample item is "The reason why I finish my English writing is to get good grades so as to get a good job."

Reliability and validity tests

We examined the reliability and validity of the Chinese college students' English writing self-efficacy scale, the Chinese college students' psychological anxiety scale, and the Chinese college students' English writing motivation scale. Unless otherwise noted, responses to all items were measured on five-point Likert-type scales.

The Chinese college students' English writing self-efficacy scale consists of 20 items, ranged from strongly disagree (1) to strongly agree (5). The mean score of all 20 items was the overall English writing self-efficacy level of college students. We conducted a confirmatory factor analysis on the whole scale. The KMO value was 0.890, the Bartlett's sphere test reached a significant level ($p < 0.001$), and the factors with characteristics greater than 1 explained 78.961% of the total variance. Exploratory factor analysis extracted two factors, namely, the English writing task self-efficacy sub-scale and the English writing skill self-efficacy sub-scale. The internal consistency coefficient of the total scale was 0.853, the internal consistency coefficient of factor 1 was 0.901, and the internal consistency coefficient of factor 2 was 0.804. Accordingly, the Chinese college students' English writing self-efficacy scale had high reliability and internal validity consistency, and was suitable to be used as a research tool for the English writing self-efficacy of Chinese college students.

The Chinese college students' English writing anxiety scale consists of 20 items, ranged from strongly disagree (1) to strongly agree (5). The mean score of all 20 items was the overall English writing psychological anxiety level of college students. We conducted a confirmatory factor analysis on the whole scale. The KMO value was 0.911, the Bartlett's sphere test reached a significant level ($p < 0.001$), and the factors with characteristics greater than 1 explained 80.738% of the total variance. Exploratory factor analysis extracted four factors, namely, the English writing project anxiety sub-scale, the English writing ideation anxiety sub-scale, the English writing avoidance anxiety sub-scale, and the English writing confidence anxiety sub-scale. The internal consistency coefficient of the total scale was 0.883, the internal consistency coefficient of factor 1 was 0.897, the internal consistency coefficient of factor 2 was 0.865, the internal consistency coefficient of factor 3 was 0.937, and the internal consistency coefficient of factor 4 was 0.900. Accordingly, the Chinese college students' English writing psychological anxiety scale had high consistency reliability and internal validity, and was suitable to be used as a research tool for English writing psychological anxiety of Chinese college students.

The Chinese college students' English writing motivation scale consists of 22 items, ranged from strongly disagree (1) to strongly agree (5). The mean score of all 22 items was the overall English writing motivation level of college students. We conducted a confirmatory factor analysis on the whole scale. The KMO value was 0.934, the Bartlett's sphere test reached a significant level ($p < 0.001$), and the factors with characteristics greater than 1 explained 82.974% of the total variance. Exploratory factor analysis extracted two factors, namely, the English writing intrinsic motivation sub-scale and the English writing extrinsic motivation sub-scale. The internal consistency coefficient of the total scale was 0.874, the internal consistency coefficient of factor 1 was 0.869, and the internal consistency coefficient of factor 2 was 0.883. Accordingly, the Chinese college students' English writing motivation scale had high reliability and internal validity consistency, and was suitable to be used as a research tool for the English writing motivation of Chinese college students.

Questionnaire distribution and data analysis

A questionnaire survey was conducted among college students in a Chinese university at the beginning of March 2022. In order to ensure the validity of the data, an electronic version of the questionnaire was sent to each college student after confirming that they fully understood the purpose and relevant requirements of the survey. A total of 700 surveys were distributed and 650 surveys were collected, of which 595 were effective, with an effective response rate of 85%.

We used SPSS 25.0 for data analysis. First, descriptive statistical analysis was applied to investigate the development level of college students' English writing self-efficacy and the status quo of college students' English writing psychological anxiety. Second, correlation analysis was applied to verify the correlation between English writing self-efficacy and psychological anxiety among college students. Finally, structural equation modeling (SEM) was used to test the mediating role of English writing motivation between college students' English writing self-efficacy and psychological anxiety.

Results and discussion

The status quo of college students' English writing self-efficacy

As shown in [Table 1](#), the results of descriptive statistical analysis of Chinese college students' English writing self-efficacy show that the scores of college students' overall English writing self-efficacy ($M = 3.0619$, $SD = 0.5978$), English writing task self-efficacy ($M = 3.1250$, $SD = 0.6184$), and English writing skill self-efficacy ($M = 2.9987$, $SD = 0.6033$) are not high, and all lower than 3.5. According to [Pajares \(1996\)](#), the mean value of English writing self-efficacy is low between 1.0 and 1.5, medium between 1.5 and 3.5, and high between 3.5 and 5.0. It can be seen that Chinese college students' English writing self-efficacy is at a moderate level. Comparing the means of English writing task self-efficacy with skill self-efficacy, we can find that the mean of English writing task self-efficacy is higher than that of skill self-efficacy, which indicates that Chinese college students are essentially more confident in English writing tasks despite the fact that they have mastered certain English writing skills after years of writing practice. It also means that Chinese college students are more inclined to regard English writing as a task and lack confidence in their own English skills. In addition, the SD of college students' overall English writing self-efficacy, English writing task self-efficacy, and English writing skill self-efficacy was less than 1, which indicates that although the overall level of English writing self-efficacy among Chinese college students is not high, the development level is relatively balanced.

Furthermore, this study conducted a descriptive analysis of each relevant dimension of college students' English writing

self-efficacy. Writing task self-efficacy includes general writing task self-efficacy and applied writing task self-efficacy. General English writing task self-efficacy refers to college students' judgment of their ability to complete general writing tasks, while applied writing task self-efficacy refers to college students' judgment of their ability to complete practical writing tasks. The mean of general writing task self-efficacy is 3.1528, and the mean of applied writing task self-efficacy is 3.3325, indicating that Chinese college students are more confident in the task of applied writing category, which may be related to the fact that college students pay more attention to the practical value of English writing more and have received more training in applied writing. For example, college students generally score higher in writing invitation letters and introduction letters. Writing skill self-efficacy includes four aspects: expressive writing, organizational writing, genre writing, and revision writing. The mean of expressive writing skill self-efficacy is significantly higher than the other three, which indicates that Chinese college students show more confidence in test-taking abilities, such as spelling words and using punctuation marks correctly, judging wordiness, and quickly examining questions. It may be related to their college entrance examination experience. According to the above clues, college teachers can select materials appropriately and flexibly when teaching English writing, which is not limited to the topics with high applicability, but also can be broadened to more general genres, so as to promote the all-round development of college students' English writing ability while maintaining their self-confidence in application-based essays.

The status quo of college students' English writing psychological anxiety

As shown in [Table 2](#), the results of descriptive statistical analysis of Chinese college students' English writing psychological anxiety show that except for project teaching anxiety ($M = 2.7912$, $SD = 0.7862$) scored low, college students' overall English writing psychological anxiety ($M = 3.2852$, $SD = 0.6831$), English writing ideation anxiety ($M = 3.0313$, $SD = 0.8277$), English writing avoidance anxiety ($M = 3.7745$, $SD = 0.5154$), and English writing confidence anxiety ($M = 3.5438$, $SD = 0.6396$) scored higher 3.0. English writing avoidance anxiety scored the highest with over 3.5. This shows that Chinese college students' English writing anxiety is generally at a moderate level. Despite the fact that they have mastered certain English writing skills after many years of writing practice, Chinese college students are still essentially more confident in English writing tasks. It also indicates that Chinese college students tend to see English writing as a task and lack confidence in their own English skills.

We further conducted a descriptive analysis of each relevant dimension of college students' psychological anxiety. Among the four dimensions of English writing psychological anxiety,

TABLE 1 Results of descriptive analysis of English writing self-efficacy.

	N	Item	Min	Max	Mean	Standard deviation
English writing self-efficacy (EWS)	595	20	1.26	4.86	3.0619	0.5978
English writing task self-efficacy (EWTS)	595	10	1.27	4.98	3.1250	0.6184
English writing skill self-efficacy (EWSS)	595	10	1.19	4.81	2.9987	0.6233

TABLE 2 Results of descriptive analysis of English writing psychological anxiety.

	N	Item	Min	Max	Mean	Standard deviation
English writing psychological anxiety (EWPA)	595	20	1	5	3.2852	0.6831
English writing project teaching anxiety (EWPTA)	595	5	1	5	2.7912	0.7862
English writing ideation anxiety (EWIA)	595	5	1	5	3.0313	0.8277
English writing avoidance anxiety (EWAA)	595	5	2.23	5	3.7745	0.5154
English writing confidence anxiety (EWCA)	595	5	1.63	5	3.5438	0.6396

TABLE 3 Results of the correlation analysis between English writing self-efficacy and psychological anxiety.

		EWPA	EWPTA	EWIA	EWAA	EWCA
EWS	<i>r</i>	−0.287**	−0.257**	−0.266**	−0.283**	−0.300**
EWTS	<i>r</i>	−0.269**	−0.191*	−0.200**	−0.336**	−0.295**
EWSS	<i>r</i>	−0.301**	−0.244**	−0.276**	−0.332**	−0.299**

** $p < 0.01$, * $p < 0.05$.

Chinese college students' avoidance anxiety is particularly serious, followed by confidence anxiety, while project teaching anxiety is relatively less serious, indicating that Chinese college students show high levels of anxiety in English writing avoidance anxiety and confidence anxiety. Facing English writing, most college students choose to avoid and lack confidence in their English writing, which may be due to the fact that Chinese college students have less daily exposure to the topic of English writing and are not interested in English writing. This enlightens that teachers in university need to reduce the assignment of writing practice tasks during English writing classes, interact more with students, and enhance college students' interest in English writing, while focusing on college students' progress in the process of English writing rather than only on their English writing scores. Project teaching anxiety and ideation anxiety belong to low-anxiety levels, and the possibility of triggering college students' anxiety is relatively small. However, in our further analysis, we also found that college students still experience writing stress, nervousness, and blankness, especially when they are faced with the requirement to complete the English writing tasks within the specified time. Project teaching anxiety mainly occurs when college students are unable to respond to the teacher's questions. It is an anxiety phenomenon arising from the protection of self-image. This suggests that teachers in university should relax time as much as possible when assigning writing tasks. They also can set up study groups

to give students the opportunity to brainstorm before writing, and ask divergent questions to encourage students to dare to answer, rather than focusing on whether the answers are correct.

Correlation between college students' English writing self-efficacy and psychological anxiety

To further understand the relationship between Chinese college students' English writing self-efficacy and psychological anxiety, the present study used the correlation coefficient to analyze the correlation between English writing task self-efficacy, skill self-efficacy, and four types of psychological anxiety.

As shown in **Table 3**, the simple correlation coefficients between English writing self-efficacy and psychological anxiety were -0.287 . The two dimensions of English writing self-efficacy and psychological anxiety were -0.269 and -0.301 , respectively. These results indicated that there was a negative correlation between English writing self-efficacy and psychological anxiety. The higher the college students' English writing efficacy, the lower the level of psychological anxiety. Furthermore, we also calculated the correlations between different dimensions of English writing self-efficacy and different dimensions of psychological anxiety, and the results showed that there was a significant negative between different dimensions of both.

We applied a combination of path analysis and SEM to verify the mediating role of English writing motivation in the relationship between English writing self-efficacy and psychological anxiety among college students. First, the predictive effect of English writing self-efficacy on psychological anxiety was examined. According to the results of path analysis, the standardized coefficient between English writing self-efficacy and psychological anxiety was -0.229 , with a standard error of 0.064 , $t = -4.138$, and a significance level of $p = 0.000$ (< 0.001), which indicates that writing self-efficacy has a significant impact on psychological anxiety, and for each unit increase in writing self-efficacy one unit, psychological anxiety decreased by 0.229 units.

We applied SEM to examine the mediating effect of writing motivation. The results showed that theoretical model ($\chi^2/df = 3.025$, RMSEA = 0.060 , CFI = 0.917 , TLI = 0.905 , SRMR = 0.058) fitted the data well. Meanwhile, the two dimensions loads of writing self-efficacy were above 0.920 , the four dimensions loads of psychological anxiety were between 0.700 and 0.950 , and the two dimensions loads of writing motivation were above 0.800 (see **Figure 1**). In terms of the standardized coefficients, the effect of writing self-efficacy on writing motivation was significantly positive ($\beta = 0.357$, $p < 0.001$), the effect of writing motivation on psychological anxiety was significantly negative ($\beta = -0.574$, $p < 0.001$), and the effect of writing self-efficacy on

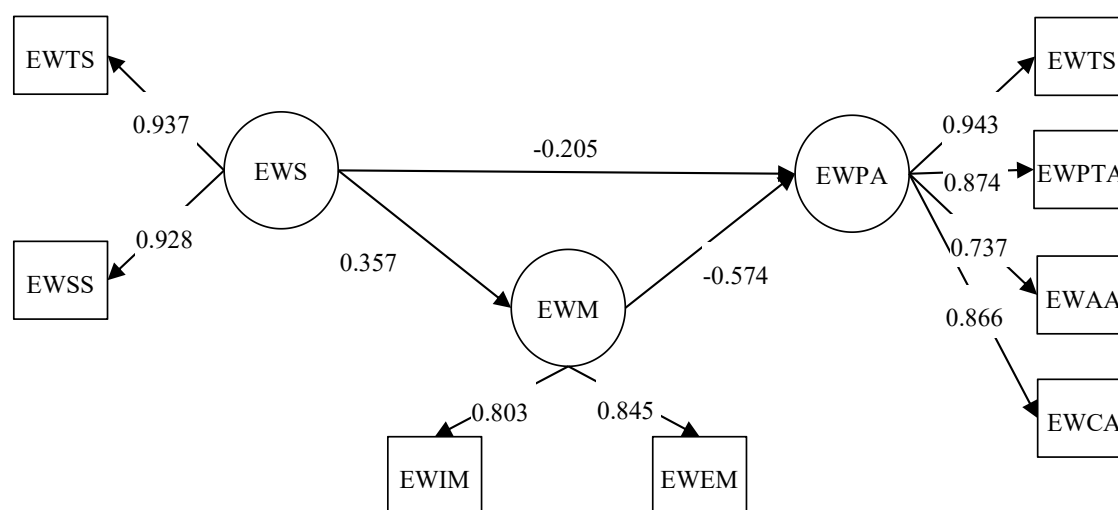


FIGURE 1
Results of the mediating effect of writing motivation.

psychological anxiety was still significantly negative ($\beta = -0.205$, $p < 0.001$), but the standardized coefficient was smaller. Thus, college students' writing motivation plays a partial mediating role in the relationship between writing self-efficacy and psychological anxiety. Writing self-efficacy can directly reduce college students' psychological anxiety, and can also alleviate psychological anxiety by enhancing their writing motivation.

Conclusion and implication

Research conclusion

The present study found that college students' overall English writing self-efficacy was at a moderate level, and Chinese college students lacked self-confidence in their English writing skills and tended to complete writing tasks. They generally had a moderate level of English writing psychological anxiety, with avoidance behavior being the most significant. College students' writing self-efficacy was significantly negatively related to psychological anxiety, and writing self-efficacy could reduce psychological anxiety by enhancing their writing motivation. The results reaffirm the previous literature on the negative effect of self-efficacy on psychological anxiety in the context of Chinese college students' English writing.

Practical implications

Our study provides pedagogical recommendations based on the research findings. First, enhance college students' English writing self-efficacy and cultivate their interest in English writing. According to the findings in our study, Chinese

college students' English writing self-efficacy level is not high, i.e., they are not confident enough in English writing. Thus, we suggest that college teachers need to pay attention to enhancing college students' self-efficacy in English writing in the teaching process (Bandura, 1977). For example, they should consolidate the basic knowledge of the English language and improve the language foundation of college students. When reviewing college students' English writing, the teacher should not only focus on word or grammatical errors, but also pay attention to the outstanding points of college students' English writing and give them positive and encouraging comments. In these ways, college students' self-confidence in English writing can be improved. In terms of cultivating college students' interest in English writing, the reason for this phenomenon that college students have a high level of English writing avoidance anxiety is the lack of interest, which means that college teachers should consciously cultivate college students' interest. In the topic selection, they should choose those writing topics that are close to the daily life of college students to reduce their avoidance psychology. In the review of writing, they should correct the English writing submitted by college students in time and give them timely feedback to enhance their learning motivation.

Second, pay attention to college students' English writing psychological anxiety. According to the high level of psychological anxiety among Chinese college students found in our study and the negative relationship between English writing self-efficacy and psychological anxiety, college teachers should pay attention to college students' attitudes toward English writing in their usual English writing teaching, actively communicate with students, and understand their emotional expressions. For example, when observing the anxiety of college students, teachers should take the initiative

to communicate with them and tell them about the normality of anxiety to weaken the psychological burden of college students. Besides, one important reason for the emergence of English writing psychological anxiety is the difficulties encountered in English writing. Pre-writing inspiration and conception or pre-writing preparation activities for college students can help to reduce their English writing psychological anxiety. In order to enrich college students' language knowledge and improve their language output skills, colleges teachers should increase comprehensible input as much as possible in the training of comprehensive language skills. At the same time, with a certain language's skills as a guarantee, college teachers should try to provide college students with various themes, and interesting content in writing materials. They also should assign different genres of writing tasks and try different forms of writing, such as poem, prose, and novel. College students should ensure enough writing output and improve the quality of written English. The improvement of writing ability is the basic way to eliminate English writing psychological anxiety.

Finally, strengthen the English writing motivation training of college students. Enhancing English writing motivation can also alleviate psychological anxiety for college students. Teachers in universities can apply the process writing teaching method in teaching, and carry out a series of idea training, such as brainstorming, to increase students' writing ideas before they write in English, so as to relieve the psychological anxiety caused by the lack of ideas (Rusinovci, 2015). In order to change college teachers' traditional evaluation method, we think teachers in colleges should reduce error-correcting composition and pay more attention to college students' writing process. This means that teachers in colleges need to shift students' attention from writing scores to the writing process. In the process of writing, what college students get and what wonderful sentences they accumulate in their brain may be more important, not just the numbers given by teachers to represent their writing ability by scores. In fact, the teacher can adopt self-evaluation, peer evaluation, and group evaluation to evaluate students' English writing.

Limitations and suggestions

Although the present study achieves the research aims and obtains some valuable findings, our study still has some limitations. As Pajares and Valiante (2012) mentioned in their study, exploring causal relationships between variables in non-experimental contexts requires sound theoretical assumptions about the model and considerable caution in interpreting the results. They also suggested that different theoretical models could be tested with more powerful statistical tools, such as SEM. Although we used SEM to suggest the positive effect of English writing self-efficacy on relieving psychological anxiety and the

mediating effect of writing motivation in the relationship from the theoretical and existing findings on English writing self-efficacy, English writing psychological anxiety, and writing motivation, and these findings were tested among Chinese college students, the results still could not clarify in the cross-sectional study. In the follow-up study, we can use the tracking research paradigm to explore the deeper causal relationships. In addition, our study is a sample from a single university, with no mention of the University courses involved, so its results are limited in terms of generalization.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

BL made significant contributions to the study concept and design. He was primarily responsible for designing the study, collecting and analyzing data, and drafting the manuscript. Also, he made several revisions and refinements to the content of the manuscript.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abolhasani, H., Golparvar, S. E., and Robatjazi, M. A. (2022). Modelling the role of L2 writing anxiety in graph-based composing performance and strategy use. *J. Psy. Res.* 51, 417–435. doi: 10.1007/s10936-022-09858-8
- Bandura, A. (1977). *Self-efficacy: the Exercise of Control*. New York, NY: Freeman. doi: 10.1891/0889-8391.13.2.158
- Bandura, A. (1986). *Social Foundations of Thought and Action: a Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall, doi: 10.2307/258004
- Cheng, Y. (2004). A measure of second language writing anxiety: scale development and preliminary validation. *J. Second Lang. Writ.* 4, 313–335. doi: 10.1515/caslar-2014-0009
- Daly, J. A., and Miller, M. D. (1975). Further studies on writing apprehension: SAT scores, success expectations, willingness to take advanced courses and sex differences. *Res. Teach. Engl.* 9, 250–256.
- Daly, J. A., and Wilson, D. A. (1983). Writing apprehension, self-esteem, and personality. *Res. Teach. Engl.* 17, 327–341.
- Fritzsche, B. A., Rapp Young, B., and Hickson, K. C. (2003). Individual differences in academic procrastination tendency and writing success. *Pers. Individ. Differ.* 35, 1549–1557. doi: 10.1016/S0191-8869(02)00369-0
- Gao, Y. H., Zhao, Y., Cheng, Y., and Zhou, Y. (2003). Motivation types of Chinese college undergraduates. *Mod. For. Lang.* 26, 28–38+6–8. doi: 10.3969/j.issn.1003-6105.2003.01.005.
- Gardner, R. C., and Lambert, W. E. (1972). *Motivation and Attitudes in Second Language Learning*. Rowley, MA: Newbury House Publishers.
- Guo, Y., and Qin, X. Q. (2010). Teaching of writing an empirical study on foreign language writing anxiety of Chinese non-English majors. *Foreign Lang. World* 2, 54–62+82.
- Hayes, J. R. (1996). “A new framework for understanding cognition and affect in writing,” in *The Science of Writing*, eds C. M. Levy and S. Ransdell (Mahwah, NJ: Lawrence Erlbaum), 1–27.
- Horwitz, E. K. (2001). Language anxiety and achievement. *Annu. Rev. Appl. Linguist.* 21, 112–126. doi: 10.1017/S0267190501000071
- Jin, Y. Y., Ahn, S., and Lee, S. M. (2022). The mediating effect of bicultural self-efficacy on acculturation and career decision-making self-efficacy for international students in South Korea. *Front. Psychol.* 13:602117. doi: 10.3389/fpsyg.2022.602117
- Jones, E. (2008). Predicting performance in first-semester college basic writers: revisiting the role of self-beliefs. *Contemp. Educ. Psychol.* 33, 209–238. doi: 10.1016/j.cedpsych.2006.11.001
- Kormos, J., and Csizér, K. (2008). Age-related differences in the motivation of learning English as a foreign language: attitudes, selves, and motivated learning behavior. *Lang. Learn.* 58, 327–355. doi: 10.1111/j.1467-9922.2008.00443.x
- Li, H., Liu, R. D., and Liu, Y. (2013). The mediating effects of EFL writing self-efficacy on the relationship between EFL writing anxiety and writing performance for college students. *Psychol. Dev. Educ.* 29, 385–390. doi: 10.16187/j.cnki.issn1001-4918.2013.04.009
- Li, J., Qi, J., Wu, L., Shi, N., Li, X., Zhang, Y., et al. (2021). The continued use of social commerce platforms and psychological anxiety—the roles of influencers, informational incentives and FoMO. *Int. J. Environ. Res. Public Health* 18:12254. doi: 10.3390/ijerph182212254
- Lowie, W., Verspoor, M., and Bot, K. D. (2009). “A dynamic view of second language development across the lifespan,” in *Language Development over the Lifespan*, (Routledge: Milton Park), 131–151. doi: 10.1002/0471264385.wei0601
- McCarthy, P., Meier, S., and Rinderer, R. (1985). Self-efficacy and writing: a different view of self-evaluation. *Coll. Compos. Commun.* 36, 465–471. doi: 10.2307/357865
- Mcleod, S. (1987). Some thoughts about feelings: the affective domain and the writing process. *Coll. Compos. Commun.* 38, 426–435. doi: 10.2307/357635
- Mendoza, L., Lehtonen, T., Lindblom-Ylänne, S., and Hyytinen, H. (2022). Exploring first-year university students’ learning journals: conceptions of second language self-concept and self-efficacy for academic writing. *System* 106:102759. doi: 10.1016/j.system.2022.102759
- Mitchell, K. M., Harrigan, T., and McMillan, D. E. (2017). Writing self-efficacy in nursing students: the influence of a discipline-specific writing environment. *Nurs. Open* 4, 240–250. doi: 10.1002/nop.2.90
- Murdoch, Y. D., and Kang, A. (2019). Direct instruction and use of online English writing software on EMI class-takers’ self-efficacy. *Int. J. Contents* 15, 97–106. doi: 10.5392/IJoC.2019.14.4.097
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Read. Writ. Quarterly* 19, 139–158. doi: 10.1080/10573560308222
- Pajares, F., and Valiante, G. (2012). Influence of self-efficacy on elementary students’ writing. *J. Educ. Res.* 90, 353–360. doi: 10.1080/00220671.1997.10544593
- Qadir, S. M., Bostanci, H. B., and Kurt, M. (2021). Writing apprehension among English as a foreign language postgraduate students. *Sage Open* 11:215824402110071. doi: 10.1177/21582440211007121
- Rusinovci, X. (2015). Teaching writing through process-genre based approach. *US-China Educ. Rev.* 5, 699–705. doi: 10.17265/2161-623X/2015.10.A.006
- Salam, A. A., and Al Dyar, M. A. (2014). Writing anxiety as a predictor of writing self-efficacy in English for special education Arab learners. *Int. Educ. Stud.* 7, 128–134. doi: 10.5539/ies.v7n6p128
- Sun, T., Wang, C., and Kim, S. Y. (2021). Psychometric properties of an English writing self-efficacy scale: aspects of construct validity. *Read. Writ.* 35, 743–766. doi: 10.1007/s11145-021-10206-w
- Torres, K. M., Arrastia-Chisholm, M. C., and Tackett, S. (2020). Perceptions of writing anxiety and self-efficacy among Spanish heritage language learners. *J. Hispanic Higher Educ.* 19, 84–98. doi: 10.1177/1538192718775175
- Wang, X., and Zhang, W. (2021). Psychological anxiety of college students’ foreign language learning in online course. *Front. Psychol.* 12:598992. doi: 10.3389/fpsyg.2021.598992
- Woodrow, L. (2011). College English writing affect: self-efficacy and anxiety. *System* 39, 510–522. doi: 10.1016/j.system.2011.10.017
- Wright, K. L., Hodges, T. S., and McTigue, E. M. (2019). A validation program for the self-beliefs, writing-beliefs, and attitude survey: a measure of adolescents’ motivation toward writing. *Assessing. Writ.* 39, 64–78. doi: 10.1016/j.asw.2018.12.004
- You, C., and Dörnyei, Z. (2016). Language learning motivation in China: results of a large-scale stratified survey. *Appl. Linguis.* 37, 495–519. doi: 10.1093/applin/amu046
- Zeng, X., Wang, S., and Zhang, F. (2020). Correlation analysis between self-efficacy and psychological anxiety of college English learners based on triadic reciprocal determinism. *Revista Argentina de Clin. Psicol.* 29, 1319–1324. doi: 10.24205/03276716.2020.190
- Zhang, H., Chen, M., and Li, X. (2021). Developmental features of lexical richness in English writings by Chinese beginner learners. *Front. Psychol.* 12:665988. doi: 10.3389/fpsyg.2021.665988
- Zhang, Y., and Guo, H. (2012). A study of English writing and domain-specific motivation and self-efficacy of Chinese EFL learners. *J. Pan-Pacific Assoc. Appl. Linguis.* 16, 101–121.
- Zhu, A. (2021). Language acquisition and anxiety: writing anxiety impedes better learning. *Int. J. Social. Sci. Educ. Res.* 4, 427–431.



OPEN ACCESS

EDITED BY
Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY
Carol A. Miller,
The Pennsylvania State University,
United States
Lillyam Gonzalez,
University of Ostrava, Czechia

*CORRESPONDENCE
Giulia Vettori
giulia.vettori@unifi.it

SPECIALTY SECTION
This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 17 May 2022
ACCEPTED 28 July 2022
PUBLISHED 23 August 2022

CITATION
Vettori G, Bigozzi L, Incognito O and
Pinto G (2022) Contribution of oral
narrative textual competence
and spelling skills to written narrative
textual competence in bilingual
language-minority children
and monolingual peers.
Front. Psychol. 13:946142.
doi: 10.3389/fpsyg.2022.946142

COPYRIGHT
© 2022 Vettori, Bigozzi, Incognito and
Pinto. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Contribution of oral narrative textual competence and spelling skills to written narrative textual competence in bilingual language-minority children and monolingual peers

Giulia Vettori*, Lucia Bigozzi, Oriana Incognito and
Giuliana Pinto

Department of Education, Languages, Intercultures, Literatures and Psychology (Psychology
Section), University of Florence, Florence, Italy

This study investigates the developmental pattern and relationships between oral narrative textual skills, spelling, and written narrative textual skills in monolingual and bilingual language-minority (BLM) children, L1-Chinese and L2-Italian. The aims were to investigate in monolingual and BLM children: (1) the developmental patterns of oral and writing skills across primary school years; (2) the pattern of relationships (direct and mediated) between oral narrative textual competence, spelling skills, and written narrative textual competence with age and socio-economic status (SES) taken under control. In total, 141 primary school children from grades 2 to 5 in Central Italy (44% BLM, 56% monolinguals) aged between 7 and 11 years ($M\text{-age} = 8.59$, $SD = 1.13$; 41% girls, 59% boys) obtained scores for oral and written narrative textual competence, spelling accuracy in dictation, and written texts. One-way ANOVA and ANOVA with robust method (Welch test) analyses and Bonferroni's correction showed that BLM children had poorer spelling skills in dictation and written narrative textual competence (i.e., text structure) than their monolingual peers. After preliminary correlation analysis, the results of hierarchical regression showed that the relationship between oral and written narrative textual competence is completely mediated by spelling accuracy in BLM children. These results suggest that adequate performance in written narrative textual competence depends on adequate spelling accuracy in writing stories. The Sobel test verified the power of this mediation. In monolinguals, the strongest predictor of written narrative textual competence is oral narrative textual competence. This relation is stronger in older children whose spelling skills are automatized. The identified pattern of relationships shows a complex network of oral and written processes. The scarce spelling skills characterizing BLM children may explain why spelling skills determine a low written narrative textual level. Scarce spelling skills absorb cognitive resources, hindering high-level cognitive processes that

regulate narrative production. In monolinguals, the medium of writing does not impact narrative textual competence. Children's oral narrative textual competence easily transfers into their written narrative productions. These findings have implications for the assessment and instruction of literacy skills in young BLM children and their monolingual peers.

KEYWORDS

bilingual language-minority children, textual competence, oral narrative, written narrative, spelling skills, Italian, Chinese

Introduction

An exponential growing number of students in school classrooms worldwide are bilingual language-minority (BLM) children. These children experience the most intensive and sustained exposure to their second language (L2) when they begin formal education because they come from homes where the language background is different from the societal language (e.g., Hoff et al., 2021). Previous studies indicate that BLM children have fewer emergent orthographic knowledge in kindergarten (Incognito et al., 2021) and writing skills in school years (Goodrich et al., 2016) with detrimental effects on achievement and school outcomes. A serious disadvantage also for online communication in written form, a widespread practice (Alarcón et al., 2020). Theory and evidence indicate that writing narrative production skills are closely linked to the development of oral and spelling skills. Therefore, this study aimed to examine the developmental pattern and (direct and mediated) relations of L2 oral and writing skills among school-age children, including BLM children and their monolingual peers. Previous research shows the development of spelling and written narrative textual competence skills in monolingual, whereas oral narrative textual competence skills are less explored, especially across primary schools. Little is known about the developmental pattern and relations between oral and writing skills in BLM children acquiring two structurally and morphologically distinct languages. To our knowledge, this is one of the few attempts to study different skills across the primary school years in the profile of BLM children acquiring L1-Chinese and L2-Italian.

Impact of oral language and spelling skills on writing skills

In this study, we refer to a model of writing development that comprehends important distal and proximal factors. Models of writing development, based on English, such as

the “simple view of writing,” describe how proximal lower-level transcription skills (e.g., spelling skills), distal higher-level processes (e.g., planning, translating, and reviewing), and general cognitive skills (e.g., working memory, reading, attention) contribute to producing written texts. The updated model of writing, the so-called “not-so-simple view,” includes executive functions (e.g., attention, planning, reviewing) as core, describing how these components evolve and interact with transcription skills (handwriting or typing and spelling) in text generation. Consistent with previous models, mainly derived from opaque orthographies, beginning writers are mainly engaged in managing “lower-level processes,” such as proximal spelling skills (Hayes, 2012). Researchers agree that this step is particularly challenging for novice writers who are engaged in translating their ideas into text (Graham and Harris, 2000) by observing the grammatical and spelling rules of the language system (Graham, 2006). Recently, oral language and reading skills have been incorporated into these models as distal factors that contribute to the production of written text (e.g., Kim and Park, 2019). The recent “Direct and Indirect Effects model of Writing” (DIEW; Kim and Schatschneider, 2017; Kim, 2020) further demonstrates that both oral text generation and spelling skills are necessary to support writing quality. In novice writers, difficulties in spelling skills may hinder writing narrative textual processes and interfere with the transfer of oral narrative textual competence to written narrative textual productions. It is important to consider that the impact of spelling skills on writing texts may vary according to the characteristics of orthographies. Most evidence derived from studies typically focuses on English, where children's spelling skills are challenged by an opaque language system with a strong inconsistency between its phoneme and grapheme correspondences. This research emphasizes the importance of investigating writing development in relation to children's linguistic and socio-economic backgrounds (e.g., Incognito et al., 2021). We still know only very little about the role of spelling skills in writing texts in school-age BLM children acquiring Italian as L2. Italian is defined as a transparent or “shallow” language because it shows consistent sound-sign matching, similar to Finnish, German, and Spanish. Research indicates that for

Italian monolingual children, the consistency between sound and sign matching sustains them in advancing quickly to writing sentences, and by the end of the first grade, children are able to write short texts (Bigozzi et al., 2016a). Further studies in Italian monolingual children (Bigozzi and Vettori, 2016) indicate that only in the case of a significant disadvantage in spelling skills do first graders' oral narrative textual competence cease to sustain written narrative textual competence. Little is known about the pattern of (direct and mediated) relationships between oral narrative textual competence, spelling skills, and written narrative textual competence in children acquiring transparent Italian orthography through primary school years.

Writing a story in L2

In studying children's written narrative textual competence, it is interesting to distinguish which factors are sustained in the early or late phase of learning to write in primary school, because the characteristics of orthography may exert a different role on novice or more expert writers and monolingual or BLM children. Previous longitudinal studies on the transition from preschool to formalized education (e.g., Pinto et al., 2015, 2016) indicate that Italian preschool children's oral narrative textual competence influences their later written narrative competence in first and second grades via a mediational effect of orthographic competence. Instead, there has been little research on how L2 spelling skills contribute to L2 written narrative textual competence in BLM children acquiring two orthographically distant languages: L1-Chinese (non-alphabet) and L2-Italian (alphabet and transparent orthography). The limited L2-Italian language input in the family context can negatively impact the acquisition of adequate L2 spelling skills (Roch and Hrzica, 2020). Prior research indicates that BLM children use word knowledge in their L1 when writing in their second language (L2) (e.g., code-switching between two languages) (e.g., Gort, 2001). This adaptive strategy sometimes conducts to a "negative transfer" across languages (e.g., Howard et al., 2012), leading children to commit L1-influenced spelling errors when writing in the societal language (e.g., "i" instead of "r"). The Italian language as a transparent orthography might reduce cognitive demands on young writers in spelling, but less is known about the relationship between spelling and narrative texts when writers must manage the demand of the morphosyntactic complexity of this language.

Aim and hypothesis

In this study, we examine the contribution of oral language and spelling skills to written narrative textual competence in a specific population of school-aged BLM children acquiring distant structural and morphological languages, L1-Chinese

(non-alphabet) and L2-Italian (alphabet and transparent orthography), and their monolingual peers. Previous research has mainly focused on the transition from preschool to formalized education (e.g., Pinto et al., 2009, 2012; Bigozzi and Vettori, 2016). In this study, we focus on the important period of primary school years when children from grades 2 to 5 are expected to have almost automatized orthographic competence in Italian. Previous results have confirmed the usefulness of adopting both dictation and spontaneous spelling of words to assess children's orthographic skills (Bigozzi et al., 2016a). In this study, two measures of spelling skills were adopted to gain a more complete picture of the involvement of spelling skills in the process of writing text by using both a measure *in the process* of writing text and an *external measure* of the level of proficiency in a dictation task.

This study contributes to closing the aforementioned gaps with the following aims.

First, we aimed to investigate the developmental patterns of BLM and monolingual children's oral narrative textual competence, spelling skills, and written narrative textual competence across primary school years. We expect an improvement in children's capacity to tell and write stories and in their spelling accuracy across primary school years. Specifically, we are interested in providing a better understanding of this improvement in the outcome measures of children in primary school by verifying floor or ceiling effects: whether there is rapid development or only gradual improvements during that period and whether there are any developmental plateaus.

Second, we aimed to examine the patterns of direct and mediated relationships between oral narrative textual competence, spelling skills, and written narrative textual competence in school-aged BLM children and their monolingual peers. We expect oral narrative textual competence and spelling skills to play a significant predictive role in written narrative textual competence. Specifically, we expect that the contribution of spelling skills is influenced by the level of proficiency, which is expected to be lower in novice writers and BLM children in comparison with monolinguals. Regarding monolingual children, we expect spelling skills to have a less significant influence on their written narrative textual competence, increasing with primary school grades, considering that most children show an adequate mastering of spelling skills in the Italian transparent orthography since the early years.

Materials and methods

Participants

A total of 141 primary school children aged 7–11 years living in Italy (M-age = 8.59, SD = 1.13; 83 girls and 58 boys)

participated in this study as follows: 79 children (56%) were monolinguals, Italian-as-L1 children, exposed to the societal language both at home and school, and 62 children (44%) were Chinese-speaking language-minority bilingual children exposed to an L1 other than Italian in the family context. All Chinese children are born in Italy to parents who are both Chinese. In this study, the great proportion of Chinese as children's minority language is related to the fact that they belong to a wide area in central Italy with a long history of Chinese immigrants. Given that, in Italy, children usually go to the public school closest to where they live, and classrooms saw a significant presence of Chinese bilingual language minorities. Children attending different primary school classrooms were included in this study: second grade $N = 30$ (monolingual = 14; BLM = 16); third grade $N = 35$ (monolingual = 20; BLM = 15); fourth grade $N = 43$ (monolingual = 25; BLM = 18); and fifth grade $N = 33$ (monolingual = 20; BLM = 13). Children with any known special educational needs or impairments/disorders were excluded to avoid any additional difficulties that could potentially affect their performance. School authorities, parents, and children provided consent to participate in the study. Background information about home language characteristics and sociocultural-economic status, defined as parents' educational level (i.e., International Standard Classification of Education; ISCED-11; [UNESCO Institute for Statistics, 2012](#)), were collected using a parental questionnaire attached to the informed consent sheet. In Italy, first-grade teachers focus primarily on the spelling component of writing, whereas second-grade teachers focus on the textual properties of writing because second graders are expected to have finalized the acquisition of orthography (National Indications of the Italian Ministry of Education, hereafter [MIUR, 2012](#)). The children's ages range from 7 to 10 years, corresponding to the second to fifth grades in school.

Measures

Narrative tasks: Oral and written narrative textual competence

Two narrative tasks were adopted to measure children's oral and written narrative textual competence. In a collective session in the classroom during school, the children individually performed a written narrative task. Each child wrote an invented story. Furthermore, in an individual session in a quiet room next to the classroom, children individually performed an oral narrative task. Each child was asked to tell a story they invented. Each story told was transcribed and codified. With respect to the aims of this study, each written and oral narrative was codified for the index of text structure. To identify the level of text structure in children's written and oral stories, in line with previous research (e.g., [Spinillo and Pinto, 1994](#); [Pinto et al., 2020](#)), the presence/absence of the 8 elements which characterize

the narrative genre (e.g., title, opening, setting, description of character/s, problem, central event, resolution of the problem, and story ending) were identified to assign the corresponding score on a five-point scale as follows:

Score 1 – no narrative: Simple description or list of events, objects, or facts;

Score 2 – sketch narrative: Opening, setting, character(s), conclusion or opening, the sketch of the problem, and resolution;

Score 3 – incomplete narrative: Opening, character(s), problem, and resolution;

Score 4 – essential narrative: Opening, character(s), problem, central event, and resolution;

Score 5 – complete narrative: Title, opening, character(s), setting, problem, central event, resolution, and narrative closing.

The rate of agreement between the judges was 99%; cases of disagreement were resolved through discussion. In summary, each child in this study obtained a score for

- “oral narrative textual competence,” derived from the score of structure in oral texts;
- “written narrative textual competence,” derived from the score of structure in written texts.

A dictation task and a written narrative task: Spelling skills

To assess children's spelling skills, a paper-and-pencil text dictation was performed individually by children in a collective session in the classroom during school time. The dictation task was taken (BVSCO) from the “Battery for the Evaluation of Writing and Orthographic Competence in Primary School” standardized for the Italian population. The children listened to a recorded text, and each child had to write down the text. To measure children's spelling skills in the two tasks (dictation and written narrative), the orthographic errors were identified based on the classification of the orthographic errors by [Pinto et al. \(2012\)](#) which covers the entire variability of orthographic errors that children may commit in the Italian language. As indicated in previous research (see e.g., [Bigozzi et al., 2017](#)), it is important to gain a comprehensive score of orthographic error including both the cases in which the pronunciation of the target word is preserved despite the spelling violation (e.g., “anno” [year] instead of “hanno” [they have]), and both the cases in which the pronunciation of the target word is changed due to a spelling violation (“mecrato” instead of “mercato”). All phonemes are possible sources of non-homophone errors such that spelling errors of this type result in phonetically implausible words. Writing accuracy was determined by the total number of

orthographic errors (see, [Pinto et al., 2012](#)), which were counted as many times as the error occurred.

Sociocultural-economic status

A parental questionnaire attached to the informed consent form was used to collect information about the children's family sociocultural-economic status (SES). For this study, the index of ISCED level was considered as follows: a score from "ISCED 1 – Primary education" till "ISCED 6 – bachelor's or master's degree" was given for the educational level of fathers and mothers. Children's SES scores were calculated based on the higher ISCED level between parents. The measure consists of the number of years of father and mother's education (see also, [Haman et al., 2017](#)).

Data analysis

First of all, we grouped the participants as follows: the youngest were the students attending second and third grades (total $N = 65$; monolingual = 34, BLM = 31); the oldest were the students attending fourth and fifth grades (total $N = 76$; monolingual = 45, BLM = 31). This division stems from the fact that second graders and third graders have similar spelling skills, in fact, it is at the end of the third grade that spelling proficiency is considered acquired. The first graders are not part of the sample because they are too immature in terms of learning. By third grade, they have acquired spelling skills, and in fourth and fifth-grades, children master spelling.

Before analyzing the data, according to [Tabachnick and Fidell's \(2013\)](#) recommendation, the presence of univariate outliers in the oral and written narrative skills scores was checked. No outliers are observed. Preliminarily, the homogeneity of variances was checked by Levene's test. To investigate the oral narrative, spelling, and written narrative text production skills in L2-Italian shown by BLM (L1-Chinese and L2-Italian) and their monolingual Italian-speaking peers, one-way ANOVA was used. In the case of non-homogeneity of variances, ANOVA with robust methods was used (Welch's Test). Moreover, we used Bonferroni's correction to control the family-wise error rate ($\alpha_{\text{new}} = \alpha_{\text{old}}/n$). Correlation analyses were then performed to investigate the relationships between variables as preliminary analyses for regression assumptions. Finally, we used hierarchical regression analysis to verify the increment in variation accounted for by the addition of predictors over a set of models ([Figure 1](#)). This is generally assessed by testing the change in R-square from one model to the next. If, after the inclusion of predictors at a given step, the R-square change is significantly greater than zero, we infer that the predictors added at that step offer incremental predictive power. The R-square change (increment) from Model 1 to Model 2 is computed as $\Delta R^2 = \text{Model 2 } R^2 - \text{Model 1 } R^2$. The R-square change (increment) from Model 2 to Model 3 is computed as $\Delta R^2 = \text{Model 3 } R^2 - \text{Model 2 } R^2$ ([Darlington and Hayes, 2017](#)).

To calculate the probability, if any, we used the Sobel test when the indirect effect of an independent variable on a dependent variable through a mediating variable is significant. Many studies rely on mediating models, and identifying whether a mediating variable significantly mediates the influence of an independent variable on a dependent variable is critical when assessing the value of such models.

Results

Before analyzing the data, based on the parents' educational level, the sample was distributed as follows: 1.4% primary school, 35.5% middle school, 25.5% 3-year professional qualification, 32.6% high school, 1.4% another higher education qualification other than a high-school diploma (conservatory, arts), and 3.5% master's degree. Preliminary descriptive statistics and comparative analyses between monolinguals and BLM (one-way ANOVA and ANOVA with the robust method – Welch's test) for the main variables are shown in [Table 1](#). Using Bonferroni's correction, the new alpha level is 0.01. The results show that statistically significant differences were found between monolingual and BLM children in spelling inaccuracy in dictation and written narrative skills. Specifically, BLM children exhibited a significantly higher number of errors than monolingual children. Regarding written narrative skills, BLM children performed statistically worse than their monolingual peers.

[Table 2](#) shows descriptive statistics and comparison analyses between younger and older children in monolinguals and BLM (one-way ANOVA and ANOVA with the robust method – Welch's test). The results show that both in monolingual and BLM, performance in written narrative skills and orthographic accuracy improve with age. However, there is no significant increase in oral narrative skills. Using the Bonferroni Correction, the new alpha level is 0.01.

Correlation analyses were conducted to determine the relationships between the variables. Scores on written narrative skills were associated with the age of participants, oral narrative skills, and spelling accuracy in narration and dictation. [Table 3](#) presents the results.

Hierarchical regression analysis was performed to determine whether performance in oral narrative skills and spelling accuracy in narrative and dictation improved participants' prediction of their written narrative skills beyond that provided by age and parents' level of education. These predictors were used in the equation because of their statistically significant correlations with written narrative skills. [Table 4](#) shows the standardized regression coefficients (β), R^2 , and change R^2 (ΔR^2) for monolingual children.

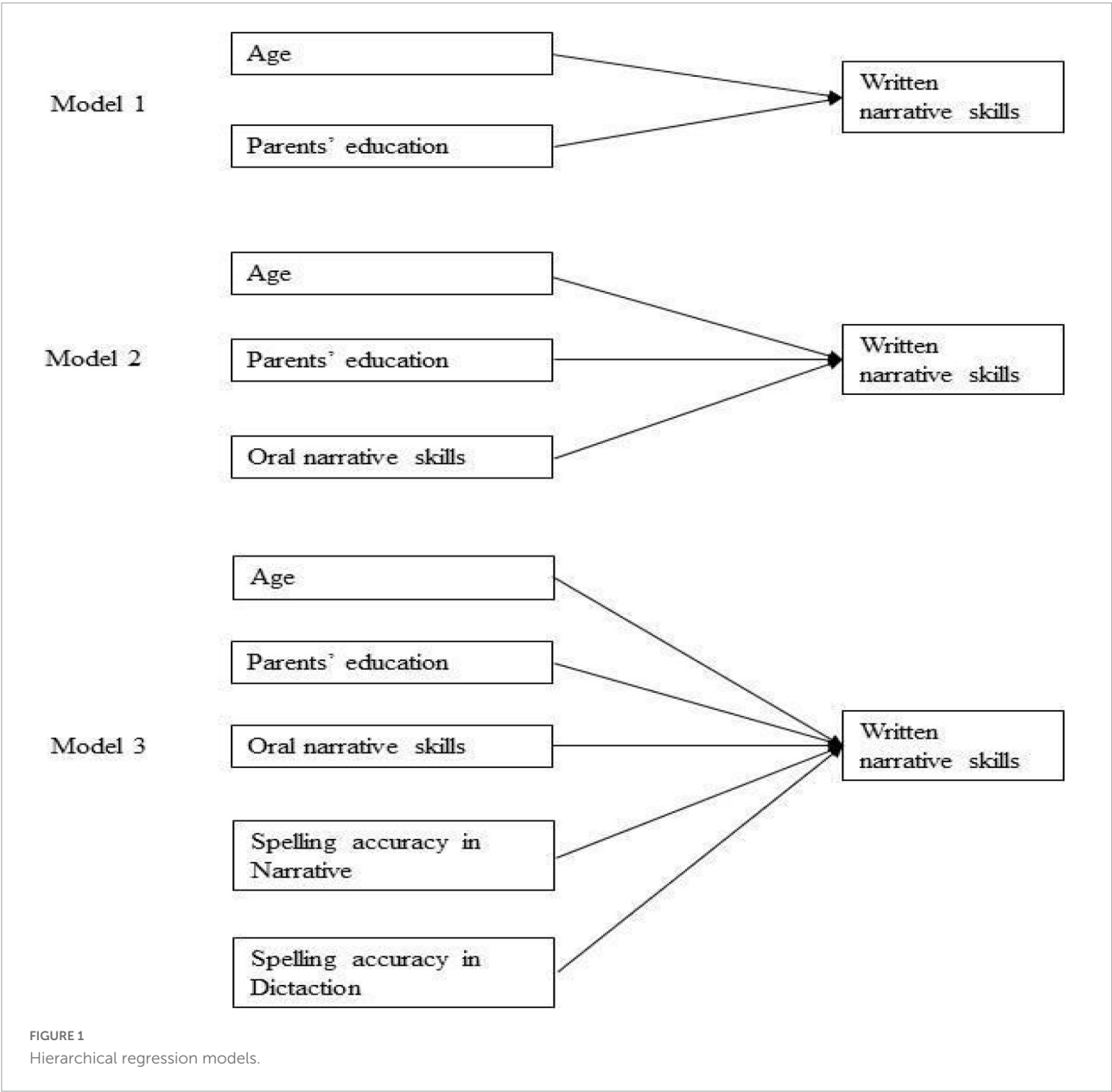


TABLE 1 Descriptive analyses and one-way ANOVA and ANOVA with robust methods results.

		Mean (SD)	Minimum	Maximum	Levene's statistic ¹	F (Welch test)	DF
Spelling accuracy in Dictation	Monolingual	0.13 (0.15)	0	0.68	6.70*	11.51***	1,123
	BLM	0.22 (0.17)	0	0.65			
Spelling accuracy in Narrative	Monolingual	0.07 (0.06)	0	0.31	25.31***	5.60	1,123
	BLM	0.11 (0.14)	0	0.50			
Written narrative skills	Monolingual	3.10 (1.16)	1	5	0.42	21.61***	1,123
	BLM	2.14 (1.16)	1	4			
Oral narrative skills	Monolingual	3.17 (1.04)	1	5	0.98	4.54	1,98
	BLM	2.69 (1.10)	1	4			

* $p < 0.05$; *** $p < 0.001$; ¹ in the case of significance of Levene's test: F statistics were computed with Welch test; Bonferroni's correction was applied, and the p -value is significant when $p < 0.01$.

TABLE 2 Descriptive analyses and one-way ANOVA and ANOVA with robust methods results for monolingual and BLM children at school level.

		Mean (SD)	Minimum	Maximum	Levene's statistic ¹	F (Welch test)	DF
Monolingual children							
Spelling accuracy in Dictation	Younger Children	0.22 (0.17)	0.03	0.68	13.99***	24.37***	1, 73
	Older Children	0.06 (0.08)	0	0.39			
Spelling accuracy in Narrative	Younger Children	0.10 (0.06)	0	0.23	1.66	10.43**	1, 65
	Older Children	0.05 (0.06)	0	0.31			
Written narrative skills	Younger Children	2.42 (1.06)	1	4	0.03	16.05***	1, 65
	Older Children	3.49 (1.03)	2	5			
Oral narrative skills	Younger Children	3.12 (1.17)	1	5	2.66	0.08	1, 64
	Older Children	3.20 (0.97)	2	5			
BLM children							
Spelling accuracy in Dictation	Younger Children	0.33 (0.15)	0.06	0.65	3.85*	34.68***	1, 55
	Older Children	0.12 (0.11)	0	0.46			
Spelling accuracy in Narrative	Younger Children	0.18 (0.16)	0	0.50	18.12***	16.46***	1, 56
	Older Children	0.05 (0.07)	0	0.32			
Written narrative skills	Younger Children	1.38 (0.77)	1	4	1.68	42.95***	1, 56
	Older Children	2.90 (0.98)	2	5			
Oral narrative skills	Younger Children	2.42 (1.24)	1	5	1.21	0.96	1, 33
	Older Children	2.83 (1.03)	2	5			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ¹ in the case of significance of Levene's test: F statistics were computed with Welch test; Bonferroni's correction was applied, and the p -value is significant when $p < 0.01$.

TABLE 3 Correlation analyses.

		Age	Parents' education	Written narrative skills	Oral narrative skills	Spelling accuracy in Narrative	Spelling accuracy in Dictation
Monolingual	Age	–	–0.11	0.45**	0.04	–0.38**	–0.54**
	Parents' education		–	0.01	0.04	0.14	0.07
	Written narrative skills			–	0.35**	–0.35**	–0.48**
	Oral narrative skills				–	–0.34**	–0.18
	Spelling accuracy in Narrative					–	0.76**
	Spelling accuracy in Dictation						–
BLM	Age	–	–0.21	0.66**	0.18	–0.48**	–0.62**
	Parents' education		–	–0.25	–0.14	0.16	0.07
	Written narrative skills			–	0.35*	–0.51**	–0.63**
	Oral narrative skills				–	–0.37*	–0.23
	Spelling accuracy in Narrative					–	0.52**
	Spelling accuracy in Dictation						–

** $p < 0.01$; * $p < 0.05$.

In Model 1, age and parents' level of education accounted for significant variations in written narrative skills (R-square = 0.19, $F(2,52) = 6.20$, $p = 0.004$). In Model 2, age, parents' level of education, and oral narrative skills accounted for significant variations in written narrative skills (R-square = 0.32, $F(3,51) = 8.18$, $p < 0.001$). The change in R-square from Model 1 to Model 2 was 0.13, reflecting a significant increase in the explained variation [$F(1,51) = 9.99$, $p < 0.01$]. In Model 3, the predictors accounted for significant variation in written narrative skills (R-square = 0.36, $F(5,49) = 5.57$, $p < 0.001$). The change in R-square from Model 2 to Model 3 was 0.04, which

does not reflect a significant increase in explained variation [$F(2,49) = 1.44$, $p = 0.247$]. These results show that the strongest predictor of written narrative skills in monolingual children is oral narrative skills. Moreover, the age of participants as a control variable is presumably contributing to independent variance, as suggested by previous correlational analyses. This is explained by the fact that the beta value decreased and the p -value increased for age, in Model 3.

Table 5 contains the standardized regression coefficients (β), R^2 , and change R^2 (ΔR^2) for BLM children. In Model 1, age and parents' level of education accounted for significant variation

TABLE 4 Stepwise regression for monolingual children (dependent variable: written narrative skills).

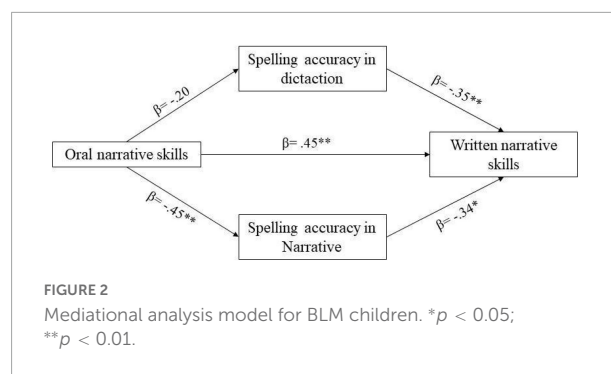
		β	R^2	ΔR^2
Step 1	Age	0.44**	0.19*	
	Parents' education	0.06		
Step 2	Age	0.41**	0.32**	0.13
	Parents' education	0.06		
	Oral narrative skills	0.36*		
Step 3	Age	0.30*	0.36**	0.04
	Parents' education	0.07		
	Oral narrative skills	0.34*		
	Spelling accuracy in Narrative	-0.28		
	Spelling accuracy in Dictation	0.09		

* $p < 0.05$; ** $p < 0.01$.**TABLE 5** Stepwise regression for BLM children (dependent variable: written narrative skills).

		β	R^2	ΔR^2
Step 1	Age	0.47*	0.30*	
	Parents' education	-0.17		
Step 2	Age	0.39*	0.39*	0.09
	Parents' education	-0.15		
	Oral narrative skills	0.31*		
Step 3	Age	-0.07	0.54**	0.15
	Parents' education	-0.21		
	Oral narrative skills	0.23		
	Spelling accuracy in Narrative	-0.37**		
	Spelling accuracy in Dictation	-0.31		

* $p < 0.05$; ** $p < 0.01$.

in written narrative skills (R -square = 0.30, $F(2,30) = 5.19$, $p = 0.004$). In Model 2, age, parent's level of education, and oral narrative skills accounted for significant variations in written narrative skills (R -square = 0.39, $F(3,29) = 6.23$, $p = 0.002$). The change in R -square from Model 1 to Model 2 was 0.09, reflecting a significant increase in explained variation [$F(1,29) = 4.24$, $p < 0.05$]. In Model 3, the predictors accounted for significant variations in written narrative skills, R -square = 0.54, $F(5,27) = 3.71$, $p < 0.001$. The change in R -square from Model 2 to Model 3 was 0.15, reflecting a significant increase in explained variation, $F(2,27) = 4.39$, $p < 0.05$. These results suggest that only spelling accuracy in written narrative texts is the best predictor of written narrative skills in BLM children; therefore, good performance in written narrative skills depends on good spelling accuracy in writing stories. Specifically, the model in step 3 improves significantly. In addition, in step 2, the oral narrative skills were significant, whereas, in step 3, they lost significance. This suggests that in the relationship between oral and written narrative skills, narrative orthographic accuracy completely mediates this relationship. We used the Sobel test to test the



power of this mediation. The results indicate that there is a total mediation of narrative orthographic accuracy performance in the relationship between oral and written narrative skills (Sobel test statistic = 2.23, $p < 0.05$; **Figure 2**). **Figure 2** explains the mediational model, and beta values and significances are shown for each distinct relationship.

Discussion

This study provides results on developmental patterns and relationships between oral narrative textual competence, spelling skills, and written narrative textual competence in two groups of school-age children, bilingual language-minority children, and their monolingual peers. The cross-sectional research design of this study provides results on the change and stability of the outcome measures throughout the primary school years.

As expected, oral narrative textual competence, spelling skills, and written narrative textual competence were found to develop throughout the primary school years in both language groups, BLM children and their monolingual peers, although some developmental trends were different within writing and oral domains and between language groups. From the comparison between young and older monolingual primary school children, a growth trend in oral narrative textual competence was observed. As presumably, older monolingual children produced oral stories with more structural information (i.e., story grammar units) than did the younger. This result shed light on a growth trend less documented in the Italian population and they are in line with previous studies on narrative development in diverse languages (Mäkinen et al., 2020). However, it is important to note that when considering BLM children's oral narrative textual competence, the results show a developmental trend from the younger to older BLM primary school children with only minor improvement from the score around 2 (sketch narrative) to the score around 3 (incomplete narrative). Although a trend of improvement in oral narrative competence was observed both in bilingual and monolingual, the results did not reach significance, and this is somewhat surprising since narrative structure generally

tends to improve significantly with age. The finding that the narrative structure in the BLM's oral narratives reported only an improving trend, which was not statistically significant, is consistent with the fact that children's oral narrative skills are poorly supported by teaching in primary school. Indeed, primary school focuses mainly on reading and writing skills, which, although related to the oral domain, are not yet able to influence the oral domain. In primary school years, children are scarcely engaged in telling oral narratives as for reading or listening to stories to the advantage of the teaching of writing (see e.g., Ministry of Education, University and Research [MIUR, 2012](#)).

Turning to spell skills, consistent growth in spelling accuracy occurred from younger to older primary school BLM and monolingual children. This is in line with previous studies on Italian children that document an age-related decline in errors, such as errors of omission, inversion, and non-phonological type together in dictation and written stories. This result can be confirmed for a specific population of BLM children acquiring the phoneme-grapheme correspondence, which is at the base of orthographic competence, in two languages with large differences. When taking a closer look at written narrative textual competence, performance was in the expected direction, with increasing primary school grades associated with improved performance in both language groups, BLM and their monolingual peers. Specifically, in monolingual children, the narrative structure in the younger group was around 2, indicating "sketch stories," instead in the older group was a little above 4 moving toward an "essential narrative" with opening, character(s), problem, central event, and resolution. These results show increasing complexity in the mental model of the story, in line with previous studies ([Pinto et al., 2015](#)). In the younger BLM children group, the narrative structure was around 1, indicating a simple "description or list" of events, objects, or facts, whereas, in the older BLM children group, it was around 3, indicating an "incomplete story" with a conventional opening, character(s), problem, and resolution.

The two language groups (i.e., BLM and monolingual) display the same changes in oral and written domains across primary school years and are united by the school experience which focuses on reading and writing instruction while leaving aside oral practice. Alongside the similarities in growth, the comparison of oral narrative textual skills, spelling, and written narrative textual skills between monolinguals and BLMs shows that BLMs show quite equal oral narrative skills to monolinguals setting around the mean-level of 2 "sketch story" and 3 "incomplete story." For what concerns the comparison of spelling skills, BLM children's level of acquisition of sound-sign correspondence was significantly lower than those of their monolingual peers in the dictation task. As reported in the literature, a disadvantage at the orthographic level in BLM children has been observed since preschool when BLM children have reported lower levels of notational awareness than their monolingual peers ([Incognito et al., 2021](#)). Notational

awareness refers to preschoolers' conceptual knowledge of the writing system assessed by the "invented spelling task" requiring the ability to process forms of writing similar to conventional spelling. The invented spelling task, administered by the researcher, consists of asking the child to write a few words in order to assess, through various indicators, to what extent to which the child is able to produce signs more or less similar to letters and to read following with her/his finger what the child has written, in an attempt to make each sign correspond more or less correctly to a sound. Notational awareness is also assessed through the ability of the child to vary the number of signs in relation to the variation of the number of sounds in the word ([Bigozzi et al., 2016a](#)). Notational awareness represents a significant predictor of later reading and writing skills of primary school children in the Italian language system ([Bigozzi et al., 2016b](#)). In addition to the evidence that clarifies that children's L2 exposure is strongly associated with their L2 language skills (e.g., [Hoff, 2018](#)), our results that BLM children were inferior in spelling skills in comparison to monolingual peers allow us to consider the influence of the characteristics of orthography, given that they were acquiring two distant structural and morphological languages, L1-Chinese (non-alphabet) and L2-Italian (alphabet and transparent orthography). While BLM children's oral narratives are preserved, the obstacle of orthographic coding absorbs cognitive resources to the construction of a rich and detailed mental model of a story, aggravating the functioning of working memory for planning the text ([Berninger and Swanson, 1994](#); [Hayes, 2012](#)). As previous writing models showed, children's automatization of spelling skills enables them to free up memory and executive function resources to be devoted to higher-level processes of writing (e.g., generation of ideas, planning, and revision). This laborious search for words and correct transcription persists in older BLMs making the arduous circulation of knowledge between oral and written. These results of the comparisons between BLM vs. monolinguals corroborate the understanding of how important it is to master spelling so that it does not represent an obstacle to more creative and ideational processes such as textual writing, as already demonstrated in previous studies in other age groups (e.g., [Pinto et al., 2015](#); [Bigozzi and Vettori, 2016](#)).

To successfully capture the pattern of the direct and mediated contributions of oral narrative textual competence and spelling skills to written narrative textual competence, we consider the results of BLM and monolingual primary school children. As predicted, spelling skills solely contributed to written narrative textual competence in BLM children, whose L2 spelling skills explained the most significant high proportion of the variance in written performance, even absorbing the influence of oral narrative textual skills in the statistical model. These results extend previous evidence about the relationship between oral language skills and sentence generation skills via spelling in the early stages of learning to write for English-speaking children to BLM children acquiring Italian

as their L2. In our sample, the results show that the group of younger BLM children exhibit significant difficulties in L2 spelling skills in comparison to monolingual peers. Italian monolingual children, as highlighted in the literature, rapidly master spelling skills and begin to be more proficient in expressing their ideas and transferring their mental models into written texts. In this way, the medium of writing did not impact narrative textual competence, and oral narrative textual competence can easily transfer and influence textual writing production. It is important to consider monolinguals' proficiency in spelling skills and related lexical knowledge in relation to their constant exposure to the societal language and the several text-hearing opportunities in the societal language since the early years. In addition, it is important to consider that, in primary school years, the teaching of writing is primarily focused on promoting the acquisition of orthographic accuracy in developing writers who are still facing lower-level transcription skills (spelling and handwriting) to progressively free working memory resources in favor of higher-level cognitive processes (Berninger et al., 1992; McCutchen, 1996; Graham and Harris, 2000) and writing strategies (planning, revising, and monitoring) (Arias-Gundín et al., 2021). Instead, BLM children show sustained L2 spelling difficulties. Such difficulties are a strong obstacle for writing texts in L2 to explain the larger proportion of variance and nullify the age effect. The BLM's spelling difficulties in L2 negatively influenced their written narrative textual competence. When writing texts, children engage in high-level cognitive processes, such as planning, translating, and reviewing, while composing text (Berninger and Swanson, 1994; Hayes, 2012). The lack of automatization of L2 spelling skills in BLM children limits the cognitive resources devoted to high-level processes. Previous results (Vettori et al., 2022) on primary school BLM children show that BLM children fall significantly behind their monolingual peers in textual structure and lexical skills. This result can be linked to limited L2 vocabulary input in BLM children beyond school. It is interesting to consider the detrimental impact of low L2 spelling skills on school-age BLM children's written narrative skills in relation to the different characteristics of Italian (L2 for BLM) and Chinese (L1 for BLM) languages. Italian is a transparent alphabetic language with complex morphology, whereas Chinese is a non-alphabetic language. BLM children are a heterogeneous group with regard to their L1 and L2 proficiency (Lonigan et al., 2018) and home literacy practices and experiences (Peets et al., 2022). As suggested by Wong (2017), children with L1 different from the societal language may show delayed development of their awareness of the adequate narrative structure qualities specific to the target language (Wong, 2017). Otherwise, they may not have the necessary linguistic resources to convey meaning appropriately (Kang, 2012).

In conclusion, our results confirm the importance of exploring direct and mediated relations between oral narrative

textual, spelling skills, and written narrative textual competence (Pinto et al., 2015) and extend previous results focused on the transition between kindergarten to second grade by investigating primary school years. Our results underline the importance of considering how different writing systems pose different challenges for developing writers, especially when considering BLM children who receive limited L2 language input beyond school. Studying how oral and written narrative textual competence relate to each other via the mediating role of spelling skills will yield important implications for educational settings, where knowledge can be employed to design interventions that most effectively sustain L2 written narrative skills in BLM children speaking L1-Chinese at home. The results for monolingual school-age children highlight the need to support oral language, which impacts the quality of children's narrative writing skills. It is also important to note that adequate mastering of spelling skills in the transparent Italian orthography no longer serves as a good discriminator of writing narrative skills in monolinguals.

Limitations and future research

Our results form an important basis for future longitudinal research to gain a more comprehensive understanding of the relations between BLM and monolingual peers' oral narrative, spelling, and written narrative skills. In a future study, a more fine-grained analysis accounting for the type of spelling and morphological errors would further inform the challenges posed by the characteristics of the Italian language to BLM children. In this study, the written narrative skills of children have been assessed using narrative text in L2. Future research can explore more in-depth BLM children's narrative skills in both L1 and L2 to examine these relationships linked to reading and writing habits and motivation (Camacho et al., 2021; De Sixte et al., 2021). Future studies should examine a more heterogeneous language background of BLM children and compare the results from a cross-linguistic perspective. In addition, further research is required to verify the stability of the model detected in this study in BLM children with different language backgrounds.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the University of Florence. Written informed

consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alarcón, R., del Pilar Jiménez, E., and de Vicente-Yagüe, M. I. (2020). Development and validation of the DIGIGLO, a tool for assessing the digital competence of educators. *Br. J. Educ. Technol.* 51, 2407–2421. doi: 10.1111/bjet.12919
- Arias-Gundín, O., Real, S., Rijlaarsdam, G., and López, P. (2021). Validation of the writing strategies questionnaire in the context of primary education: A multidimensional measurement model. *Front. Psychol.* 12:700770. doi: 10.3389/fpsyg.2021.700770
- Berninger, V. W., and Swanson, H. L. (1994). "Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing," in *Children's writing: Toward a process theory of the development of skilled writing*, ed. E. Butterfield (Greenwich, CT: JAI Press), 57–81.
- Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., and Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Read. Writ.* 4, 257–280. doi: 10.1007/BF01027151
- Bigozzi, L., and Vettori, G. (2016). To tell a story, to write it: Developmental patterns of narrative skills from preschool to first grade. *Eur. J. Psychol. Educ.* 31, 461–477. doi: 10.1007/s10212-015-0273-6
- Bigozzi, L., Tarchi, C., and Pinto, G. (2016a). Spelling across tasks and levels of language in a transparent orthography. *PLoS One* 11:e0163033. doi: 10.1371/journal.pone.0163033
- Bigozzi, L., Tarchi, C., Caudek, C., and Pinto, G. (2016b). Predicting reading and spelling disorders: A 4-year prospective cohort study. *Front. Psychol.* 7:337. doi: 10.3389/fpsyg.2016.00337
- Bigozzi, L., Tarchi, C., and Pinto, G. (2017). Consistency and stability of Italian children's spelling in dictation versus composition assessments. *Read. Writ. Q.* 33, 109–122. doi: 10.1080/10573569.2015.1102111
- Camacho, A., Alves, R. A., De Smedt, F., Van Keer, H., and Boscolo, P. (2021). Relations among motivation, behaviour, and performance in writing: A multiple-group structural equation modeling study. *Br. J. Educ. Psychol.* 91, 1456–1480. doi: 10.1111/bjep.12430
- Darlington, R. B., and Hayes, A. F. (2017). *Regression analysis and linear models*. New York, NY: Guilford, 603–611.
- De Sixte, R., Fajardo, I., Mañá, A., Jáñez, Á., Ramos, M., García-Serrano, M., et al. (2021). Beyond the educational context: Relevance of intrinsic reading motivation during COVID-19 confinement in Spain. *Front. Psychol.* 12:703251. doi: 10.3389/fpsyg.2021.703251
- Goodrich, J. M., Farrington, A. L., and Lonigan, C. J. (2016). Relations between early reading and writing skills among Spanish-speaking language minority children. *Read. Writ.* 29, 297–319. doi: 10.1007/s11145-015-9594-8
- Gort, M. (2001). *On the threshold of biliteracy: Bilingual writing processes of English-dominant and Spanish-dominant first graders in a two-way bilingual education program*. Boston, MA: Boston University.
- Graham, S. (2006). "Writing," in *Handbook of educational psychology*, eds P. Alexander and P. Winne (Mahwah, NJ: ER), 457–478.
- Graham, S., and Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educ. Psychol.* 35, 3–12. doi: 10.1207/S15326985EP3501_2
- Haman, E., Wodniecka, Z., Marecka, M., Szweczyk, J., Bialecka-Pikul, M., Otwinowska, A., et al. (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from polish-english migrants to the United Kingdom. *Front. Psychol.* 8:1444. doi: 10.3389/fpsyg.2017.01444
- Hayes, J. R. (2012). Modeling and remodeling writing. *Writ. Commun.* 29, 369–388. doi: 10.1177/0741088312451260
- Hoff, E. (2018). Bilingual development in children of immigrant families. *Child Dev. Perspect.* 12, 80–86. doi: 10.1111/cdep.12262
- Hoff, E., Tulloch, M. K., and Core, C. (2021). Profiles of minority-majority language proficiency in 5-year-olds. *Child Dev.* 92, 1801–1816. doi: 10.1111/cdev.13591
- Howard, E. R., Green, J. D., and Arteagoitia, I. (2012). Can you rid guat ay rot? A developmental investigation of cross-linguistic spelling errors among Spanish-English bilingual students. *Biling. Res. J.* 35, 164–178. doi: 10.1080/15235882.2012.703637
- Incognito, O., Bigozzi, L., Vettori, G., and Pinto, G. (2021). Efficacy of two school-based interventions on notational ability of bilingual preschoolers: A group-randomized trial study. *Front. Psychol.* 12:686285. doi: 10.3389/fpsyg.2021.686285
- Kang, J. Y. (2012). Do bilingual children possess better phonological awareness? Investigation of Korean monolingual and Korean-English bilingual children. *Read. Writ.* 25, 411–431. doi: 10.1007/s11145-010-9277-4
- Kim, Y. S. G. (2020). Structural relations of language and cognitive skills, and topic knowledge to written composition: A test of the direct and indirect effects model of writing. *Br. J. Educ. Psychol.* 90, 910–932. doi: 10.1111/bjep.12330
- Kim, Y. S. G., and Park, S. H. (2019). Unpacking pathways using the direct and indirect effects model of writing (DIEW) and the contributions of higher order cognitive skills to writing. *Read. Writ.* 32, 1319–1343. doi: 10.1007/s11145-018-9913-y
- Kim, Y. S. G., and Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *J. Educ. Psychol.* 109:35. doi: 10.1037/edu0000129
- Lonigan, C. J., Goodrich, J. M., and Farver, J. M. (2018). Identifying differences in early literacy skills across subgroups of language-minority children: A latent profile analysis. *Dev. Psychol.* 54:631. doi: 10.1037/dev0000477
- Mäkinen, L., Gabbatore, I., Loukusa, S., Kunnari, S., and Schneider, P. (2020). A comparison of picture-based narratives by Finnish, Italian and English-speaking children. *Early Educ. Dev.* 31, 395–410. doi: 10.1080/10409289.2019.1666446

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.946142/full#supplementary-material>

- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educ. Psychol. Rev.* 8, 299–325. doi: 10.1007/BF01464076
- MIUR (2012). Indicazioni nazionali per il curricolo della scuola dell'infanzia del primo ciclo di istruzione. (National indications for the curriculum of the nursery school and the first cycle of education). *Ministero dell'istruzione, Univ. e della Ricerca. Ann. Pubblica Istruzione* 2012, 3–88.
- Peets, K. F., Yim, O., and Bialystok, E. (2022). Language proficiency, reading comprehension and home literacy in bilingual children: The impact of context. *Int. J. Biling. Educ. Biling.* 25, 226–240. doi: 10.1080/13670050.2019.1677551
- Pinto, G., Bigozzi, L., Gamannossi, B. A., and Vezzani, C. (2009). Emergent literacy and learning to write: A predictive model for Italian language. *Eur. J. Psychol. Educ.* 24, 61–78. doi: 10.1007/BF03173475
- Pinto, G., Bigozzi, L., Gamannossi, B. A., and Vezzani, C. (2012). Emergent literacy and early writing skills. *J. Genet. Psychol.* 173, 330–354. doi: 10.1080/00221325.2011.609848
- Pinto, G., Tarchi, C., and Bigozzi, L. (2015). The relationship between oral and written narratives: A three-year longitudinal study of narrative cohesion, coherence, and structure. *Br. J. Educ. Psychol.* 85, 551–569. doi: 10.1111/bjep.12091
- Pinto, G., Tarchi, C., and Bigozzi, L. (2016). Peer interaction does not always improve children's mental state talk production in oral narratives. A study in 6-to 10-year-old Italian children. *Front. Psychol.* 7:1669. doi: 10.3389/fpsyg.2016.01669
- Pinto, G., Tarchi, C., and Bigozzi, L. (2020). Improving children's textual competence in kindergarten through genre awareness. *Eur. J. Psychol. Educ.* 35, 137–154. doi: 10.1007/s10212-019-00418-8
- Roch, M., and Hržica, G. (2020). Narrative comprehension by Croatian-Italian bilingual children 5–7 years old. *Dev. Narrat. Comprehension Multiling. Assess. Instrum. Narrat.* 61:171. doi: 10.1075/sibil.61.06roc
- Spinillo, A. G., and Pinto, G. (1994). Children's narratives under different conditions: A comparative study. *Br. J. Dev. Psychol.* 12, 177–193. doi: 10.1111/j.2044-835X.1994.tb00627.x
- Tabachnick, B. G., and Fidell, L. S. (2013). *Using multivariate statistics*, 6th Edn. Boston, MA: Allyn & Bacon.
- UNESCO Institute for Statistics (2012). *International standard classification of education: ISCED 2011*. Montreal, QC: UNESCO Institute for Statistics.
- Vettori, G., Bigozzi, L., Incognito, O., and Pinto, A. G. (2022). Relationship between oral narrative and vocabulary skills among bilingual language-minority children and their monolingual peers in primary school. *Int. J. Biling.* doi: 10.1177/13670069221089064
- Wong, Y. K. (2017). Relationships between reading comprehension and its components in young Chinese-as-a-second-language learners. *Read. Writ.* 30, 969–988. doi: 10.1007/s11145-016-9708-y



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Fabián Román,
Costa University
Corporation, Colombia
Malissa Maria Mahmud,
Sunway University, Malaysia

*CORRESPONDENCE

Nieves Gutiérrez-Ángel
nga212@ual.es
Jesús-Nicasio Sánchez-García
jn.garcia@unileon.es

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 15 March 2022

ACCEPTED 23 May 2022

PUBLISHED 06 September 2022

CITATION

Gutiérrez-Ángel N,
Sánchez-García J-N,
Mercader-Rubio I, García-Martín J and
Brito-Costa S (2022) Digital literacy in
the university setting: A literature
review of empirical studies between
2010 and 2021.
Front. Psychol. 13:896800.
doi: 10.3389/fpsyg.2022.896800

COPYRIGHT

© 2022 Gutiérrez-Ángel,
Sánchez-García, Mercader-Rubio,
García-Martín and Brito-Costa. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction
in other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021

Nieves Gutiérrez-Ángel^{1*}, Jesús-Nicasio Sánchez-García^{2*},
Isabel Mercader-Rubio¹, Judit García-Martín³ and
Sonia Brito-Costa^{4,5}

¹Departamento de Psicología, Área de Psicología Evolutiva y de la Educación, Universidad de Almería, Almería, Spain, ²Departamento de Psicología, Sociología y Filosofía, Universidad de León, León, Spain, ³Departamento de Psicología Evolutiva y de la Educación, Universidad de Salamanca, Salamanca, Spain, ⁴Instituto Politécnico de Coimbra, Coimbra, Portugal, ⁵Coimbra Education School, Research Group in Social and Human Sciences Núcleo de Investigação em Ciências Sociais e Humanas da ESEC (NICSH), Coimbra, Portugal

The impact of digital devices and the Internet has generated various changes at social, political, and economic levels, the repercussion of which is a great challenge characterized by the changing and globalized nature of today's society. This demands the development of new skills and new learning models in relation to information and communication technologies. Universities must respond to these social demands in the training of their future professionals. This paper aims to analyze the empirical evidence provided by international studies in the last eleven years, related to the digital literacy of university students, including those pursuing degrees related to the field of education. Our findings highlight the fact that the digital literacy that is offered in universities to graduate/postgraduate students, in addition to treating digital literacy as a central theme, also focuses on perceived and developed self-efficacy. This is done by strengthening competencies related to digital writing and reading, the use of databases, the digital design of content and materials, and the skills to edit, publish or share them on the web, or applications aimed at treating digital literacy as emerging pedagogies and educational innovation. Secondly, we found studies related to digital competencies and use of the Internet, social networks, web 2.0, or the treatment of digital risks and their relationship with digital literacy. Thirdly, we found works that, in addition to focusing on digital literacy, also focused on different psychological constructs such as motivation, commitment, attitudes, or satisfaction.

Systematic review registration: <https://www.scopus.com/home.uri>; <https://www.recursoscientificos.fecyt.es/>.

KEYWORDS

digital literacy, pre-service & teacher education, higher education, teachers', transversal competences

Introduction

The concept of digital literacy (DL) appears for the first time in the works of Zurkowski (1974), for whom it is an ability to identify, locate, and examine information. However, despite its novelty, the conceptions it encompasses have been changing (Lim and Newby, 2021). Proof of this are the contributions of Gilster (1997) who combines the idea that DL is also closely linked to skills such as access, evaluation, and management of information used in learning processes. Digital learning is understood as the set of technical-procedural, cognitive, and socio-emotional skills necessary to live, learn, and work in a digital society (Eshet-Alkalai, 2012; European Commission, 2018). It is related to reading, writing, calculation skills, and effective use of technology in personal, social, and professional areas. It is also considered inseparable from the social and educational needs of the society in which we live (Larraz, 2013; Brata et al., 2022). Therefore, we refer to a concept that has several aspects including the technological aspect, the informative and multimedia aspect, and the communicative aspect. It involves a complete process and multiple literacies (Gisbert and Esteve, 2011; Lázaro, 2015; Valverde et al., 2022). It requires mastery of certain competencies related to the identification of training needs, access to information in digital environments, the use of ICT tools to manage information, interpretation, and representation of information, and the evaluation of information and the transmission of information (Covello and Lei, 2010; Walsh et al., 2022).

Digital literacy in university students

In recent years, society has undergone enormous changes with the digitalization of many of its spheres at the information level, the communication level, the level of knowledge acquisition, the level of the establishment of social relations, and even the level of leisure. Thus, our habits and means of accessing, managing, and transforming information have also changed (European Union, 2013; Cantabrana and Cervera, 2015; Allen et al., 2020; López-Meneses et al., 2020).

These developments have also had a great impact on the educational field, in which we have to rethink firstly what kind of students we are training in terms of the skills they need in today's society, and secondly, whether we are training a profile of future teachers capable of training a student body that uses information and communication technologies as something inherent to their own personal and social development. In short, digital communication has changed practices related to literacy and has gained great relevance in the development of knowledge in the twenty-first century (Comisión Europea, 2012, 2013; European Commission, 2012; OECD, 2012; Unión Europea, 2013; Instituto Nacional de Tecnologías Educativas y Formación

del Profesorado, 2017; Gudmundsdottir and Hatlevik, 2018; Pérez and Nagata, 2019; Fernández-de-la-Iglesia et al., 2020).

The European Commission (2013) indicates that initial teacher training (ITT) should integrate teachers' digital literacy, betting on the pedagogical use of digital tools, enabling them to use them in an effective, appropriate, and contextualized manner. This teaching competence should be characterized by having a holistic, contextualized, performance-, function-, and development-oriented character. In short, it is about incorporating and adequately using ICT as a didactic resource (Cantabrana and Cervera, 2015; Castañeda et al., 2018; Tourón et al., 2018; Chow and Wong, 2020; Vodá et al., 2022).

In this sense, according to the work of Krumsvik (2009), the CDD (*competencia digital docente de los profesores*—digital competency training for teachers) is composed of four components: basic digital skills (Bawden, 2008), didactic competence with ICT (Koehler and Mishra, 2008; Gisbert and Esteve, 2011), learning strategies, and digital training or training.

While at the Spanish level, the Common Framework of Digital Teaching Competence of the National Institute of Educational Technologies and Teacher Training (INTEF, 2017) standardizes it in five areas: information and information literacy, communication and collaboration, digital content creation, security, and problem solving (López-Meneses et al., 2020). Recently, they have been consolidated as competencies that must be acquired by any university student, along with the knowledge, skills, and attitude that make up a digitally competent citizen (Recio et al., 2020; Indah et al., 2022).

Digital literacy in future teachers

Several efforts have been made to equip future teachers with these competencies through different standards and frameworks to the level of learning acquired (Fraser et al., 2013; INTEF, 2017; UNESCO, 2018). However, how to work these competencies in initial training is still a hotly debated topic, in which special attention is paid to the promotion of experiences of a pedagogical and innovative nature to transform teaching practices, involving the integration of technologies in the classroom, as stated in the Horizon Report 2019 for the Higher Education (Educause, 2019; Le et al., 2022).

Universities are in a moment of transformation, from a teacher-focused teaching model to a model based on active learning through the use of digital technologies, giving rise to a new type of education in which the use of digital devices is intrinsic (Area, 2018; Aarsand, 2019). If digital resources and devices are an inescapable part of current and future teaching practice, digital competency training for future teachers becomes extremely relevant, given that teachers need to acquire these competencies in their initial training to integrate them into their practices as future teachers. That is, the digital competence (DC) acquired during their initial training significantly predicts

the integration of technologies in future teaching practice (Nikou and Aavakare, 2021), which could range from basic digital literacy to the integration of technologies in their daily teaching practice (Gisbert et al., 2016; Alanoglu et al., 2022). Several studies have defined the different indicators that make up DC (Siddiq et al., 2017; González et al., 2018; Rodríguez-García et al., 2019; Cabero-Almenara and Palacios-Rodríguez, 2020).

This calls for a new paradigm, in which future teachers must be digitally literate, in terms of the application of active methodologies, digital competencies, and the use of innovative strategies, styles, and approaches (García-Martin and García-Sánchez, 2017; Gómez-García et al., 2021).

Currently, literacy workshops for future professionals are being carried out in a timely and precise manner from customized short training capsules to specific semester-long subjects in undergraduate or postgraduate studies. The training is focused on several specific aspects of digital literacy, but there is a lack of experience in imparting comprehensive digital training. In addition, there are just a few interactions with professional experts in such literacy (Ata and Yildirim, 2019; Campbell and Kapp, 2020; Domingo-Coscolla et al., 2020; Tomczyk et al., 2020; Vinokurova et al., 2021).

The present study

For the present study, we based our approach on quality and current education, in which DC was postulated as a key element for the development of students. The educational system was tasked with preparing them for their full development and participation in society (OECD, 2011). For this reason, digital literacy is understood as an essential requirement for development in the society in which we live, based on the promotion of strategies related to searching, obtaining, processing, and communicating information. All these aspects have been consolidated as the dimensions of literacy in the twenty-first century (Piscitelli, 2009; Martín and Tyner, 2012). It is, therefore, necessary to understand the reality of this subject and to investigate how these practices are being developed in the context of work. And secondly, it is equally necessary to implement new interventions and lines of research that respond to this urgent need for literacy required by today's society. Therefore, we posed the following research questions: What psychoeducational and learning variables are key in digital literacy? What is the current situation internationally regarding digital literacy in all disciplines in pre-service teacher education? What are the differences in digital literacy requirements pre and post pandemic?

Objective

The objective of this study is to analyze the empirical evidence provided by international studies from 2010 to 2021

related to the digital literacy of university students, including those who are pursuing careers related to the educational field.

Relevant differences will be observed in the contributions in empirical evidence from international studies pre-post-pandemic; and drawn from diverse cultural backgrounds (Spanish-Latin, Portuguese, Finnish, etc.), gender, and personal digital resources.

Materials and methods

The systematic review is composed of four phases, following the model of Miller et al. (2016) and Scott et al. (2018).

PHASE 1: Search terms: In this phase, we developed a schematic of search terms from Web of Science and Scopus databases. We also accessed the databases to locate specific studies that were referenced in the publications that we found in the databases during our initial search. The schematic of terms and thematic axes that were used as a starting point for scanning both databases for anything related to the descriptor “digital” and the descriptor “literacy” is presented in Figure 1.

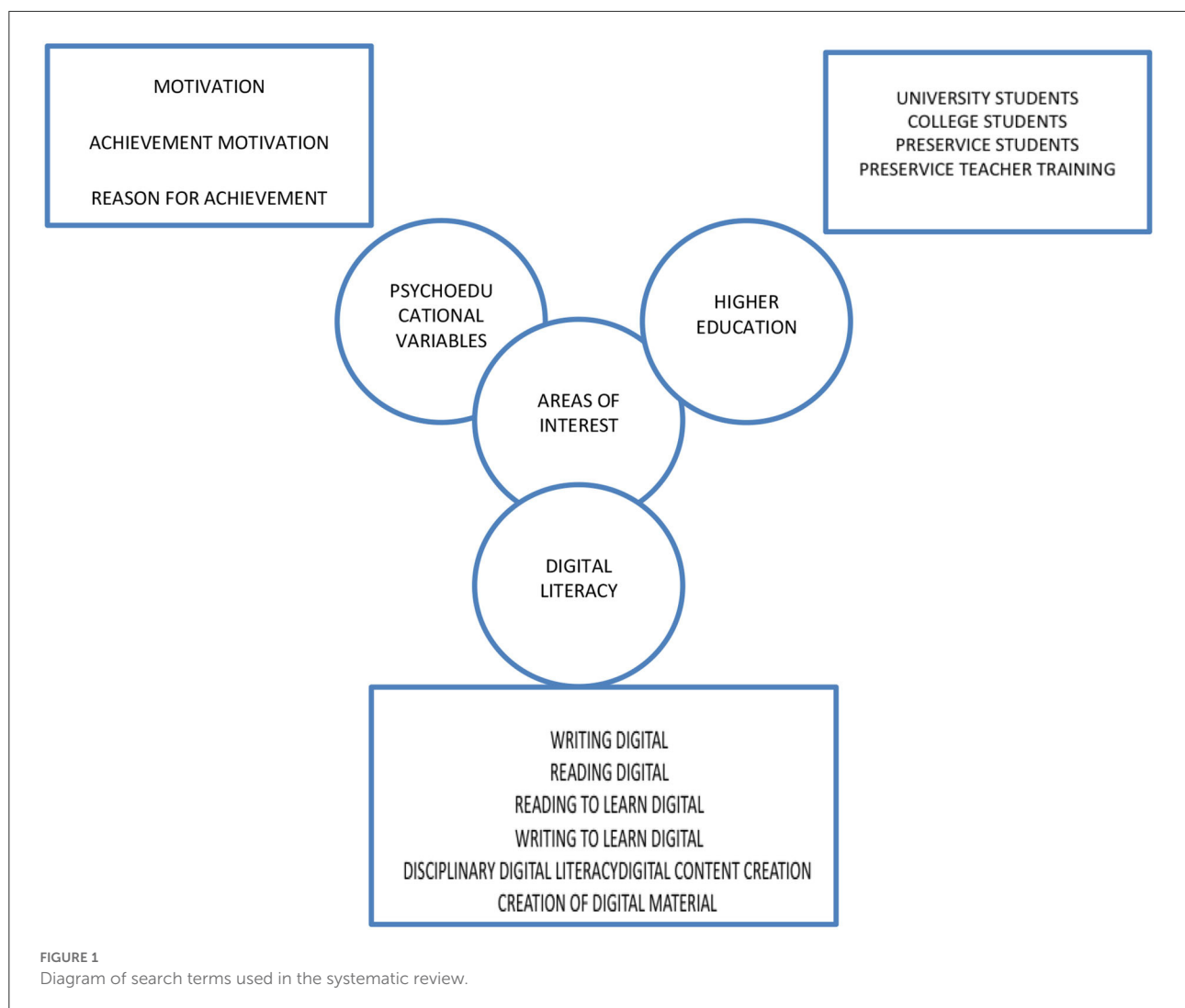
PHASE 2: Selection process based on inclusion and exclusion criteria. The following selection criteria were applied: year of publication between 2010 and 2021, availability of full text, and language of publication in English, Portuguese, or Spanish. Once the first results were obtained, they were selected based on title, abstract, and the use of standardized instruments in their methodology. We rejected the studies that used “*ad hoc*” instruments to measure digital competence.

In addition, the selection indicators provided by Cooper and Hedges (1994) and Cooper (2009) were used, such as peer-reviewed journals, referenced databases, and citation indexes.

PHASE 3: Analysis of methodological quality and indicators based on scientific evidence. Following Torgerson (2007) and Risko et al. (2008) and taking into consideration the MQQn (Risko et al., 2008), we used seven indicators to analyze the quality and effectiveness of the studies (Acosta and Garza, 2011). These were: alignment of theory, findings, reliability and validity, descriptive details of participants and the study, sample, and consistency of findings and conclusions with the data (Risko et al., 2008). Alternatively, evidence-based indicators were also used along with study effect sizes (Díaz and García, 2016; Canedo-García et al., 2017).

PHASE 4: Reliability and outcomes. Reliability was established for both the selection criteria and the coding criteria during each phase, to evidence the replicability of the results. In addition, the results entailed a qualitative analysis of the selected studies, the central arguments, and the evidence provided in a modulated way to address the research questions.

Therefore, the procedure to be followed was documented and charted according to the PRISMA statement (Moher et al., 2009; Page et al., 2021) (see Figure 2). Likewise, an analysis was



undertaken of the key foci in the various studies to highlight the relevant findings and evidence they provided in this regard. The key focus of our work was: first, to analyze the documents related to the digital literacy of university students; second, to identify which variables affect digital literacy; and third, to undertake a comparative analysis between the different variables that were analyzed.

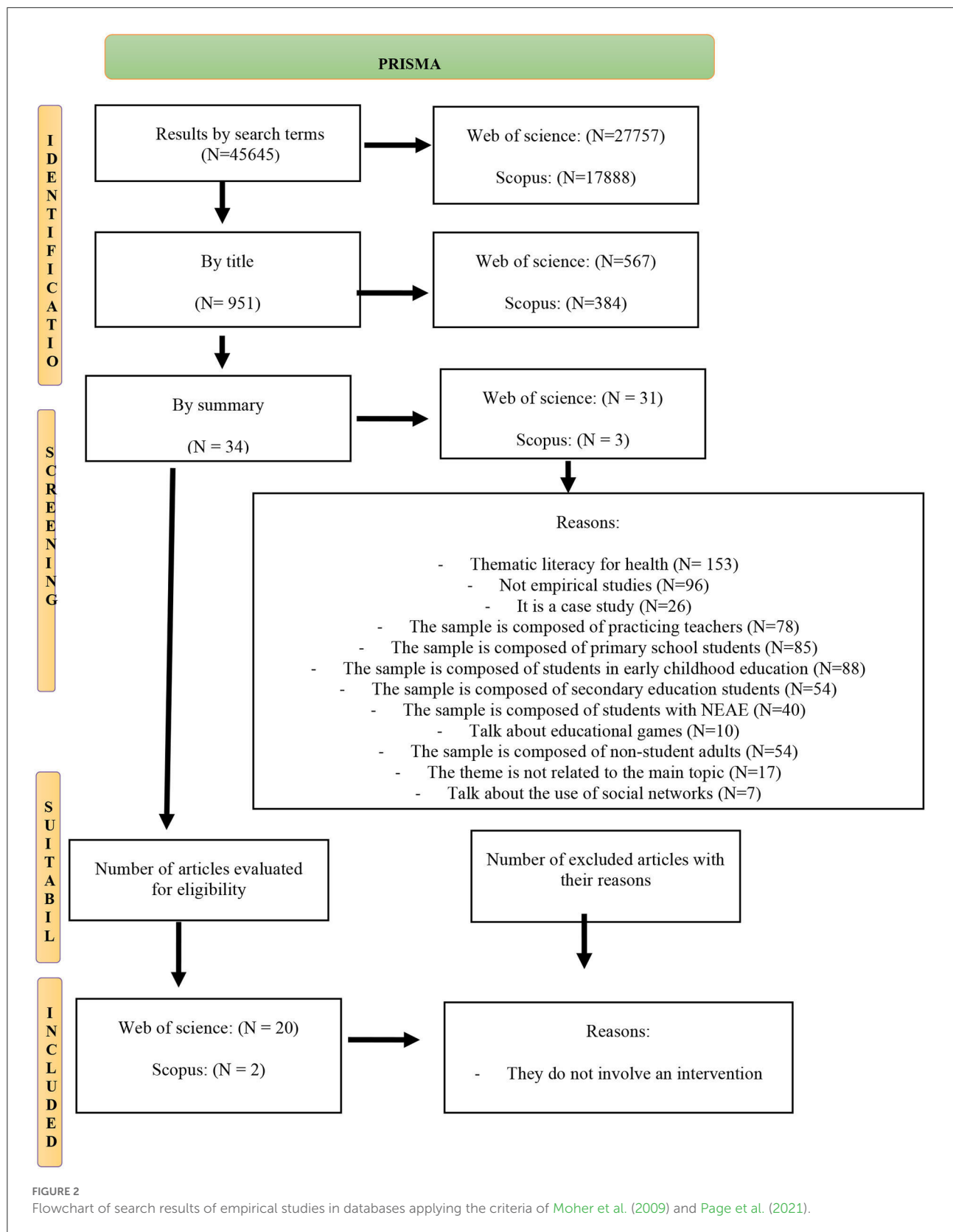
Results

All the selected studies had as samples university students who were pursuing some type of degree or postgraduate degree related to education, and therefore, studying to become future teachers. An intervention design was presented that corresponds to a pre-intervention, the intervention itself, and a post-intervention using techniques such as the activation of prior knowledge, instructions, emulation, and subsequent

tests. We also found studies that had an experimental design assessing control groups and experimental groups (Kajee and Balfour, 2011; Kuhn, 2017; Pequeño et al., 2017; Sharp, 2018; Lerdpornkulrat et al., 2019).

In the case of those responsible for the intervention, practically in all cases, the teacher acts as such, with one or two of them taking the lead. Although the presence of specialized personnel should also be highlighted, as is the case of the work elaborated by Alfonzo and Batson (2014) and Elliott et al. (2018) in which a professional librarian also intervened. Or, in the work detailed by Ball (2019), where a consultant who is not a teacher but a professional expert in the use of digital devices and trained for such an occasion by a responsible brand (Apple) carried out the training at the center.

If we examine the constructs or competencies covered by the works selected in our search, we find that all of them, in addition to dealing with digital literacy, also focus on self-efficacy perceived and developed through digital literacy.



The results of our study could be understood under different themes.

First, we found studies that referred to digital competence and other educational issues. Within them, we found a series of competencies that are emphasized such as digital writing and reading. Research developed from digital media, such as databases, web, or applications aimed at the treatment of digital literacy was noted as emerging pedagogies and educational innovation. The digital design of content and materials and the skills to edit, publish or share them, and competencies related to mathematics and its digital literacy, formed part of digital literacy.

Second, we found studies related to digital competence and the use and employment of the Internet, social networks, web 2.0, and the treatment of digital risks and their relationship with digital literacy.

Third, we found works that in addition to focusing on digital literacy, also focused on different psychological constructs such as motivation, commitment, attitudes, or satisfaction (Tables 1, 2).

Regarding instructional literature, we found a large number of results on mass training programs or courses in which digital literacy was the focus. Examples include a course offered in which students could sign up to, or modules taught during the teaching of a subject. We also found investigations on interventions that had been carried out through different subjects in the study program from where the sample was taken. In this case, the samples were taken on an *ad hoc* basis from a specific student body which the researcher intentionally decided based on a previous intervention experience with them (Ata and Yildirim, 2019; Ball, 2019; Campbell and Kapp, 2020; Domingo-Coscolla et al., 2020; Tomczyk et al., 2020; Vinokurova et al., 2021).

In terms of material resources, all the studies used some type of documentation (digital or not) with instructions on the development of the activities, in which the students were provided with what to do and the steps to follow. In this case, the development scenario was both online and face-to-face, based on different activities given through workshops or seminars for their development.

It should also be noted that in those investigations in which the intervention itself required a specific application or program, the same was used, specifically, and even the intervention had a specific scenario since it was carried out in person in specialized laboratories where experts and specific material was available for this purpose. As an example of these specific materials, in our results, we found the use of the Photo Story 3, Dashboard, and Wikipedia, as well as the EMODO program or the SELI platform (Kajee and Balfour, 2011; Robertson et al., 2012; Ball, 2019; Hamutoglu et al., 2019; Tomczyk et al., 2020).

Regardless of the setting and the program or application employed, we can classify the duration of these interventions into two broad groups: those that had a duration of <1 semester,

and those that had an intervention whose duration ranged from one semester to one academic year.

Regarding the instruments used, it should be noted that most of them used survey forms as an evaluation instrument, either by the researcher or by the students. In addition, it is usually used as a resource to collect information of a personal nature and about one's own experience throughout the intervention. We must also highlight the fact that in many of the results found, this form was used digitally or virtually, abandoning the old paper forms (Kajee and Balfour, 2011; Robertson et al., 2012; Carl and Strydom, 2017; Elliott et al., 2018; Ball, 2019; Lerdpornkulrat et al., 2019; Campbell and Kapp, 2020).

Regarding the use of questionnaires, scales or self-reports, we found several works that used participants' digital literacy histories as instruments. Through them, the researcher could learn first-hand about the sample's personal experience of digital literacy, the previous knowledge they possess, the digital skills they had mastered, those they lack, or those they consider they should improve. It also included the sample's vision regarding the use and employment of digital resources in teaching practice (Kajee and Balfour, 2011; Robertson et al., 2012; Pequeño et al., 2017; Elliott et al., 2018).

In the case of scales, we found two papers that employed a Likert-scale elaborated *ad hoc*. We also found studies that employed standardized scales like the Information Literacy Assessment Scale for Education (ILAS-ED), the Digital Literacy Scale, or the E-Learning Attitudes Scale.

Some of the studies we reviewed used semi-structured interviews as a means of monitoring and providing feedback to the students Table 3; (Kajee and Balfour, 2011; Alfonzo and Batson, 2014; Gill et al., 2015; Carl and Strydom, 2017; Elliott et al., 2018; Elphick, 2018; Ata and Yildirim, 2019; Campbell and Kapp, 2020).

As for the sequence through which the different interventions were developed, we found two types—first, those that divided the contents in time, as is the case of the work of Kajee and Balfour (2011), who covered a first semester digital writing from online classes, self-instructions and face-to-face classes in a specific laboratory, and in a second semester was exposed to different digital research techniques, following the same methodology. In contrast, we spotted the second type, where the same technique was followed throughout the study, as is the case of Robertson et al. (2012). They applied digital stories as a tool for the development of the activity, but also the evaluation of the competency. In the research carried out by Lerdpornkulrat et al. (2019), it is apparent that with the use of the rubric, the teacher gave them an example of the work and asked them all to practice evaluating and grading this work. In this way, they could check if they understood how to use a rubric. They then used the rubric to self-assess their work. After receiving feedback, both groups of students revised and resubmitted their completed projects again.

TABLE 1 Summary of the results found.

Research	Participants					Construct and competence	Instructional procedure	Instructional techniques	Instructional strategies
	Sample	Groups	Design	Sampling and inclusion and exclusion criteria	Teachers				
Alfonzo and Batson (2014)	N = 20 university doctoral students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 2. A teacher and a librarian	Digital literacy/digital research/research software/sdigital databases/self-efficacy	Digital search—apa standards—applications Resource management	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Specific grants Colloquium Planning-Reinforcement Review Selection
Ata and Yildirim (2019)	N = 295 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/internet/social media/perception/digital reading/digital writing/self-efficacy	Training course	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Ball (2019)	Do not specify	Do not specify	Pre-post intervention	Do not specify	Specialized personnel	Digital literacy/digital writing/digital material/creation/editing//media literacy/cybersecurity/self-efficacy	BA Writing and Publishing Program. emphasis on writing, researching, evaluating and reviewing articles in a digital environment	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Botturi (2019)	N = 26 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/access to information/digital content creation/content sharing/self-efficacy	Specific face-to-face program of 2 credits DML education course with 12 2-h sessions	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Campbell and Kapp (2020)	N = 4 university students (future teachers)	Do not specify	Pre-post intervention	Do not specify	N Teachers = 1	Digital literacy/self-efficacy/motivation	Training course Graduate Certificate in Education (PGCE)	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection

(Continued)

TABLE 1 (Continued)

Research	Participants					Construct and competence	Instructional procedure	Instructional techniques	Instructional strategies
	Sample	Groups	Design	Sampling and inclusion and exclusion criteria	Teachers				
Carl and Strydom (2017)	N = 11 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/E-portfolio/self-efficacy/motivation	Digital content design—digital material design	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Domingo-Coscolla et al. (2020)	N = 11 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 11	Digital literacy/diversity/innovation/self-efficacy/motivation	FIMTD project	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Elliott et al. (2018)	N = 48 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Support staff—library staff	Digital literacy/digital writing/digital material/self-efficacy	Module focused on theories of learning and development—sociological module focused on educational inequalities	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Elphick (2018)	N = 949 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/attitude/motivation/and on a day-to-day basis self-efficacy	Use of iPad in education on a day-to-day basis	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection

(Continued)

TABLE 1 (Continued)

Research	Participants					Construct and competence	Instructional procedure	Instructional techniques	Instructional strategies
	Sample	Groups	Design	Sampling and inclusion and exclusion criteria	Teachers				
Gabriele et al. (2019)	N = 141 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/attitude/web 2.0/gamification/self-efficacy	Training course	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Gill et al. (2015)	N = 11 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/pre-preparation/digital knowledge/self-efficacy	Application of practical knowledge from different subjects of the career	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Specific grants Colloquium Planning-Reinforcement Review Selection
Hamutoglu et al. (2019)	N = 47 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/attitude/digital learning/self-efficacy/motivation	Training course once a week for 3 h per week	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Istemic et al. (2016)	N = 115 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/digital content design/digital mathematics/self-efficacy	Creation of digital stories—design of digital content—design of digital materials	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Specific grants Colloquium Planning-Reinforcement Review Selection

(Continued)

TABLE 1 (Continued)

Research	Participants					Construct and competence	Instructional procedure	Instructional techniques	Instructional strategies
	Sample	Groups	Design	Sampling and inclusion and exclusion criteria	Teachers				
Kajee and Balfour (2011)	N = 20 university students (future teachers)	GE = 10 GC = 10	Pre-post intervention	Intentional sampling	N Teachers = 1	Academic Literacy/Digital Writing/Digital Research/Self-Efficacy	Self-instructional/online classes in specific labs	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Specific grants Colloquium Planning-Reinforcement Review Selection
Kuhn (2017)	N = 20 university students (future teachers)	GE = 12 GE2 = 5 GC = 3	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/attitude/digital skills/motivation/autonomy/self-efficacy	Digital Practice and PLE	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Lerdpornkulrat et al. (2019)	N = 584 university students (future teachers)	GE = 321 GC = 263	Pre-post intervention	Intentional sampling	N Teachers = 1	Digital literacy/motivation/self-efficacy	Training course	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Paige et al. (2016)	N = 31 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/digital content design/digital mathematics	Creation of digital stories—design of digital content—design of digital materials	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Slowmation—digital narratives—round tables—interviews—oral evaluations
Pequeño et al. (2017)	N = 54 university students (future teachers)	GE = 31 GC = 24	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/digital narrative/self-efficacy	Application of practical knowledge from different subjects of the career	Activation of previous knowledge-Scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection

(Continued)

TABLE 1 (Continued)

Research	Participants					Construct and competence	Instructional procedure	Instructional techniques	Instructional strategies
	Sample	Groups	Design	Sampling and inclusion and exclusion criteria	Teachers				
Robertson et al. (2012)	N = 150 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	N Teachers = 2	Digital literacy/new pedagogies/multiliteracy/selfstories—thoughtful writing efficacy	Creation of digital	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Specific aid Colloquium Planning-Reinforcement Review Selection -Sharing
Sharp (2018)	N = 51 university students (future teachers)	GE = 20 GE2 = 20 GC = 11	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/attitude/digital skills/motivation/autonomy/—wiki, —microblog self-efficacy	Creation of a blog, —asynchronous discussion,	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Tomczyk et al. (2020)	N = 227 university students (future teachers)	Do not specify	Pre-post intervention	Intentional sampling	Do not specify	Digital literacy/digital inclusion/digital risks/digital content/self-efficacy	SELI Platform	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection
Vinokurova et al. (2021)	Do not specify	Do not specify	Do not specify	Do not specify	Do not specify	Digital literacy/self-efficacy	Training course	Activation of previous knowledge-scaffolding Self-instructions Collaborative/individual emulation Visualization	Colloquium Planning-Reinforcement Review Selection

TABLE 2 Summary of the interventions found.

Research	Materials	Instructor role	Student role	Student grouping	Implementation/ Context	Program duration	Intervention results	Comments
Alfonzo and Batson (2014)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	For 4 days	Greater use of digital tools than before training	Has a sparse sample
Ata and Yildirim (2019)	Does not specify	Teacher—Researcher	Developer of each activity	Great group	Researcher/face-to-face	An academic year	Increasing digital competence	It should apply more evaluation tools
Ball (2019)	Dashboard—training modules—Wikipedia guidelines and rules	Teacher—Researcher	Developer of each activity	Small group	Researcher/face-to-face	An academic year	Increasing digital competence	Does not indicate the method
Botturi (2019)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Great group	Researcher	4 months	Increasing digital competence	Has a sparse sample
Campbell and Kapp (2020)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Great group	Researcher/virtual	5 months	Increasing digital competence	Has a sparse sample
Carl and Strydom (2017)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	Do not specify	Great interest and motivation on the part of the participants	Does not use standardized instruments
Domingo-Coscolla et al. (2020)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Great group	Researcher	Do not specify	Increasing digital competence	Has a sparse sample/does not indicate duration
Elliott et al. (2018)	Weekly Lectures-seminars-online resources-library	Teacher—Researcher	Developer of each activity	Small group	Researcher/face-to-face	An academic year	Increased digital expertise and dominance	Has a sparse sample
Elphick (2018)	Conferences and seminars	Teacher—Researcher	Developer of each activity	Great group	Researcher/face-to-face	One semester	Increased digital expertise and dominance	Does not use standardized instruments

(Continued)

TABLE 2 (Continued)

Research	Materials	Instructor role	Student role	Student grouping	Implementation/ Context	Program duration	Intervention results	Comments
Gabriele et al. (2019)	Power point presentations—introductory videos of the software—brochures—applications created <i>ad hoc</i>	Teacher—Researcher	Developer of each activity	Great group	Researcher/face-to-face	10 months	Increasing digital competence	Has a sparse sample
Gill et al. (2015)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	For 3 years	Practical knowledge of the application of ICT as a learning tool	Has a sparse sample
Hamutoglu et al. (2019)	Texts/documents—EDMODO	Teacher—Researcher	Developer of each activity	Great group	Researcher/face-to-face	5 weeks	Increasing digital competence	Has a sparse sample
Istemic et al. (2016)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	An educational technology course in the academic year 2011–2012	Creation of digital content for the teaching of mathematics	Does not use standardized instruments
Kajee and Balfour (2011)	Texts/documents—computer applications—Laboratory with computers—standalone server—printer	Teacher—Researcher through 40 workstations	Developer of each activity through 40 workstations	Small group/face-to-face	Researcher Specific laboratory	Two semesters of 14 weeks duration	GE improvements greater than GC	Has a sparse sample
Kuhn (2017)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	An academic year	GE1 and GE2 improvements greater than GC	Has a sparse sample
Lerdpornkulrat et al. (2019)	Power point presentations—introductory videos of the software—brochures	Teacher—Researcher	Developer of each activity	Small group	Researcher/face-to-face	13 sessions	Increased self-efficacy in relation to standards and expectations	It should apply more evaluation tools

(Continued)

TABLE 2 (Continued)

Research	Materials	Instructor role	Student role	Student grouping	Implementation/ Context	Program duration	Intervention results	Comments
Paige et al. (2016)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	Do not specify	Creation of digital content for the teaching of mathematics	Does not use standardized instruments
Pequeño et al. (2017)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	An academic year	GE improvements greater than GC	Has a sparse sample
Robertson et al. (2012)	Texts/documents—computer applications—Photo Story 3 program	Teacher—Researcher	Developer of each activity	Small group	Researcher/virtual	For 3 years: 10 months	New learning and means of expression	Has a sparse sample
Sharp (2018)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Small group	Researcher/face-to-face	Two semesters	GE1 and GE2 improvements greater than GC	Has a sparse sample
Tomczyk et al. (2020)	Texts/documents—SELI platform	Teacher—Researcher	Developer of each activity	Great group	Researcher/virtual	Do not specify	Increasing digital competence	Does not indicate the process
Vinokurova et al. (2021)	Texts/documents—specific computer applications—material with indications	Teacher—Researcher	Developer of each activity	Great group	Researcher/virtual	Do not specify	Increasing digital competence	Omits data for possible replicability

TABLE 3 Assessment intervention in the reviewed studies.

Research	Timetable for the implementation of each instrument	Direct comments	Task-specific performance	Overall task performance
Alfonzo and Batson (2014)	Pre-evaluation, post-evaluation and follow-up evaluation using Qual-trics software	Comparison and improvement of the results obtained through the Qual-trics software	Learning the ZOTERO platform at the end of the invention	Mastery of digital bibliographic research and ZOTERO
Ata and Yildirim (2019)	During the intervention	Does not specify	Does not specify	Carecen of digital skills to find, evaluate, create, and communicate
Ball (2019)	During the intervention	Tests throughout the development of the subject through portfolios	Feedback of the results of the questionnaires at the end of each module that showed improvements	Progressive mastery of digital skills
Botturi (2019)	Before and after the intervention	Agree with the participants on the contents and the evaluation	Yields are analyzed practice and evolution	Limited space in the curriculum
Campbell and Kapp (2020)	Before and after the intervention	Learning models and tasks to apply in the classroom	Inclusion of digital competences in curriculum design and monitoring of their development	Differences between resources in centers and in households
Carl and Strydom (2017)	Before and after the intervention	Assessment through direct observation and class visits	Digital learning as part of teacher training	Digital writing support required
Domingo-Coscolla et al. (2020)	Before and after the intervention	Documentary analysis. Discussion groups and finally questionnaires	Digital literacy and content creation	Not all aspects of CDD are measured
Elliott et al. (2018)	Before and after the intervention	Through the delivery of weekly activities	Increased capacity to identify, select and apply digital reading	Not all students developed these skills
Elphick (2018)	Before and after the intervention	Performance is measured through direct observation and scales	Increasing the dominance of digital competence with iPads	A single discipline with a smaller number of staff and students
Gabriele et al. (2019)	Before and after the intervention	feedback on your programming experience and skills from questionnaires	Medium-high level of CT skills, combining design and programming skills	It must be applied in educational practice and not only at the laboratory level
Gill et al. (2015)	Before and after the intervention	3 stages of ict teaching capacity development in which each phase is evaluated	Practice itself as a learning tool	Minimal development where there is no real use of ICT for learning and teaching
Hamutoglu et al. (2019)	Before and after the intervention	Before and after the introduction by standardized instruments	Increased attitudes and skills	Only through EDMODO
Istemic et al. (2016)	Before and after the intervention	Describes the statement design framework and evaluation criteria for solving mathematical and digital problems	Their conceptions changed during the course of passive recipients to active producers of media content.	Control group without intervention
Kajee and Balfour (2011)	Before and after the intervention	Evaluates the results by semesters from accounts or observations	Increasing digital capacity	Large differences in terms of resources
Kuhn (2017)	Before and after the intervention	Evaluate performance through student presentations	Improving your digital skills and abilities	Scarcity of digital tools
Lerdpornkulrat et al. (2019)	Before and after the intervention	Formative assessment and feedback	Increased ability to search, evaluate, process and communicate information	Only the students of the experimental group participated in a formalized activity in the classroom
Paige et al. (2016)	Before and after the intervention	Development of conceptual and semiotic understandings.	Increasing digital literacy in content creation	It is only done with one app
Pequeño et al. (2017)	Before and after the intervention	Narrative research with digital ethnography,	Technological and social mediation	Focused solely on one degree

(Continued)

TABLE 3 (Continued)

Research	Timetable for the implementation of each instrument	Direct comments	Task-specific performance	Overall task performance
Robertson et al. (2012)	Before, during, and after the intervention	Throughout the process, personal reflections on their own experience are requested.	New understanding of literacy, particularly when digital stories are shared as part of the adult classroom experience	Only uses digital stories to gather information from the sample
Sharp (2018)	Before and after the intervention	Performance is evaluated after each practice	Increased perceived levels of confidence and importance of digital literacy	Does not indicate assessment instruments
Tomczyk et al. (2020)	Before and after the intervention	Reflections and own experiences on e-learning at the end of each course	Increasing digital competence	Does not indicate assessment instruments
Vinokurova et al. (2021)	Before, during, and after the intervention	Observation, analysis and pedagogical design and surveys during the intervention	Increasing professional skills, information culture and digital literacy	Insufficient digital resources

In the investigation by Elliott et al. (2018), the intervention was structured in work modules with the following sequence of sessions: they were introduced in the first session with opportunities for group discussions and questions. Essential module reading was provided in weekly online study units and module workshops integrated academic reading and writing activities, such as paraphrasing and referencing, with module content.

In the study by Ball (2019), in the first year, the students took modules on publishing history, culture, markets, and media. In the second year, the intervention was based on their publishing skills, reading for writing development, and grammar and general literacy.

Hamutoglu et al. (2019) organized their intervention in different weeks, such that during the first week of the 14-week semester, the instructor oriented the students for the course and administered pre-tests. In the following week, students were provided with a session on the Edmodo platform and orientation training on the course content.

In the work of Gabriele et al. (2019), the experimental research plan (i.e., activities to be performed, methodology to be adopted) was established over 4 months followed by the organization of the reading material (power point presentations, introductory videos of the software, handouts, *ad hoc* created applications as examples).

We also found interventions that had very short time durations, but provide daily detail of the contents and interventions. Similarly, Alfonzo and Batson (2014) dedicate 1 day to the search and orientation in digital resources, 1 day to the APA standards, and 3 days to develop and use a specific application.

In the research by Istenic et al. (2016), the intervention was based on six different types of tasks related to a variety of mathematical problems, including problems with redundant data, problems with multiple solutions, problems with multiple paths to the solution, problems with no solution, mathematical problems in logic, and problems with insufficient information.

In some interventions, the sequence through which they are developed is the very development of the subject of the degree course from which they are implemented, as is the case of the work of Gill et al. (2015).

In the work of Carl and Strydom (2017), students were first familiarized with the devices and then introduced to electronic portfolios, which helped them to create blogs that serve as platforms for electronic portfolios, and guided them on how to collect artifacts and how to reflect and share content.

In one work we found narrative was used as a technique so that the students could later present their work, analyze it in groups, rework it and present it again to their classmates. Kuhn (2017), Pequeño et al. (2017), and Elphick (2018) followed this model.

Adopting a novel consultative approach, Botturi (2019) co-designed the intervention with his students in two steps: they were surveyed 4 weeks before the start of the course and asked to choose between two options: an overview of different topics/methods/experiences, or an in-depth exploration of one or two topics/methods/experiences. All respondents indicated a preference for the first option and provided indications of the topics they wished to cover (see Tables 4, 5).

The limitations of our search are listed in Table 6. At the theoretical level, we encountered studies that were not very

TABLE 4 Assessment instruments used in the instructional intervention in the reviewed studies.

Research	Questionnaires-self-reports-rating scales-semantic differential	Wallet physical/virtual	Interviews-Reports	Evaluation of the effects of the intervention	Satisfaction	Comments-Individual-Group
Alfonzo and Batson (2014)	Information literacy assessment scale for education (ILAS-ED)	Observations on student work	Does not specify	Post-evaluation of the competencies from the qualtrics software	Learning and satisfaction for participating students	Significant effects on previous methods of instruction
Ata and Yildirim (2019)	Digital literacy scale	Does not specify	Does not specify	The final evaluation confirms the mastery of digital competences	Attitudinal, cognitive and are predictors of digital literacy	Domain alto and positive perceptions of digital literacy
Ball (2019)	Article editing of at least 1,500 words of additional content to the article–500–word report detailing the choice of edits made and the approach used	Edited portfolio	Weekly blog through Pebblepad (an electronic portfolio platform), detailing and explaining the work done that week	1,090 edits in 124 articles, creating six new articles	High capacity for digital editing and publication of content	Mastery and monitoring of competencies after the training course
Botturi (2019)	<i>Ad hoc</i> elaborate Likert scale	Does not specify	Follow-up interviews	Greater digital self-efficacy	Critical assessment of obstacles to implementing DML	Ability to integrate DML
Campbell and Kapp (2020)	Questionnaires that provide background on participants' biographies, perceptions, and experiences with technology	Reflections - justification of their use of technology - narratives of the difficulties experienced	Video recording, semi-structured - focus group interview	Increasing understanding of digital learning possibilities	Complementary tool and means to participate and not as an intentional remedy	Digital non-competition is a barrier today
Carl and Strydom (2017)	<i>Ad hoc</i> elaborate Likert scale	Individual and virtual	Recorded interviews: reflection, training, professional development, and social dimensions of the e-portfolio	Integration of electronic portfolios as tools for reflection	High institutional expectations	Digital growth and development through the use of digital portfolios
Domingo-Coscolla et al. (2020)	<i>Ad hoc</i> elaborate Likert scale	Does not specify	Focus groups	Promoting digital literacy and digital content creation	Insufficient C DD proficiency	Three institutional actions on CDD to be considered in university curricula

(Continued)

TABLE 4 (Continued)

Research	Questionnaires-self-reports-rating scales-semantic differential	Wallet physical/virtual	Interviews-Reports	Evaluation of the effects of the intervention	Satisfaction	Comments-Individual-Group
Elliott et al. (2018)	Essay of 3,000 words on the theories of learning—group oral presentation	Portfolio of 3,000 words. The portfolio was divided into three sections that required students to relate different phases of their personal education experiences to theory.	Semi-structured questionnaires, mainly quantitative, at the beginning and end of the academic year	Difficulties as part of the process	Students' expectations of achievement as the course progressed	Scaffolding strategies with a positive effect on digital self-efficacy
Elphick (2018)	Free text surveys— <i>ad hoc</i> elaborate Likert scale	Does not specify	Semi-structured interview with small groups	Correlations between classrooms rich in technology and digital self-efficacy	The use of iPads has a positive impact on digital behaviors and perceptions about digital skills	Digital competence as a key skill in teachers
Gabriele et al. (2019)	<i>Ad hoc</i> elaborate Likert scale	Does not specify	Tests to check the level of abstraction, parallelism, logistics, synchronization, and control	practical applicability of the intervention	Elaboration of digital material from games with Scratch Software	Increased knowledge and digital skills
Gill et al. (2015)	Interviews developed in 6 phases	Does not specify	Interviews developed in 6 phases	development is proportionate to opportunities to observe and/or use ICT for learning	Classroom experience enables and accelerates the development of digital literacy	The development of digital literacy as a key challenge for future donors
Hamutoglu et al. (2019)	E-Learning attitudes scale—digital literacy scale	Does not specify	Does not specify	Relevant results in terms of avoidance	The trend is one of the most significant predictors of digital literacy skills.	Effectiveness of treatment on participants' attitudes toward e-learning platforms
Istemic et al. (2016)	Performance analysis—analysis of written reflections—pre- and post-test scores-reflections of the participants	Does not specify	Does not specify	Increases in digital pedagogical competences	Instructional approach with digital storytelling and multi-mode design to facilitate learning	Transfer of ICT competencies and their integration into teaching
Kajee and Balfour (2011)	Digital literacy stories of the participants (collected at the beginning of the semester)	Remarks of student work—access and sufficiency	Semi-structured interviews	Digital practice as valuable and social knowledge	Influence of the social context	Digital literacy as a contribution and influence to learning

(Continued)

TABLE 4 (Continued)

Research	Questionnaires-self-reports-rating scales-semantic differential	Wallet physical/virtual	Interviews-Reports	Evaluation of the effects of the intervention	Satisfaction	Comments-Individual-Group
		surveys—journal of researcher's reflections				
Kuhn (2017)	<i>Ad hoc</i> elaborate likert scale	Does not specify	Focus groups	Obtaining new literacies from digital practice	Need for support and guidance in these contents	Redesign of the PLE of the students.
Lerdpornkulrat et al. (2019)	Questionnaires developed <i>ad hoc</i> —standardized questionnaires	Rubric	Does not specify	Developing self-efficacy related to digital literacy	Increase in self-efficacy in information literacy	The rubric as an appropriate tool to measure learning outcomes related to information literacy
Paige et al. (2016)	<i>Ad hoc</i> elaborate Likert scale	Does not specify	Does not specify	experiences and reflections of the PST on Slowmation as an educational tool	Modeling of best practice evaluation tools.	Digital literacy skills development
Pequeño et al. (2017)	Transmedia narratives	Does not specify	Comments and recommendations made in the group work	Transmedia education as a process of technological mediation and social	Digital skills that students incorporate into internships design, analysis, production, and dissemination of transmedia content	Creation and dissemination of transmedia content
Robertson et al. (2012)	Personal digital story	Remarks of student work—journal of researcher's reflections	Does not specify	Digital stories as an appropriate tool for evaluation and reflection	Multi-literacy	Evidence of transformative pedagogy
Sharp (2018)	<i>Ad hoc</i> elaborate likert scale	Does not specify	Does not specify	Increasing prevalence of digital learning environments.	Greater involvement in digital practices	Collaborative digital literacy practices
Tomczyk et al. (2020)	<i>Ad hoc</i> elaborate likert scale	Does not specify	Does not specify	Need for more training	Need for more studies to identify digital gaps	Achievement Learning Autonomy Adaptation
Vinokurova et al. (2021)	Does not specify	Does not specify	Does not specify	Educational paradigm shift in terms of the content of education	Digital transformation	Increased opportunities for teachers to offer and disseminate ICTs if they have good digital literacy

TABLE 5 Treatment fidelity.

Research	Pertinence	Meetings	Feedback	Reliability and validity assessment	Maintenance and generalization	Other controls	Feedback
Alfonzo and Batson (2014)	Horizontal relevance	Does not specify	Feedback to the student at the end of the course	Does not specify	Pre-post-follow-up evaluation	Agreement between observers collecting data	The duration of the workshops is short
Ata and Yildirim (2019)	Horizontal relevance	Does not specify	Feedback to students after the completion of each phase	Reliability Validity	Pre-post-intervention evaluation	A single researcher	Does not indicate the process or sessions
Ball (2019)	Horizontal relevance	Does not specify	Feedback to students after each module	Consistency	Pre-post-intervention evaluation	A single researcher	Does not use standardized instruments
Botturi (2019)	Horizontal relevance	Does not specify	Continuous feedback to students on each task	Consistency	Pre-post-intervention evaluation	A single researcher	Does not use records such as interviews or portfolios
Campbell and Kapp (2020)	Horizontal relevance	Does not specify	Feedback at the end of the intervention	Does not specify	Pre-post-intervention evaluation	A single researcher	Does not indicate the process or sessions
Carl and Strydom (2017)	Horizontal relevance	Does not specify	Feedback to students at the end of the course	Does not specify	Pre-post-intervention evaluation	A single researcher	Does not specify the duration
Domingo-Coscolla et al. (2020)	Horizontal relevance	Does not specify	Feedback to students at the end of the intervention	Reliability Validity	Pre-post-intervention evaluation	Agreement between observers collecting data	Does not use records such as interviews or portfolios
Elliott et al. (2018)	Horizontal relevance	Does not specify	Feedback to students after each session	Reliability Validity Consistency Exploratory factor analysis	Pre-post-intervention evaluation	Agreement between observers collecting data	Does not use standardized instruments
Elphick (2018)	Horizontal relevance	Does not specify	Feedback to students after each session	Consistency	Pre-post-intervention evaluation	A single researcher	Does not use standardized instruments
Gabriele et al. (2019)	Horizontal relevance	Does not specify	feedback on your programming experience and skills from questionnaires	Reliability Consistency Validity	Pre-post-intervention evaluation	Does not specify	Does not use records such as interviews or portfolios
Gill et al. (2015)	Horizontal relevance	Does not specify	Feedback to students in each subject	Reliability Consistency Validity Exploratory factor analysis	Pre-post-follow-up evaluation	Do not specify	Does not apply any self-assessment scale
Hamutoglu et al. (2019)	Horizontal relevance	Does not specify	Feedback to students with the scores of each standardized instrument	Reliability Validity	Pre-post-intervention evaluation	A single researcher	Does not use records such as interviews or portfolios
Istemic et al. (2016)	Horizontal relevance	Does not specify	Feedback to students after completing each task (6)	Reliability Validity	Pre-post-intervention evaluation	Do not specify	Does not apply any self-assessment scale

(Continued)

TABLE 5 (Continued)

Research	Pertinence	Meetings	Feedback	Reliability and validity assessment	Maintenance and generalization	Other controls	Feedback
Kajee and Balfour (2011)	Horizontal relevance	Does not specify	Student feedback at the end of each semester	Does not specify	Pre-post-intervention evaluation	A single researcher	Only applicable within the university and within the laboratory itself
Kuhn (2017)	Horizontal relevance	Does not specify	Continuous feedback after each student presentation	Vaqlidez	Pre-post-follow-up evaluation	Do not specify	Does not use standardized instruments
Lerdpornkulrat et al. (2019)	Horizontal relevance	Does not specify	Feedback from the researcher and self-assessment	Reliability Consistency Validity Exploratory factor analysis	Pre-post-intervention evaluation	A single researcher	Does not use records such as interviews or portfolios
Paige et al. (2016)	Horizontal relevance	Does not specify	Feedback after the intervention	Validity	Pre-post-intervention evaluation	Do not specify	Does not specify the duration
Pequeño et al. (2017)	Horizontal relevance	Does not specify	Feedback after the intervention	Consistency Validity	Pre-post-intervention evaluation	Do not specify	Does not use standardized instruments
Robertson et al. (2012)	Horizontal relevance	Does not specify	Continuous feedback from their own experiences	Does not specify	Pre-post-follow-up evaluation	Agreement between observers collecting data	Does not apply any self-assessment scale
Sharp (2018)	Horizontal relevance	Does not specify	Feedback after the intervention	Consistency Exploratory factor analysis	Pre-post-intervention evaluation	Do not specify	Does not use standardized instruments
Tomczyk et al. (2020)	Horizontal relevance	Does not specify	Feedback after the intervention	Reliability Consistency Validity Exploratory factor analysis	Pre-post-intervention evaluation	Do not specify	Does not use records such as interviews or portfolios
Vinokurova et al. (2021)	Horizontal relevance	Does not specify	Feedback from students through their own experience	Validity	Pre-post-follow-up evaluation	Do not specify	Does not indicate the process or sessions

Indicators and controls used in the instructional intervention in the empirical studies reviewed II.

current, missing research questions or hypotheses, or even missing objectives. At the statistical level, we found several studies had a small or unrepresentative sample.

Analyzing the interventions themselves, we identified a few limitations, especially in those studies that neither indicates the tasks, record the entire process, or lack key information to replicate the intervention. In some studies, key information relating to the person carrying out the intervention was missing, particularly on whether they had the specific training for this purpose. Another limitation that was identified was that very few evaluation strategies were in place to evaluate the interventions (see Table 7).

Similarly, gaps were found regarding ethical controls, where in some studies the main limitation was that ethical controls were non-existent or not specified (Robertson et al., 2012; Istenic et al., 2016; Kuhn, 2017; Elphick, 2018; Ata and Yildirim, 2019; Tomczyk et al., 2020).

Figure 3 shows the evolution over the years of the samples used in each of the studies from 2011 to 2020.

Figure 4 shows the evolution over the years of the controls used in each of the studies from 2011 to 2021.

Discussion

This work aimed to analyze the empirical evidence found in international studies between 2011 to 2021 related to the digital literacy of university students, including those pursuing degrees in education. This objective has been met.

Regarding the first focus related to literacy, this paper highlighted the fact that studies from the West are the most prevalent in this field (Çoklar et al., 2017; Ata and Yildirim, 2019; Hamutoglu et al., 2019; Sujarwo et al., 2022), which correspond to cross-sectional studies, mostly employing instruments such as “the Digital Literacy Scale” developed by Ng (2012), and “the information literacy self-efficacy scale (ILS)” developed by Kurbanoglu et al. (2006). Regarding the level of mastery, the results showed an upper intermediate level of competence in information and digital literacy, communication, and collaboration, but a low intermediate level in terms of digital content creation, particularly in the creation and dissemination of multimedia content using different tools (López-Meneses et al., 2020; Moreno et al., 2020).

Regarding the second focus, digital literacy in university students, this study reviewed the various contributions of other works and found the presence of a competent group in this field, which makes efficient use of both the Internet and digital media (Çoklar et al., 2016; Ata and Yildirim, 2019; Lim and Newby, 2021). However, differences were also found in this collective relating to gender, where women were more competent than men in digital literacy, information literacy, technological literacy, and communicative literacy (Hamutoglu et al., 2019; López-Meneses et al., 2020; Navarro, 2020). However, on the

other hand, we also found studies that revealed particular gender gaps where men showed a higher propensity for DL, while women outperform men in the overall digital literacy test (Ata and Yildirim, 2019). Ata and Yildirim (2019) also found differences in DL between students where university students studying science or mathematics-related majors had higher levels of digital literacy than students majoring in social sciences or psychology fields (Ata and Yildirim, 2019; Chow and Wong, 2020).

And as for the third focus, digital literacy in future teachers, we found a dual use of digital literacy, in its social and leisure aspect (searching or maintaining friendships through social networks, sharing digital content, downloading content, or playing online games), and in its academic aspect (searching in search engines, working through online documents, organizing or synthesizing information from different processors, using computer programs to make presentations, edit images or content, or create audiovisual content (López-Meneses et al., 2020).

The main contribution of this review lies in its comparison between pre/post-pandemic studies, which show a great increase in the use of technologies in the educational world (across the curriculum), and research work focused on measuring the competencies of these devices (Baber et al., 2022). These new investigations have not only followed the line of previous ones but focused on the measurement of digital literacy and its influence on it by variables such as the degree of origin, gender, age, or being a digital native or immigrant (Castañeda-Peña et al., 2015; Çoklar et al., 2016; Castañeda et al., 2018; Ata and Yildirim, 2019; Gür et al., 2019; Hamutoglu et al., 2019; Lerdpornkulrat et al., 2019; González et al., 2020; Navarro, 2020; De Sixte et al., 2021). But there has been an expansion of the topics and variables that are studied in conjunction with digital literacy, among which we find as a novelty, the study of psycho-educational variables such as academic motivation (Chow and Wong, 2020), self-efficacy and motivation (Lerdpornkulrat et al., 2019), effort expectations (Nikou and Aavakare, 2021), and self-concept as a student and as a teacher (Yeşilyurt et al., 2016). The importance attached to the educational field, the identification of different roles or behaviors within the concept of digital literacy that is delimited, or even the types of uses within the concept of digital literacy (López-Meneses et al., 2020; Moreno et al., 2020; Navarro, 2020; Lim and Newby, 2021) are new trends.

Therefore, we can affirm that in this study the research predictions are fulfilled, in that the results found show relevant differences from international studies pre-post pandemic; and by different cultural backgrounds (Spanish Latin, Portuguese, Finnish...), gender, and personal digital resources. In terms of applications for educational practice, these results do not indicate that university students are competent in terms of digital literacy, although they demonstrate some competencies like online information search, information evaluation, information processing, information communication, and

TABLE 6 Limitations of the instructional interventions described in the empirical studies reviewed.

Research	Background limitations	Limitations on participants	Limitations of the instrument	Program limitations	Limitations of results	Discussion on limitations and conclusions	General limitations	Comments
Alfonzo and Batson (2014)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data	Non-grouping	No graphs or tables They do not analyze each variable Not analyzing generalization effects	Does not indicate reliability and validity assessment	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Ata and Yildirim (2019)	The research question is missing	Lack of inclusion and exclusion criteria	No tasks Do not record the entire process	Non-grouping	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Few evaluation strategies
Ball (2019)	The research question is missing Missing assumptions or forecasts	No method	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	No graphs or tables They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Does not indicate the sample
Botturi (2019)	The research question is missing	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Few evaluation strategies
Campbell and Kapp (2020)	The research question is missing Missing assumptions or forecasts	Lack of inclusion and exclusion criteria	No tasks Do not record the entire process	Non-grouping	They do not analyze each variable	Does not indicate reliability and validity assessment current previews	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Carl and Strydom (2017)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping No duration	No graphs or tables They do not analyze each variable Not analyzing generalization effects	Does not indicate reliability and validity assessment	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger

(Continued)

TABLE 6 (Continued)

Research	Background limitations	Limitations on participants	Limitations of the instrument	Program limitations	Limitations of results	Discussion on limitations and conclusions	General limitations	Comments
Domingo-Coscolla et al. (2020)	The research question is missing Missing assumptions or forecasts	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Elliott et al. (2018)	The research question is missing Missing assumptions or forecasts	Lack of inclusion and exclusion criteria Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	No graphs or tables They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Elphick (2018)	The research question is missing Missing assumptions or forecasts	Lack of inclusion and exclusion criteria	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	No number of sessions	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	The application of standardized chords and instruments is lacking. Few evaluation strategies
Gabriele et al. (2019)	Obsolete fonts	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	Only the publication is compared	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Gill et al. (2015)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data	Non-grouping	No graphs or tables They do not analyze each variable Not analyzing generalization effects	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger

(Continued)

TABLE 6 (Continued)

Research	Background limitations	Limitations on participants	Limitations of the instrument	Program limitations	Limitations of results	Discussion on limitations and conclusions	General limitations	Comments
Hamutoglu et al. (2019)	The research question is missing	Lack of inclusion and exclusion criteria	No tasks	Non-grouping	Only the publication is compared	The answer to the research question is not indicated	No ethical controls (informed acceptance to participate, confidentiality...)	Few evaluation strategies
Istemic et al. (2016)	The research question is missing Missing assumptions or forecasts Missing targets	Lack of inclusion and exclusion criteria	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	No graphs or tables They do not analyze each variable Not analyzing generalization effects	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	The application of standardized chords and instruments is lacking. Few evaluation strategies
Kajee and Balfour (2011)	Obsolete fonts The research question is missing Missing assumptions or forecasts	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Not who implemented	No graphs or tables They do not analyze each variable Not analyzing generalization effects	Does not indicate Reliability and Validity Assessment	Key information to replicate the intervention is missing	Sample must be larger
Kuhn (2017)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	No number of sessions Not who implemented	No graphs or tables They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Lerdpornkulrat et al. (2019)	Missing assumptions or forecasts	Lack of inclusion and exclusion criteria	No tasks Do not record the entire process	Does not indicate instruction procedure	No practical and theoretical applications	No explicit limitations	No ethical controls (informed acceptance to participate, confidentiality...)	Does not use the wallet
Paige et al. (2016)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	Non-grouping	No graphs or tables They do not analyze each variable Not analyzing generalization effects	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger

(Continued)

TABLE 6 (Continued)

Research	Background limitations	Limitations on participants	Limitations of the instrument	Program limitations	Limitations of results	Discussion on limitations and conclusions	General limitations	Comments
Pequeño et al. (2017)	The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	No number of sessions Not who implemented	No graphs or tables They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Sample must be larger
Robertson et al. (2012)	Obsolete fonts The research question is missing Missing assumptions or forecasts Missing targets	Reduced sample Non-representative sample	Non-validity and reliability of instruments with their own data Inadequacy of the age course of the instruments Instruments unknown and not provided for in the Annex	Not who implemented	No graphs or tables They do not analyze each variable Not analyzing generalization effects	Does not indicate Reliability and Validity Assessment	It's not an experimental intervention study, it's just a pre-post group Key information to replicate the intervention is missing No ethical controls (informed acceptance to participate, confidentiality...)	The application of standardized chords and instruments is lacking. Few evaluation strategies
Sharp (2018)	The research question is missing Missing assumptions or forecasts	Lack of inclusion and exclusion criteria	Non-validity and reliability of instruments with their own data Instruments unknown and not provided for in the Annex	No number of sessions Not who implemented	No graphs or tables They do not analyze each variable	They do not compare with previous current studies	Key information to replicate the intervention is missing	The application of standardized chords and instruments is lacking. Few evaluation strategies
Tomczyk et al. (2020)	Missing research question Missing assumptions or forecasts	Lack of inclusion and exclusion criteria	No tasks Do not record the entire process	Non-grouping	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Few evaluation strategies
Vinokurova et al. (2021)	The research question is missing Missing assumptions or forecasts Missing targets	Lack of inclusion and exclusion criteria	No tasks Do not record the entire process	Non-grouping	They do not analyze each variable	They do not compare with previous current studies	No ethical controls (informed acceptance to participate, confidentiality...)	Does not indicate the procedure or the participants or the sessions

TABLE 7 Treatment fidelity.

Research	Moment	Comparison of the control group	Sequence of instruction	Previous written protocol	Comparable instructor training	File	Uniform and standard application
Alfonzo and Batson (2014)	Pre During Expose Follow	Evaluate the group in general	3 workshops: Library Orientation, APA style, ZOTERO	Day 1: Library orientation, APA style. Day 2, 3, and 4: ZOTERO	Does not specify	Pre-evaluation, post-evaluation, and follow-up evaluation using qualtrics software	Equal application of the program to all students: same duration, sequence, tasks, and context
Ata and Yildirim (2019)	During Expose Follow	Evaluate the group in general	Does not specify	Does not specify	Does not specify	Does not specify	Equal application of the program to all students: same duration, sequence, tasks, and context
Ball (2019)	During Expose Follow	Evaluate the group in general	Modules of history and editorial culture, markets, and media. Editorial Skills Module, Reading for Writing, and Grammar Development and General Literacy	Does not specify	Does not specify	Portfolios and weekly blog	Equal application of the program to all students: same duration, sequence, tasks, and context
Botturi (2019)	During Expose Follow	Evaluate the group in general	Agreed with students that provided instructions on the topics they wished to cover	Does not specify	Does not specify	Balance	Equal application of the program to all students: same duration, sequence, tasks, and context
Campbell and Kapp (2020)	During Expose Follow	Evaluate the group in general	Does not specify	Does not specify	Does not specify	Questionnaires, portfolio, and interviews	Equal application of the program to all students: same duration, sequence, tasks, and context
Carl and Strydom (2017)	Pre During Expose Follow	They evaluate the group in general although I am divided into two subgroups	Stages: familiarization, indexing, graphing and cartography, and interpretation	Familiarization -blo-share	Does not specify	-Recorded interviews - portfolio	Equal application of the program to all students: same duration, sequence, tasks, and context
Domingo-Coscolla et al. (2020)	During Expose Follow	Evaluate the group in general	Does not specify	Does not specify	Does not specify	Scales and focus groups	Equal application of the program to all students: same duration, sequence, tasks, and context

(Continued)

TABLE 7 (Continued)

Research	Moment	Comparison of the control group	Sequence of instruction	Previous written protocol	Comparable instructor training	File	Uniform and standard application
Elliott et al. (2018)	During Expose Follow	Evaluate the group in general	Sessions with opportunities for group discussions and questions. Module essential reading was provided in weekly online study units	Does not specify	Broader university support from support staff specializing in academic skills in the “learning development team” and library staff.	Questionnaires, essays, and portfolio	Equal application of the program to all students: same duration, sequence, tasks, and context
Elphick (2018)	During Expose Follow	Evaluate the group in general	Conferences and seminars—direct observation—scales—interviews	Does not specify	Training sessions facilitated by an Apple professional Authorized Development Coach	Narratives—presentations—classroom observations—comments and feedback—audiovisual recordings	Equal application of the program to all students: same duration, sequence, tasks, and context
Gabriele et al. (2019)	During Expose Follow	Evaluate the group in general	1. Experimental research plan 2. The reading material was organized (power point presentations, introductory videos of the software, brochures, applications created <i>ad hoc</i> as examples)	Does not specify	Does not specify	Scales and individual tests	Equal application of the program to all students: same duration, sequence, tasks, and context
Gill et al. (2015)	Pre During Expose Follow	Evaluate the group in general	Of the different subjects related to ICT in the career	Of the different subjects related to ICT in the career	Does not specify	Interviews	Equal application of the program to all students: same duration, sequence, context tasks
Hamutoglu et al. (2019)	During Expose Follow	Evaluate the group in general	Preliminary tests of the first week. In the following week session on the Edmodo platform and an orientation training on the content of the course	Does not specify	Does not specify	Two standardized scales	Equal application of the program to all students: same duration, sequence, context tasks
Istemic et al. (2016)	Pre During Expose Follow	Evaluate the group in general	Six tasks	Students completed the pre-test before the start of the study and the subsequent test 15 days later.	Does not specify	Digital Literacy Stories—Pre and Post-Assessment	Equal application of the program to all students: same duration, sequence, context tasks
Kajee and Balfour (2011)	Pre During Expose Follow	Evaluation of the intervention group and another equivalent control group to verify differential efficacy	Semester 1: Digital Writing Semester 2: Digital Research	Does not specify	Does not specify	Digital literacy stories—semi-structured interviews—observations—access and sufficiency surveys—journal of researchers' reflections	Equal application of the program to all students: same duration, sequence, context tasks

(Continued)

TABLE 7 (Continued)

Research	Moment	Comparison of the control group	Sequence of instruction	Previous written protocol	Comparable instructor training	File	Uniform and standard application
Kuhn (2017)	During Expose Follow	Evaluation of the intervention group and another equivalent control group to verify differential efficacy	Scales—exhibition—discussion groups	Does not specify	Does not specify	Narratives—exhibitions—classroom observations—comments and feedback—audiovisual recordings	Equal application of the program to all students: same duration, sequence, context tasks
Lerdpornkulrat et al. (2019)	During Expose Follow	Only the GC participates in a formalized face-to-face activity based on the use of the course rubric as a self-assessment tool	Through the rubric they were able to self-evaluate your own work After receiving feedback, both groups of students reviewed and resubmitted their feedback Complete projects again	Does not specify	Does not specify	Questionnaires developed <i>ad hoc</i> —standardized questionnaires	only the students of the experimental group participated in a formalized activity in the classroom
Paige et al. (2016)	Pre During Expose Follow	Evaluate the group in general	Slowmotion, vivas, digital narratives, roundtables, interviews and oral assessments	Slow	Does not specify	Pre- and post- intervention test—Scale	Equal application of the program to all students: same duration, sequence, context tasks
Pequeño et al. (2017)	During Expose Follow	Evaluation of the intervention group and another equivalent control group to verify differential efficacy	Narrative—characteristics—exhibition—analysis—reworking—exhibition and possibilities	Digital ethnography for examine relations with technologies and the media and how they mediate in the configuration of subjectivities	Does not specify	Narratives—exhibitions—classroom observations—comments and feedback—audiovisual recordings	Equal application of the program to all students: same duration, sequence, context tasks
Robertson et al. (2012)	Pre During Expose Follow	Evaluate the group in general	Digital stories. After the presentation, you are asked to write a written reflection describing your experience	Content analysis and categorization	Does not specify	Digital literacy stories of the—observations—journal of researcher's reflections	Equal application of the program to all students: same duration, sequence, context tasks

(Continued)

TABLE 7 (Continued)

Research	Moment	Comparison of the control group	Sequence of instruction	Previous written protocol	Comparable instructor training	File	Uniform and standard application
Sharp (2018)	During Expose Follow	Evaluation of the intervention group and another equivalent control group to verify differential efficacy	Does not specify	Does not specify	Does not specify	Scales	Equal application of the program to all students: same duration, sequence, context tasks
Tomczyk et al. (2020)	During Expose Follow	Evaluate the group in general	Unspecified	Does not specify	Does not specify	Scale	Equal application of the program to all students: same duration, sequence, context tasks
Vinokurova et al. (2021)	During Expose Follow	Evaluate the group in general	Does not specify	Does not specify	Does not specify	Theoretical analysis of the pedagogical experience, interpretation of scientific data, pedagogical design method (planning, modeling, and conducting classes), and analysis of empirical data in the form of a survey	Equal application of the program to all students: same duration, sequence, context tasks

Indicators and controls used in the instructional intervention in the empirical studies reviewed.

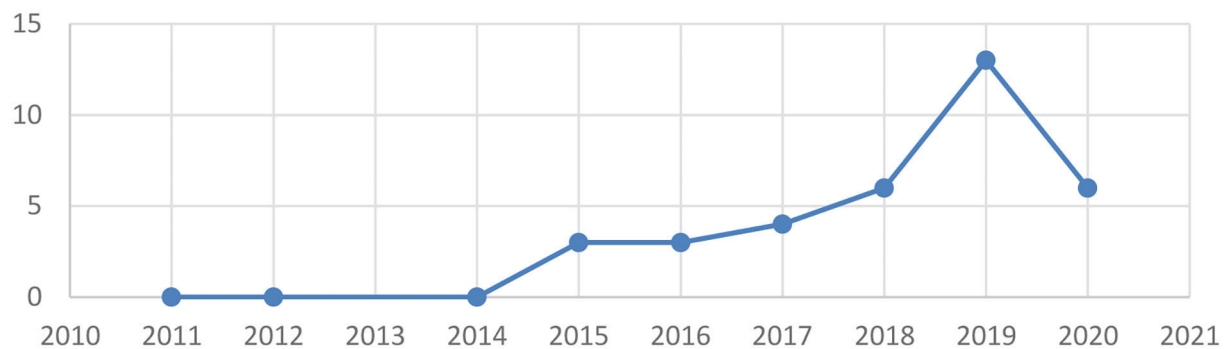


FIGURE 3
Evolution over years of the samples used in the studies from 2010 to 2021.

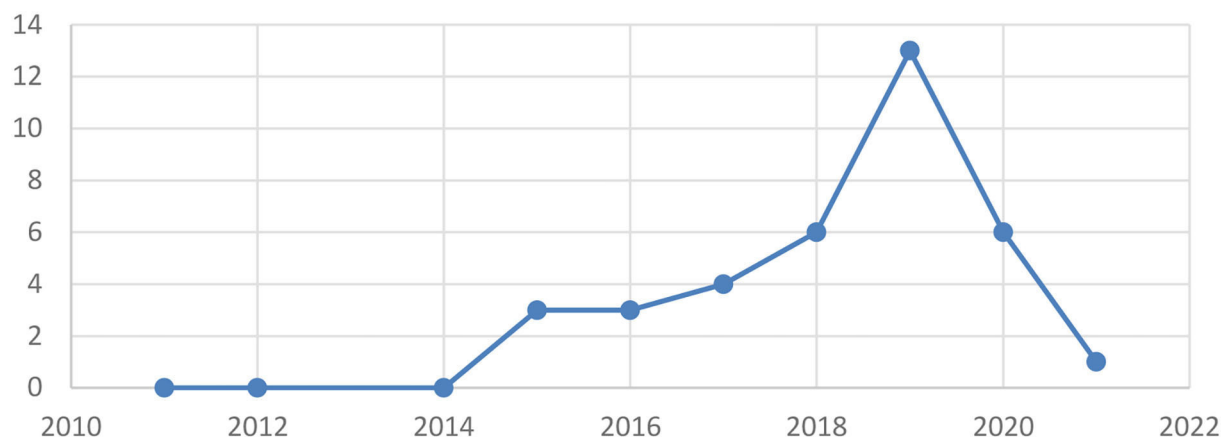


FIGURE 4
Evolution over years of the controls used in studies from 2010 to 2021.

dissemination skills (Çoklar et al., 2016; Lerdpornkulrat et al., 2019). Therefore, there is the risk of training an incomplete student body in digital competence. For complete and comprehensive digital literacy for university students, especially future teachers, there is an urgent need to invest in digital literacy programs. This will ensure that the comprehensive digital competence of students corresponds to the use and employment of the Internet and digital devices in their teaching tasks (Gisbert et al., 2016), and be a guarantee of their integration into teaching practice (Aslan and Zhu, 2016; Nikou and Aavakare, 2021).

As for the limitations of this work, they are closely related to the seven indicators for analyzing study quality and effectiveness (Acosta and Garza, 2011), which are: alignment of theory, findings, reliability and validity, descriptive details of participants, and the study, sample, and consistency of findings and conclusions with the data (Risko et al., 2008). Along with evidence-based indicators, and effect sizes of studies (Díaz and

García, 2016; Canedo-García et al., 2017). So future lines of research or work, should take into account overcoming these limitations, and embrace them in the face of their development.

The number of studies found in the systematic review is comparable to what is usual in this type of study and even higher. For example, in the exemplary systematic review by Scott et al. (2018), they identified only 29 studies that met the quality criteria, reviewing 50 years of studies published in the US, and of these, only four were quantitative. In the study by Borgi et al. (2020), they only found ten studies that fit the criteria in a very good analysis. Other systematic reviews go along the same lines, and in the same journal and section *Frontiers in Psychology*. For example, Dickson and Schubert (2020) and Liu et al. (2022) found only six studies in a review of great interest; the study by Nguyen et al. (2021) identified 18 eligible articles; Shou et al. (2022) with 12 studies included; or Tarchi et al. (2021), Huang (2022) found seven studies for quantitative analysis and eight for indirect evidence; Coxen et al. (2021) with 21 articles

included in the focal analyzes of the systematic review. The number of studies to be representative is not defined by the number but by the existence of such studies. In a systematic review, all studies are reviewed, thus the population of published studies that fit the indicated criteria. With these studies, it was possible to do an analysis of objective indicators in a general comparison between studies; assessing the instruments used; examining the characteristics of the interventions such as strategies, instructional procedure, and psychological variables considered; comparing the fidelity controls of the treatments, which guarantees their rigor and their application in the terms prescribed by the empirical validation of the interventions; and reviewing the limitations of the studies and their contributions by years. These contributions were based on objective data from the studies and have been represented in tables and figures. In addition, a qualitative analysis is provided that highlights the value of intervention studies in relation to digital competence, and the key psychological variables that have been used. It is true that the studies published since 2010 were used, and that there could have been more studies before, but considering the evolution of this type of focus in relation to digital competence and the psychological variables involved, it is evident that the most interesting thing is to consider the recent years which is when its need and use has been generalized throughout the population.

Conclusions

In general, the results show that university students are digitally literate and make efficient use of both the Internet and digital media. In this sense, we found an intermediate or higher level in skills related to communication and collaboration, such as through different chat rooms, platforms, and communication applications. But an intermediate-low level in terms of digital content creation, especially in the creation and dissemination of multimedia content. So, this should be one of the future competencies to increase in this group. Although there are differences according to gender, age, or degree of origin.

We have to invest in comprehensive digital literacy programs for teachers in initial training, which appears implicit in the training plans of their official studies. Digital literacy needs to be a part of the official curriculum, and be developed rather quickly as a separate subject but in an interdisciplinary manner throughout their training. In this way, they become digitally literate people capable of creating and generating digital content and possessing the necessary competencies and skills to use and share such content.

We must also invest in assessing teachers' self-perception. Only by knowing their opinion, skills, and shortcomings, can digital training programs be designed. Digital literacy is a predictor of good digital use and a predictor of the good use

and employment of digital devices and the Internet in the future when they would be teaching.

The findings of this study compel us to consider the following: first, we need to rethink the form and manner in which future teachers are capacitated in digital literacy, if we are doing it in the best way, or if on the contrary there are gaps that should be solved. Second, we should take into account the contributions of the results found and their consequences to formulate effective intervention designs and strategies to effectively capacitate pre-service teachers in digital literacy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

J-NS-G, NG-Á, IM-R, JG-M, and SB-C: conceptualization, methodology, software, writing—review and editing, visualization, supervision, and validation. NG-A: formal analysis, investigation, and resources: UAL, ULE, USAL, IPC, data curation, writing—original draft preparation, and funding acquisition. J-NS-G and NG-A: project administration. All authors contributed to the article and approved the submitted version.

Funding

The general operating funds of the universities have been used Universidad de León (Spain), Universidad de Almería (Spain), Universidad de Salamanca (Spain), Instituto Politécnico de Coimbra and NICSH (Portugal).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Aarsand, P. (2019). Categorization activities in norwegian preschools: digital tools in identifying, articulating, and assessing. *Front. Psychol.* 10, 973. doi: 10.3389/fpsyg.2019.00973
- Acosta, S., and Garza, T. (2011). *The Podcasting Playbook: A Typology of Evidence-Based Pedagogy for PreK-12 Classrooms with English Language Learners*.
- Alanoglu, M., Aslan, S., and Karabatak, S. (2022). Do teachers' educational philosophies affect their digital literacy? The mediating effect of resistance to change. *Educ. Inf. Technol.* 27, 3447–3466. doi: 10.1007/s10639-021-10753-3
- *Alfonzo, P. M., and Batson, J. (2014). Utilizing a co-teaching model to enhance digital literacy instruction for doctoral students. *Int. J. Doc. Stud.* 9, 61–71. doi: 10.28945/1973
- Allen, J., Belfi, B., and Borghans, L. (2020). Is there a rise in the importance of socioemotional skills in the labor market? Evidence from a trend study among college graduates. *Front. Psychol.* 11, 1710. doi: 10.3389/fpsyg.2020.01710
- Area, M. (2018). Hacia la universidad digital: dónde estamos y a dónde vamos? *Rev. Iberoam. Educ. Dist.* 21, 25–30. doi: 10.5944/ried.21.2.21801
- Aslan, A., and Zhu, C. (2016). Investigating variables predicting turkish pre-service teachers' integration of ICT into teaching practices. *Br. J. Educ. Technol.* 48, 552–570. doi: 10.1111/bjet.12437
- *Ata, R., and Yildirim, K. (2019). Exploring Turkish pre-service teachers' perceptions and views of digital literacy. *Educ. Sci.* 9, 40–56. doi: 10.3390/educsci9010040
- Baber, H., Fanea-Ivanovici, M., Lee, Y.-T., and Tinmaz, H. (2022). A bibliometric analysis of digital literacy research and emerging themes pre-during COVID-19 pandemic. *Inform. Learn. Sci.* 123, 214–232. doi: 10.1108/ILS-10-2021-0090
- *Ball, C. (2019). WikiLiteracy: enhancing students' digital literacy with wikipedia. *J. Inform. Liter.* 13, 253–271. doi: 10.11645/13.2.2669
- Bawden, D. (2008). Revisión de los conceptos de alfabetización informacional y alfabetización digital: traducción. *Anal. Document.* 5, 361–408. Available online at: <https://revistas.um.es/analedoc/article/view/2261>
- Borgi, M., Collacchi, B., Giuliani, A., and Cirulli, F. (2020). Dog visiting programs for managing depressive symptoms in older adults: a meta-analysis. *Gerontologist* 60, e66–e75. doi: 10.1093/geront/gny149
- *Botturi, L. (2019). Digital and media literacy in pre-service teacher education: a case study from Switzerland. *Nordic J. Dig. Liter.* 14, 147–163. doi: 10.18261/issn.1891-943x-2019-03-04-05
- Brata, W., Padang, R., Suriani, C., Prasetya, E., and Pratiwi, N. (2022). Student's digital literacy based on students' interest in digital technology, internet costs, gender, and learning outcomes. *Int. J. Emerg. Technol. Learn.* 17, 138–151. doi: 10.3991/ijet.v17i03.27151
- Cabero-Almenara, J., and Palacios-Rodríguez, A. (2020). Marco europeo de competencia digital docente "DigCompEdu". Traducción y adaptación del cuestionario "DigCompEdu Check-In". *Edmeti* 9, 213–234. doi: 10.21071/edmeti.v9i1.12462
- *Campbell, E., and Kapp, R. (2020). Developing an integrated, situated model for digital literacy in pre-service teacher education. *J. Educ.* 79, 18–30. doi: 10.17159/2520-9868/i79a02
- Canedo-García, A., García-Sánchez, J. N., and Pacheco-Sanz, D. I. (2017). a systematic review of the effectiveness of intergenerational programs. *Front. Psychol.* 8, 1882. doi: 10.3389/fpsyg.2017.01882
- Cantabrana, J. L., and Cervera, M. G. (2015). El desarrollo de la competencia digital docente a partir de una experiencia piloto de formación en alternancia en el Grado de Educación. *Educar* 51, 321–348. doi: 10.5565/rev/educar.725
- *Carl, A., and Strydom, S. (2017). e-Portfolio as reflection tool during teaching practice: the interplay between contextual and dispositional variables. *S. Afr. J. Educ.* 37, 1–10. doi: 10.15700/saje.v37n1a1250
- Castañeda, L., Esteve, F., and Adell, J. (2018). Por qué es necesario repensar la competencia docente para el mundo digital? *Rev. Educ. Distan.* 56, 2–20. doi: 10.6018/red/56/6
- Castañeda-Peña, H., Barbosa-Chacón, J. W., Marcales, G., and Barreto, I. (2015). Profiling information literacy in higher education: traces of a local longitudinal study. *Univer. Psychol.* 14, 445–458. doi: 10.11144/Javeriana.upsy14-2.pilh
- Chow, S. K. Y., and Wong, J. L. K. (2020). Supporting academic self-efficacy, academic motivation, and information literacy for students in tertiary institutions. *Educ. Sci.* 10, 361. doi: 10.3390/educsci10120361
- Çoklar, A., Efilci, E., Şahin, Y., and AkÇay, A. (2016). Determining the reasons of technostress experienced by teachers: A qualitative study. *Turkish Online J. Qual. Inq.* 7, 71–96. doi: 10.17569/tojqi.96082
- Çoklar, A. N., Yaman, N. D., and Yurdakul, I. K. (2017). Information literacy and digital nativity as determinants of online information search strategies. *Comput. Hum. Behav.* 70, 1–9. doi: 10.1016/j.chb.2016.12.050
- Comisión Europea (2012). *Informe Conjunto de 2012 del Consejo y de la Comisión sobre la Aplicación del Marco Estratégico Para la Cooperación Europea en el Ámbito de la Educación y la Formación (ET 2020)*. Comisión Europea. Available online at: [https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52012XG0308\(01\)&from=EN](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52012XG0308(01)&from=EN)
- Comisión Europea (2013). *Monitor Education and Training 2013*. Comisión Europea. Available online at: <https://op.europa.eu/es/publication-detail/-/publication/25626e01-1bb8-403c-95da-718c3cfcdf19/language-en>
- Cooper, H. (2009). *Research Synthesis and Meta-Analysis: A Step-By-Step Approach*. London: Sage.
- Cooper, H., and Hedges, L. V. (1994). *The Handbook of Research Synthesis*. New York: Russell Sage.
- Covello, S., and Lei, J. (2010). *A Review of Digital Literacy Assessment Instruments. IDE-712 Front-End Analysis Research*. Syracuse: Analysis for Human Performance Technology Decisions.
- Coxen, L., van der Vaart, L., Van den Broeck, A., and Rothmann, S. (2021). Basic psychological needs in the work context: a systematic literature review of diary studies. *Front. Psychol.* 12, 698526. doi: 10.3389/fpsyg.2021.698526
- De Sixte, R., Fajardo, I., Mañá, A., Jáñez, Á., Ramos, M., García-Serrano, M., et al. (2021). Beyond the educational context: relevance of intrinsic reading motivation during COVID-19 confinement in Spain. *Front. Psychol.* 12, 703251. doi: 10.3389/fpsyg.2021.703251
- Díaz, C., and García, J. N. (2016). Identification of relevant elements for promoting efficient interventions in older people. *J. Psychodidact.* 21, 157–173. doi: 10.1387/RevPsicodidact.13854
- Dickson, G. T., and Schubert, E. (2020). Music on prescription to aid sleep quality: a literature review. *Front. Psychol.* 11, 1695. doi: 10.3389/fpsyg.2020.01695
- *Domingo-Coscolla, M., Bosco, A., Carrasco Segovia, S., and Sánchez Valero, J. A. (2020). Fomentando la competencia digital docente en la universidad: percepción de estudiantes y docentes. *Rev. Invest. Educ.* 38, 167–782. doi: 10.6018/rie.340551
- Educause (2019). *Horizon Report 2019—Higher Education Edition*. Boulder: Educause.
- *Elliott, S., Hendry, H., Ayres, C., Blackman, K., Browning, F., Colebrook, D., et al. (2018). 'On the outside I'm smiling but inside I'm crying.' Communication successes and challenges for undergraduate academic writing. *J. Furth. High. Educ.* 45, 758–770. doi: 10.1080/0309877X.2018.1455077
- *Elphick, M. (2018). The impact of embedded iPad use on student perceptions of their digital capabilities. *Educ. Sci.* 8, 102–125. doi: 10.3390/educsci8030102
- Eshet-Alkalai, Y. (2012). Thinking in the digital era: a revised model for digital literacy. *Iss. Inform. Sci. Inform. Technol.* 9, 267–276. doi: 10.28945/1621
- European Comission. (2012). *The European Commission and bureaucratic autonomy: Europe's custodians*. Cambridge: Cambridge University Press.
- European Comission. (2013). *The European Commission of the Twenty-First Century*. Oxford: Oxford University Press.
- European Commission (Ed) (2018). *Proposal for a Council Recommendation on Key Competences for Lifelong Learning*. New York: European Commission.
- European Union (2013). Digital protest skills and online activism against copyright reform in France and the European Union. *Policy Int.* 5, 27–55.
- Fernández-de-la-Iglesia, J. C., Fernández-Morante, M. C., Cebreiro, B., Soto-Carballo, J., Martínez-Santos, A. E., and Casal-Otero, L. (2020). Competencias y actitudes para el uso de las TIC de los estudiantes del grado de maestro de Galicia. *Publicaciones* 50, 103–120. doi: 10.30827/publicaciones.v50i1.11526
- Fraser, J., Atkins, L., and Richard, H. (2013). *Digilit Leicester. Supporting Teachers, Promoting Digital Literacy, Transforming Learning*. Leicester: Leicester City Council.
- *Gabriele, L., Bertacchini, F., Tavernise, A., Vaca-Cárdenas, L., Pantano, P., and Bilotta, E. (2019). Lesson planning by computational thinking skills in Italian pre-service teachers. *Inform. Educ.* 18, 69–104. doi: 10.15388/infedu.2019.04

- García-Martin, J., and García-Sánchez, J. N. (2017). Pre-service teachers' perceptions of the competence dimensions of digital literacy and of psychological and educational measures. *Comp. Educ.* 107, 54–67. doi: 10.1016/j.compedu.2016.12.010
- *Gill, L., Dalgarno, B., and Carlson, L. (2015). How does pre-service teacher preparedness to use ICTs for learning and teaching develop through their degree program? *Austral. J. Teach. Educ.* 40, 36–59. Available online at: https://www.scopus.com/record/display.uri?eid=2-s2.0-84870756234&origin=inward&txGid=927e8280a577067fddaa8df02fa0b665&featureToggles=FEATURE_NEW_DOC_DETAILS_EXPORT:1
- Gilster, P. (1997). *Digital Literacy* (p. 1). New York, NY: Wiley Computer Pub.
- Gisbert, M., and Esteve, F. (2011). Digital Learners: la competencia digital de los estudiantes universitarios. *La Cuestión Univer.* 7, 48–59.
- Gisbert, M., Esteve, F., and Lázaro, J. (2016). La competencia digital de los futuros docentes: “cómo se ven los actuales estudiantes de educación” *Perspect. Educ.* 55, 34–52. doi: 10.4151/07189729-Vol.55-Iss.2-Art.412
- Gómez-García, G., Hinojo-Lucena, F. J., Fernández-Martín, F. D., and Romero-Rodríguez, J. M. (2021). Educational challenges of higher education: validation of the information competence scale for future teachers (ICS-FT). *Educ. Sci.* 12, 14. doi: 10.3390/educsci12010014
- González, M. J. M., Rivoir, A., Lázaro-Cantabrana, J. L., and Gisbert-Cervera, M. (2020). “Cuánto importa la competencia digital docente? Análisis de los programas de formación inicial docente en Uruguay. *Int. J. Technol. Educ. Innov.* 6, 128–140. doi: 10.24310/innoeduca.2020.v6i2.5601
- González, V., Román, M., and Prendes, M. P. (2018). Formación en competencias digitales para estudiantes universitarios basada en el modelo DigComp. *Rev. Electr. Tecnol. Educ.* 65, 1–15.
- Gudmundsdottir, G. B., and Hatlevik, O. E. (2018). Newly qualified teachers' professional digital competence: implications for teacher education. *Euro. J. Teach. Educ.* 41, 214–231. doi: 10.1080/02619768.2017.1416085
- Gür, D., Canan, Ö., Hamutoglu, N. B., Kaya, G., and Demirtas, T. (2019). The Relationship between lifelong learning trends, digital literacy levels and usage of web 2.0 tools with social entrepreneurship characteristics. *Croat. J. Educ.* 21, 45–76. Available online at: <https://hrcak.srce.hr/220701>
- *Hamutoglu, N. B., Savaşçı, M., and Sezen-Gültekin, G. (2019). Digital literacy skills and attitudes towards e-learning. *J. Educ. Fut.* 16, 93–107. doi: 10.30786/jef.509293
- Huang, C. (2022). Self-Regulation of learning and EFL learners' hope and joy: a review of literature. *Front. Psychol.* 13, 833279. doi: 10.3389/fpsyg.2022.833279
- Indah, R. N., Budhiningrum, A. S., and Afifi, N. (2022). The research competence, critical thinking skills and digital literacy of Indonesian EFL students. *J. Lang. Teach. Res.* 13, 315–324. doi: 10.17507/jltr.1302.11
- Instituto Nacional de Tecnologías Educativas y Formación del Profesorado (2017). *Marco Común de Competencia Digital Docente Octubre 2017*. Instituto Nacional de Tecnologías Educativas y Formación del Profesorado. Available online at: https://aprende.intef.es/sites/default/files/2018-05/2017_1020_Marco-Com%C3%B3n-de-Competencia-Digital-Docente.pdf
- INTEF (Ed.) (2017). *Marco Común de Competencia Digital Docente*. INTEF. Available online at: https://aprende.intef.es/sites/default/files/2018-05/2017_1020_Marco-Com%C3%B3n-de-Competencia-Digital-Docente.pdf
- *Istemic, A., Cotic, M., Solomonides, I., and Volk, M. (2016). Engaging preservice primary and preprimary school teachers in digital storytelling for the teaching and learning of mathematics. *Br. J. Educ. Technol.* 47, 29–50. doi: 10.1111/bjet.12253
- *Kajee, L., and Balfour, R. (2011). Students' access to digital literacy at a South African university: privilege and marginalisation. *S. Afr. Linguist. Appl. Lang. Stud.* 29, 187–196. doi: 10.2989/16073614.2011.633365
- Koehler, M. J., and Mishra, P. (2008). “Introducing tpck,” in *The Handbook of Technological Pedagogical Content Knowledge (tpck) for Educators*, ed AACTE Committee on Innovation and Technology (Mahwah, NJ: Lawrence Erlbaum Associates), 3–29.
- Krumsvik, R. J. (2009). Situated learning in the network society and the digitised school. *Euro. J. Teach. Educ.* 32, 167–185. doi: 10.1080/02619760802457224
- *Kuhn, C. (2017). Are students ready to (re)-design their personal learning environment? The case of the e-dynamic space. *J. N. Approach. Educ. Res.* 6, 11–19. doi: 10.7821/naer.2017.1.185
- Kurbanoglu, S. S., Akkoyunlu, B., and Umay, A. (2006). Developing the information literacy scale. *J. Doc.* 62, 730–743. doi: 10.1108/00220410610714949
- Larraz, E. F. (2013). Competencia digital en la educación superior: instrumentos de evaluación y nuevos entornos. Enl@ ce: *Revista Venezolana de Información. Tecnología y Conocimiento*. 10, 29–43.
- Lázaro, J. (2015). *La Competència Digital Docent Com a Eina per Garantir la Qualitat en l'ús de les TIC en un Centre Escolar* (Tesis doctoral), Tarragona: Universitat Rovira i Virgili.
- Le, B., Lawrie, G. A., and Wang, J. T. (2022). Student self-perception on digital literacy in STEM blended learning environments. *J. Sci. Educ. Technol.* 31, 303–321. doi: 10.1007/s10956-022-09956-1
- *Lerdpornkulrat, T., Poondje, C., Koul, R., Khiawrod, G., and Prasertsirikul, P. (2019). The positive effect of intrinsic feedback on motivational engagement and self-efficacy in information literacy. *J. Psychoeduc. Assess.* 37, 421–434. doi: 10.1177/0734282917747423
- Lim, J., and Newby, T. J. (2021). Preservice teachers' attitudes toward Web 2.0 personal learning environments (PLEs): considering the impact of self-regulation and digital literacy. *Educ. Inform. Technol.* 26, 3699–3720. doi: 10.1007/s10639-021-10432-3
- Liu, W., Xu, Y., Xu, T., Ye, Z., Yang, J., and Wang, Y. (2022). Integration of neuroscience and entrepreneurship: a systematic review and bibliometric analysis. *Front. Psychol.* 13, 810550. doi: 10.3389/fpsyg.2022.810550
- López-Meneses, E., Sirignano, F. M., Vázquez-Cano, E., and Ramírez-Hurtado, J. M. (2020). University students' digital competence in three areas of the DigCom 2.1 model: a comparative study at three European universities. *Austral. J. Educ. Technol.* 36, 69–88. doi: 10.14742/ajet.5583
- Martin, A., and Tyner, K. (2012). Media education, media literacy and digital competence. *Co-municar* 19, 31–39. doi: 10.3916/C38-2012-02-03
- Miller, D. M., Scott, C. E., and McTigue, E. M. (2016). Writing in the secondary-level disciplines: a systematic review of context, cognition and content. *Educ. Psychol. Rev.* 30, 83–120. doi: 10.1007/s10648-016-9393-z
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6, e1000097. doi: 10.1371/journal.pmed.1000097
- Moreno, A. J., Fernández, M. A., and Godino, A. L. (2020). Información y alfabetización digital docente: influencia de la rama formativa. *J. Educ. Teach. Train.* 10, 140–150. Available online at: <https://digibug.ugr.es/handle/10481/59591>
- Navarro, J. A. M. (2020). La competencia digital de los estudiantes universitarios latinoamericanos. *Int. J. Educ. Res. Innov.* 14, 276–289. doi: 10.46661/ijeri.4387
- Ng, W. (2012). Can we teach digital natives digital literacy? *Comput. Educ.* 59, 1065–1078. doi: 10.1016/j.compedu.2012.04.016
- Nguyen, H. T. T., Hoang, A. P., Do, L. T. K., Schiffer, S., and Nguyen, H. T. H. (2021). The rate and risk factors of postpartum depression in vietnam from 2010 to 2020: a literature review. *Front. Psychol.* 12, 731306. doi: 10.3389/fpsyg.2021.731306
- Nikou, S., and Aavakare, M. (2021). An assessment of the interplay between literacy and digital technology in higher education. *Educ. Inform. Technol.* 26, 3893–3915. doi: 10.1007/s10639-021-10451-0
- OECD (2011). *Informe Habilidades y Competencias Del Siglo xxi Para Los Aprendizajes Del Nuevo Milenio En los Países de la OECD*. Available online at: http://recursostic.educacion.es/blogs/europa/media/blogs/europa/informes/Habilidades_y_competencias_siglo21_OCDE.pdf
- OECD (2012). *Education at a Glance 2012. OECD Indicators. Panorama de la educación. Indicadores de la OCDE 2014*. Informe español. Available online at: https://www.oecd-ilibrary.org/education/education-at-a-glance-2012_eag-2012-en
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71. doi: 10.1136/bmj.n71
- *Paige, K., Bentley, B., and Dobson, S. (2016). Slowmation: an innovative twenty-first century teaching and learning tool for science and mathematics pre-service teachers. *Austral. J. Teach. Educ.* 2, 1–15. doi: 10.14221/ajte.2016v41n2.1
- *Pequeño, J. M. G., Rodríguez, E. F., and de la Iglesia Atienza, L. (2017). Narratives transmedia amb joves universitaris. Una etnografia digital en la societat hiperconnectada. *Anàlisi* 57, 81–95. Available online at: <https://raco.cat/index.php/Analisi/article/view/330248>
- Pérez, T. A., and Nagata, J. J. (2019). The digital culture of students of pedagogy specialising in the humanities in Santiago de Chile. *Comput. Educ.* 133, 1–12. doi: 10.1016/j.compedu.2019.01.002
- Piscitelli, A. (2009). *Nativos Digitales*. Mexico: Santillana. doi: 10.26439/contratexto2008.n016.782
- Recio, F., Silva, J., and Marchant, N. A. (2020). Análisis de la competencia digital en la formación inicial de estudiantes universitarios: un estudio de meta-análisis en la web of science. *Rev. Med. Educ.* 59, 125–146. doi: 10.12795/pixelbit.77759

Risko, V. J., Roller, C. M., Cummins, C., Bean, R. M., Block, C. C., Anders, P. L., et al. (2008). A critical analysis of research on reading teacher education. *Read. Res. Q.* 43, 252–288. doi: 10.1598/RRQ.43.3.3

*Robertson, L., Hughes, J., and Smith, S. (2012). “Thanks for the assignment!”: digital stories as a form of reflective practice. *Lang. Liter.* 14, 78–90. doi: 10.20360/G2S88D

Rodríguez-García, A. M., Raso, F., and Ruiz-Palmero, J. (2019). Competencia digital, educación superior y formación del profesorado: un estudio de meta-análisis en la web of Science. *Pixel Bit* 54, 65–81. doi: 10.12795/pixelbit.2019.i54.04

Scott, C. E., McTigue, E. M., Miller, D. M., and Washburn, E. K. (2018). The what, when, and how of preservice teachers and literacy across the disciplines: a systematic literature review of nearly 50 years of research. *Teach. Teach. Educ.* 73, 1–13. doi: 10.1016/j.tate.2018.03.010

*Sharp, L. A. (2018). Collaborative digital literacy practices among adult learners: levels of confidence and perceptions of importance. *Int. J. Instruct.* 11, 153–166. doi: 10.12973/iji.2018.11111a

Shou, S., Li, Y., Fan, G., Zhang, Q., Yan, Y., Lv, T., et al. (2022). The efficacy of cognitive behavioral therapy for tic disorder: a meta-analysis and a literature review. *Front. Psychol.* 13, 851250. doi: 10.3389/fpsyg.2022.851250

Siddiq, F., Gochyyev, P., and Wilson, M. (2017). Learning in digital networks – ICT literacy: a novel assessment of students’ 21st century skills. *Comput. Educ.* 109, 11–37. doi: 10.1016/j.compedu.2017.01.014

Sujarwo, S., Trisanti, T., and Kusumawardani, E. (2022). Digital literacy model to empower women using community-based education approach. *World J. Educ. Technol. Curr. Issues* 14, 175–188. doi: 10.18844/wjet.v14i1.6714

Tarchi, C., Ruffini, C., and Pecini, C. (2021). The contribution of executive functions when reading multiple texts: a systematic literature review. *Front. Psychol.* 12, 716463. doi: 10.3389/fpsyg.2021.716463

*Tomczyk, Ł., Potyrała, K., Włoch, A., Wnek-Gozdek, J., and Demeshkant, N. (2020). Evaluation of the functionality of a new e-learning platform vs. previous experiences in e-learning and the self-assessment of own digital literacy. *Sustainability* 12, 10219. doi: 10.3390/su122310219

Torgerson, K. (2007). A genetic study of the acute anxious response to carbon dioxide stimulation in man. *J. Psychiatric Res.* 41, 906–917. doi: 10.1016/j.jpsychires.2006.12.002

Tourón, J., Martín, D., Navarro, E., Pradas, S., and Íñigo, V. (2018). Validación de constructo de un instrumento para medir la competencia digital docente de los profesores (CDD). *Rev. Española Pedagog.* 76, 25–54. doi: 10.22550/REP76-1-2018-02

UNESCO (2018). *ICT Competency Framework for Teachers*. UNESCO. Available online at: <https://en.unesco.org/themes/ict-education/competency-framework-teachers>

Unión Europea (2013). *Comprender las Políticas de la Unión Europea: Una Nueva Revolución Industrial*. Unión Europea. Available online at: <https://www.europarl.europa.eu/factsheets/es/sheet/61/los-principios-generales-de-la-politica-industrial-de-la-union>

Valverde, J., González, A., and Acevedo, J. (2022). Desinformación y multialfabetización: Una revisión sistemática de la literatura. *Comunicar* 7, 97–110. doi: 10.3916/C70-2022-08

*Vinokurova, N., Mazurenko, O., Prihodchenko, T., and Ulanova, S. (2021). Digital transformation of educational content in the pedagogical higher educational institution. *Rev. Invest. Apuntes Univer.* 11, 1–20. doi: 10.17162/au.v11i3.713

Vodá, A. I., Cautisanu, C., Grădinaru, C., Tănăsescu, C., and de Moraes, G. H. S. M. (2022). Exploring digital literacy skills in social sciences and humanities students. *Sustainability* 14, 2483. doi: 10.3390/su14052483

Walsh, K., Pink, E., Ayling, N., Sondergeld, A., Dallaston, E., Tournas, P., et al. (2022). Best practice framework for online safety education: results from a rapid review of the international literature, expert review, and stakeholder consultation. *Int. J. Child Comput. Interact.* 33, 100474. doi: 10.1016/j.ijcci.2022.100474

Yeşilyurt, E., Ulaş, A. H., and Akan, D. (2016). Teacher self-efficacy, academic self-efficacy, and computer self-efficacy as predictors of attitude toward applying computer-supported education. *Comput. Hum. Behav.* 64, 591–601. doi: 10.1016/j.chb.2016.07.038

Zurkowski, P.G. (1974). The information service environment relationships and priorities. *Relat. Pap.* 5, 27.

* The asterisk of focal references are APA standards.



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Jinfen Xu,
Huazhong University of Science and
Technology, China
Tahereh Taherian,
Yazd University, Iran

*CORRESPONDENCE

Fakieh Alrabai
falrabai@kku.edu.sa

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 30 May 2022

ACCEPTED 22 September 2022

PUBLISHED 14 October 2022

CITATION

Alrabai F and Alamer A (2022) The role
of learner character strengths and
classroom emotions in L2 resilience.
Front. Psychol. 13:956216.
doi: 10.3389/fpsyg.2022.956216

COPYRIGHT

© 2022 Alrabai and Alamer. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

The role of learner character strengths and classroom emotions in L2 resilience

Fakieh Alrabai^{1*} and Abdullah Alamer²

¹Department of English, Faculty of Languages and Translation, King Khalid University, Abha, Saudi Arabia, ²Department of English, King Faisal University, Alhasa, Saudi Arabia

This study aimed to examine a theory-driven model to explain how language learner's trait emotional intelligence (TEI) and effort as two learner character strengths predict learner enjoyment as a positive emotion and anxiety and boredom as two negative classroom emotions, and how these variables, collectively, predict resilience in language learning. The underlying relationship between these variables was tested *via* a comprehensive model within a positive psychology perspective using the partial least squares structural equation modeling (PLS-SEM) approach. The paths in the final structural model indicated that L2 learner TEI did not significantly explain their resilience directly but rather completely indirectly through the mediation of learner negative and positive emotions. Learner effort, directly and indirectly, predicted L2 resilience and its predictive power in it was much larger than that of TEI. In addition, enjoyment and boredom directly influenced L2 resilience and also mediated the relationship between learner character strengths and resilience. Anxiety did not significantly predict learner L2 resilience directly since its influence was rather dependent on the role of enjoyment and boredom in L2 resilience. These findings widely support the claims within positive psychology domain that recognize the vital role of character strengths and learner emotions in enhancing L2 learner resilience.

KEYWORDS

effort, emotional intelligence, negative emotions, positive psychology, positive emotions, resilience

Introduction

Based on the view of [Seligman and Csikszentmihalyi \(2014\)](#), the notion of positive psychology (PP) was founded on three pillars: (1) positive experiences/emotions, (2) positive personality character traits/strengths, and (3) positive institutions. Emotions; the first PP pillar; are multifaceted affective, physiological, behavioral, and cognitive reactions to the different situations experienced in learning situations ([Bielak and Mystkowska-Wiertelak, 2020](#)). According to the broaden-and-build theory ([Fredrickson, 2001, 2009](#)), emotions can be classified into two groups: positive (PEs) and negative (NEs). These different emotions serve different functions and usually have opposing impacts on learning. According to [Fredrickson \(2001, 2009\)](#), NEs such as anxiety and boredom restrict learner experience, decrease resilience and narrow down focus; whereas

PEs like enjoyment broaden experience, build future emotional and cognitive resources (Bielak and Mystkowska-Wiertelak, 2020), and enhance learner resilience (Dewaele and MacIntyre, 2014).

Character strengths are the second pillar of positive psychology (Seligman and Csikszentmihalyi, 2014). The VIA inventory of character strengths (see Park et al., 2004) comprises a list of 24 character strengths classified under 6 broad virtues. Under the virtue of courage in this inventory is the character strength of persistence which involves, according to Kim and Kim (2017), maintaining effort by learners to solve problems in the face of obstacles and difficulties. In light of this definition, effort is a character strength that empowers learners to successfully acquire and learn the language (Hiver et al., 2021; Alamer, 2022a). In the same vein, emotional intelligence (EI), which is conceptualized by Salovey et al. (2009) as the ability to understand feelings in the self and others and to use these feelings as informational guides for thinking and action, is viewed as another key concept of character strengths within PP domain and so is its sub-component trait emotional intelligence (TEI) which is conceived by Petrides et al. (2007, p. 449) as “a constellation of emotional self-perceptions located at the lower levels of personality hierarchies.” Persistent effort in language learning is conceptualized by Lake (2013) as “the amount of time and frequency one spends studying the L2 and persisting in the face of obstacles and difficulties” (p. 230) and TEI have been found to be closely linked to learner emotions. However, the role of these two character strengths in language learner resilience, which is, according to Kim and Kim (2019, p. 3), “the ability to bounce back from adversity,” is yet to be established in second language research. So, relative to the volume of research defining the role of character strengths on LL and because the relationship between L2 learners’ character strengths and classroom emotions identified by previous studies was established solely based on the findings of correlational studies, and also due to the lack of available research on the association between L2 learner character strengths and resilience, investigating the nature of this relationship in greater depth is crucial. This relationship requires to be investigated in one model that would allow unveiling the underlying processes in the relationship among learner character strengths, classroom emotions, and L2 resilience and therefore meaningfully substantiate the mechanisms that underlie such relationship. Hence, by using a structural equation modeling (SEM) approach we examine a theory-driven model to explain how two character strengths (TEI and effort) predict the positive and negative classroom emotions, and how these variables, collectively, predict resilience in language learning. This study has focused on efforts and TEI among the character strength because of the strong connections they have with learner emotions and L2 resilience as will be presented and discussed in the subsequent section of this article. In this regard, the findings of such an examination should support the theoretical claims

that the concepts of TEI and effort are inherently emotion-related constructs and that learner emotions are usually subject to both learner-internal (e.g., TEI and effort) and learner-external regulators as suggested by Dewaele and MacIntyre (2019). A significant contribution of the present study is that it includes, besides learner emotions, learner character strengths and resilience as two important themes that have rarely been discussed in relation to language learning (Oxford, 2016). To the author’s best knowledge, this study is among the first attempts in the L2 domain to investigate the interrelationship between TEI, learner emotions, and learner resilience in a single model. The study findings are anticipated to unveil the complex relationship between EFL learner character strengths and emotions and the role of these variables in predicting learner resilience for language learning.

Literature review

Trait emotional intelligence

TEI is one component of EI besides ability and emotional intelligence. Wellbeing, emotionality, self-control, and sociability are the four sub-factors of general TEI (see Petrides and Furnham, 2003). While wellbeing refers to being happy with life, demonstrating self-confidence, and being optimistic about life, emotionality, nonetheless, represents being empathic, clear about people’s feelings, and having the ability to communicate feelings to others. The self-control factor is characterized by the ability to regulate emotions and stress. Sociability, on the other hand, pertains to the ability to influence other people’s feelings, defend your own rights, and maintain social awareness.

TEI has a very vital role in all domains of knowledge. In the domain of language learning, it has been found to be mostly related to language learner emotions. In this regard, higher levels of TEI were found to be linked to higher levels of learner positive emotions like enjoyment (Dewaele and Mercer, 2018; Li and Xu, 2019; Li, 2020; Resnik and Dewaele, 2020) and lower levels of negative emotions, such as anxiety (Dewaele et al., 2008; Dewaele, 2013; Shao et al., 2013; Li and Xu, 2019; Li, 2020) and boredom (Li et al., 2020). A positive relationship between TEI and enjoyment has been acknowledged by past research. In a 6-week EI intervention study among 56 high school Chinese students, Li and Xu (2019) found that intervention was effective in boosting learners’ positive emotions (e.g., enjoyment), counteracting their negative ones (e.g., anxiety), and improving their EI. In addition, the results of the study by Li (2020) demonstrated that emotional intelligence is a positive predictor of enjoyment, and that enjoyment partially mediated the relationship between TEI and students’ self-reported and actual language performance (Shao et al., 2020).

In addition to its positive links with PEs, research emphasized that TEI is significant in regulating emotions in

that L2 learners with higher emotional intelligence perceived themselves as more capable of gauging the emotions of their interlocutor, controlling their stress, and feeling self-confident compared with those with lower emotional intelligence (e.g., Dewaele et al., 2008; Dewaele, 2013). In the same vein, Lake (2013) added that TEI helps language learners to recognize their own and others' strengths, overcome language obstacles, and obtain optimal affective and learning experiences in the L2 classroom.

TEI does not only help in regulating emotions but further enables individuals to transform negative emotions into positive emotions to reduce stress, anxiety, and conflict; improve relationships; and increase achievement, stability, self-motivation, social awareness, and harmony (see Goleman, 2012; Oxford, 2016). Furthermore, TEI has been established to be positively associated with learner resilience in learning (see Fiorilli et al., 2020; Trigueros et al., 2020).

Language emotions

While past research (e.g., Gkonou et al., 2017) confirms the adverse effects of NEs (e.g., anxiety and boredom) on foreign language (FL) learning outcomes, the role of PEs (e.g., enjoyment) has only recently gained increasing momentum, which has been inspired by advances in the role of PP (Seligman and Csikszentmihalyi, 2000; Dewaele and MacIntyre, 2014, 2016; Gregersen and MacIntyre, 2014; Oxford, 2016; Gkonou et al., 2017). In this regard, research acknowledged that PEs enhance learners' resilience and persistence in facing the problems they encounter while learning a foreign language (Dewaele and MacIntyre, 2014; Dewaele et al., 2019). Additionally, MacIntyre and Gregersen (2012) recognized that PEs support broadening language learner cognition, controlling negative emotions, endorsing resilience, building personal and social resources, and generating greater wellbeing. One of the positive emotions mostly experienced by L2 learners is foreign language enjoyment. Seligman and Csikszentmihalyi, 2000 conceptualized enjoyment as "the good feelings people experience when they break through the limits of homeostasis—when they do something that stretches them beyond what they were" (p. 12). The positive effect of enjoyment in language learning goes beyond creating an enjoyable and safe psychological atmosphere for language learners to promoting their resilience and persistence in dealing with the difficulties they go through in FL learning. A vast line of research in the L2 domain acknowledged that EFL learners with higher levels of enjoyment appear more resilient in L2 learning, thereby acknowledging a positive connection between these two variables (Frederickson et al., 2003; MacIntyre and Gregersen, 2012; Dewaele and MacIntyre, 2014; Oxford, 2016; Dewaele and Alfawzan, 2018; Dewaele et al., 2018; MacIntyre et al., 2019; Shao et al., 2020). The influence of enjoyment in the

course of L2 learning is usually further extended to alleviate the undesirable influence of NES with which enjoyment is often negatively connected. In this respect, learners who demonstrate high enjoyment usually display low levels of foreign language classroom anxiety as recognized by a vast line of L2 research (Dewaele and MacIntyre, 2014, 2016; Dewaele and Dewaele, 2017; Khajavy et al., 2018; Li et al., 2018; Jiang and Dewaele, 2019; Li and Xu, 2019; Elahi et al., 2021) and boredom (Dewaele et al., 2018; Dewaele and Li, 2022; Li and Wei, 2022). Furthermore, enjoyment has also been found to be linked to the effort learners expend in learning a foreign language (Pekrun and Linnenbrink-Garcia, 2014).

One of the most experienced NEs in the language classroom is anxiety. According to, Horwitz (1986), language anxiety is a specific state of anxiety learners experience when they participate in learning and/or using a language. Past studies have shown that anxiety is negatively related to different variables in the process of language learning (Gregersen and MacIntyre, 2014; Jiang and Li, 2017; Saito et al., 2018; Li et al., 2020). In such a process, anxiety has been found to be negatively correlated with positive emotions, such as enjoyment as established by a vast body of literature (e.g., Dewaele and Alfawzan, 2018; Jin and Zhang, 2018; Saito et al., 2018; Li et al., 2020). The research investigated the interplay between these two emotions and found a moderate negative correlation between anxiety and enjoyment. The findings of their study showed that it is possible to experience both a high level of enjoyment and anxiety or neither one nor the other (Resnik and Dewaele, 2020) concluding that enjoyment and anxiety are separate emotions instead of two opposite ends along the same continuum (Li et al., 2020) since high score on one variable does not automatically imply a low score on the other emotion.

Besides its negative associations with PEs, negative correlations have been established by earlier research between TEI and anxiety. For example, in a study involving 464 multilingual learners, Dewaele et al. (2008) found that those who scored higher on TEI reported significantly lower anxiety when using their different languages in a variety of situations attributing that to increased learner confidence in the ability to convey and read their emotions. Research shows that students' higher levels of trait emotional intelligence (TEI) were linked to lower levels of anxiety. In addition, TEI was revealed as a significant predictor of language anxiety (Resnik and Dewaele, 2020).

Similar to its negative connections with PEs and TEI, anxiety holds a negative relationship with L2 learner resilience. According to Chaffee et al. (2014) and Kim and Kim (2017), higher levels of language anxiety are often represented in low resilience in language learning. In this study, researchers maintained that the other side of the relationship is correct in that higher levels of resilience could play an important role in reducing learner anxiety associated with learning an L2. Although anxiety has been recognized as significantly

adversely related to language achievement, researchers contested its negative effect over time (Alamer and Lee, 2021; Sparks and Alamer, 2022). For example, using cross-lagged panel analysis, Alamer and Lee (2021) found that a decrease in anxiety depends on increasing achievement and not the other around.

Foreign Language Classroom Boredom which was conceptualized by Pawlak et al. (2020) “as a negative emotion composed of disengagement, dissatisfaction, attention deficit, altered time perception and decreased vitality” (p. 2) is another NE that is highly associated with negative feelings, such as declined self-regulation, demotivation (Kruk and Zawodniak, 2018), low persistence, low activation and lack of interest, restlessness, anxiety, frustration, helplessness, dislike, unpleasant state of passiveness, guilt, tiredness, sleepiness, disengagement, and dissatisfaction (Li et al., 2021). Li et al. (2021), illuminated that such negative feelings and behaviors usually arise from class activities that are perceived by learners as over-challenging or under-challenging (repetitive, monotonous, undiversified) and/or unimportant, irrelevant, uninteresting, meaningless, and insufficiently stimulating. Past research has recognized the positive association between the negative emotions of anxiety and Boredom in that higher levels of anxiety are usually linked with higher boredom and vice versa (Kruk and Zawodniak, 2018; Li and Dewaele, 2020; Pawlak et al., 2020).

The negative emotions of anxiety and Boredom are often negatively linked with learner resilience whereas highly anxious and bored EFL learners exhibit lower resilience in L2 learning (e.g., Connor and Davidson, 2003; Chen and Padilla, 2019; Shao et al., 2020). In addition, Li and Dewaele (2020) examined the predictive effects of TEI and online learning achievement perceptions on the Boredom of 348 Chinese tertiary students. In this study, TEI and achievement perceptions co-predicted Boredom negatively confirming the negative association among these three constructs.

Effort

Karabiyik and Mirici (2018) conceptualized effort in language learning as the amount of time and energy students invest to learn a foreign language and engaging students to fulfill the requirements of learning this language. According to Alamer (2021) and Hiver et al. (2021), language learners become more successful in learning a foreign language when they invest more effort and penitence in learning this language. In another study, Alamer (2022a) found that the nature and quality of the expended effort are much more important to recognize than the amount of effort being exerted. Investing effort alone could not, nonetheless, lead to successful language learning as established by earlier studies (Oxford and Shearin, 1994;

Gardner, 2010; Alamer, 2022a,b) unless such effort is associated with strong positive psychological emotions for learning and engagement. With regard to learner emotions, the effort that learners invest in learning an L2 has been found to positively affect their enjoyment (Pekrun and Linnenbrink-Garcia, 2014; MacIntyre et al., 2019; Shao et al., 2020) and resilience (Kim and Kim, 2017). In this respect, we believe that effort is generally thought to be a basis for resilience that precedes it, and that the degree of resilience students demonstrate in an L2 is somehow dependent on the amount and quality of effort those students exert in learning this language. On the other hand, the effort is usually negatively related to learner negative emotions in that learners who devote much effort to language learning usually demonstrate lower levels of anxiety (Pinel and Csizér, 2013) and boredom (Pawlak et al., 2020) in the language classroom. However, these conclusions are only theoretically maintained, and it is necessary to examine the complex relationship between effort and language learner emotions and to assess how such a relationship might account for L2 resilience.

L2 resilience

The psychological construct of resilience has been recently adopted in the EFL learning context (Kim et al., 2017). Kim and Kim (2019) emphasized that resilience augments people's ability to overcome difficulties and adversity and to interpret adversity positively through efforts rather than giving up. Among the characteristics that denote L2 learners' classroom resilience, according to Kim et al. (2017), are learner emotional positivity/ happiness (i.e., learners' perceptions of their lives as positive and satisfactory), persistence (paying continuous efforts to solve problems in the face of difficulties), and self-regulation (the ability to regulate one's own thoughts, feelings, and emotions). In EFL learning contexts, resilience has been found to have positive links with positive emotions like enjoyment, negative correlations with negative emotions like anxiety, and positive links with learner TEI as established earlier in this section.

Hypothesized conceptual model

Theoretical links have been established by past research between enjoyment as a positive learner emotion and the character strengths of TEI (Dewaele and Mercer, 2018; Li and Xu, 2019; Li, 2020; Resnik and Dewaele, 2020), and effort (Pekrun and Linnenbrink-Garcia, 2014; MacIntyre et al., 2019; Shao et al., 2020). Accordingly, direct positive paths were depicted from TEI and effort to enjoyment leading to resilience. Besides, direct positive paths were also anticipated from TEI, effort, and enjoyment to resilience in light of the theoretical

assumptions grounded in the literature that higher levels of TEI (Fiorilli et al., 2020; Trigueros et al., 2020), effort (Kim et al., 2017), and enjoyment (e.g., Frederickson et al., 2003; MacIntyre and Gregersen, 2012; Dewaele and MacIntyre, 2014; Oxford, 2016; Dewaele et al., 2018; MacIntyre et al., 2019; Shao et al., 2020) are usually associated with higher levels of resilience in L2 learning.

Because learners with a higher sense of character strengths like TEI and effort were generally found to experience a lower level of negative emotions, such as anxiety and Boredom (see e.g., Dewaele et al., 2008; Dewaele, 2013; Piniel and Csizér, 2013; Shao et al., 2013; Li and Xu, 2019; Li, 2020; Pawlak et al., 2020), negative direct paths from TEI and effort to anxiety and to boredom were drawn.

Since the feelings of language anxiety are usually equipped with a sense of boredom in L2 classes (Kruk and Zawodniak, 2018; Li and Dewaele, 2020; Pawlak et al., 2020), a direct positive path was drawn between these two variables. In addition, due to the well-acknowledged negative associations between negative and positive emotions, negative paths were predicted from anxiety to enjoyment (e.g., Dewaele and Alfawzan, 2018; Saito et al., 2018; Li et al., 2020; Moskowitz and Dewaele, 2020; among many others), as well as from boredom to enjoyment (Dewaele and Li, 2022; Li and Wei, 2022). Likewise, similar negative paths were estimated from anxiety and boredom to resilience based on theoretical argumentations made by past investigations (e.g., Connor and Davidson, 2003; Chen and Padilla, 2019; Shao et al., 2020).

Method

Participants

The present study involved 484 (308 female and 176 male) Saudi undergraduate students studying English at a public Saudi university. They were aged between 18 and 25 years, with a mean age of 20.2 years ($SD = 0.48$). Participants hold similar EFL learning experience ($M = 8.85$ years, $SD = 0.49$) and were studying at similar levels at the department and thus are believed to demonstrate similar levels of English proficiency.

Participants were invited *via* email to participate in this study by completing an online survey. Once consent to participate in the research was granted, a web link containing the online survey was provided to the study participants. Those who showed a willingness to withdraw from the study were advised to simply refrain from completing the questionnaire and leave the web page. An email was sent to all students in the English department, inviting them to participate in the study by completing an online survey *via* a link attached to that email. Those who started filling out the questionnaire and changed their mind about participation were asked to leave the web page.

Measures

Foreign language enjoyment

Learner enjoyment in the FL classroom was assessed based on three factors: private/personal, Social, and Teacher support enjoyment. The 18 items used in this scale were adopted from previous studies on enjoyment (e.g., Dewaele and Alfawzan, 2018; Jin and Zhang, 2018; Li et al., 2018; Jiang and Dewaele, 2019).

Boredom

An 18-item scale was adapted by Li et al. (2021) to assess three factors pertaining to learner boredom in foreign language classes: foreign language classroom boredom, teacher-dislike boredom, and task boredom.

Anxiety

A total of 8 items that were extracted from the foreign language classroom anxiety scale by Horwitz (1986) were used in this study to capture the physical symptoms of learner anxiety, nervousness, and lack of confidence in the FL class. Two anxiety items were phrased to indicate low anxiety and were thus reverse coded so that high scores point to high anxiety for all items on this measurement.

Effort

Learners' effort was assessed using 10 items taken from Gardner's (2010) scale of effort and desire to learn English.

Students' enjoyment, boredom, anxiety, and effort were rated based on a seven-point agreement scale that ranges from totally agree (7) to totally disagree (1); with high scores indicating a high degree of enjoyment and effort and low degree of boredom and anxiety and vice versa. Negatively-worded items were assigned the opposite values in all scales.

Trait emotional intelligence

The Trait emotional intelligence questionnaire—short form (TEIQue-SF) developed by Petrides (2009) and employed by Dewaele (2019) in the EF learning domain was used in this study. This 30-item scale comprises 15 facets (two items per facet) to assess five TEI factors: wellbeing, self-control, sociability, emotionality, and global TEI. Similar to the other scales, students were asked to display the degree of their agreement or disagreement with each statement in this measurement base on a 7-point Likert type scale ranging from "Completely Disagree" (1) to "Completely Agree" (7). Fifteen items of the scale were negatively worded and had to be reverse-coded and, consequently, high scores equal high TEI.

Resilience

A total of 15 items were originally adapted from Shin et al. (2009) and later validated by Kim et al. (2017) and were used in this study to obtain information about L2 learners' resilience. These items pertain to learner happiness (9 items), persistence (3 items), and self-regulation (3 items) and they were measured by a five-point Likert scale, ranging from *strongly disagree* (1) to *strongly agree* (5).

The full survey is available in the online [Supplementary material](#).

Statistical analysis

The current research applied the partial least squares structural equation modeling (PLS-SEM) to assess the hypothesized model. PLS-SEM is an alternative approach to the widely used covariance-based SEM (CB-SEM) in that it focuses on explaining the variance in the outcome variables in the model (Hair et al., 2019). As such, PLS-SEM is most suitable when the research goal is explaining and predicting the outcome variables (exploratory) rather than theory testing (confirmatory) (Hair and Alamer, 2022). It is also a suitable SEM alternative when the model involves formative constructs, i.e., constructs that are defined by their indicators and not the other way around (see Sparks and Alamer, 2022 for a detailed example). In contrast to reflective constructs, formative constructs are evaluated by two major steps: (1) the indicators multicollinearity with the variance inflation factor (VIF) with values below 5 indicating no issue of multicollinearity in the model, and (2) the size and significance of the indicator weights and loadings (Hair and Alamer, 2022).

Because PLS-SEM is a variance-based method, the evaluation of the model relies on several prediction measures, not on the goodness of fit (Hair et al., 2019). However, we report the most commonly used model fit index within the PLS domain, the standardized root mean square residual (SRMR), and its exact fit, with values below 0.08 indicating acceptable fit (Hair and Alamer, 2022). Two major criteria were used: the explanatory power, which is assessed *via* the coefficient of determination (the R^2 value), and the predictive power, which is assessed through the PLS_{predict} procedure (Shmueli et al., 2019). The R^2 value is interpreted using Hair and Alamer (2022) as follows: values of 0.06, 0.16, and 0.36 are indicative of small, medium, and strong explanatory power, respectively. To provide information about model out-of-sample predictive capability, PLS_{predict} procedures "incorporate model assessments based on an initial training sample (randomly drawn separate sub-sample of the total sample) and estimate the predictive power of the model on a second hold-out sample of data—other than that used in calculating the initial PLS-SEM solution" (Hair and Alamer, 2022, p. 8). In this way, PLS structural model has a predictive power if the errors (i.e., RMSE) it produces are

less than the errors produced by the naïve linear regression model (LM) (for greater details see Shmueli et al., 2019). An assessment of collinearity was also considered using the VIF index. Also, following Hair and Alamer's (2022) guidelines in interpreting the effect size of the structural model, β values in the ranges of 0–0.1, 0.1–0.3, and 0.3–0.5, and those that are >0.5 are indicative of weak, modest, moderate, and strong effect sizes, respectively.

Results

Missing values, outliers, and normality

The data of the present study did not contain any missing values. Another check through Q-Q plots has been done to check if there were any outlier values in the data that depart significantly from the rest. The results did not show any problematic values. Normality was assessed by inspecting skewness and kurtosis values following the ± 2 guidelines (Hair et al., 2022). The data did not show a departure from these cut-off values.

Assessing the measurement model

As shown in Table 1, the data appears to be relatively normally distributed. In the structural model, the constructs were measured formatively. That is, specific constructs (e.g., private/personal enjoyment, social enjoyment, and Teacher support enjoyment) are specified to affect the general construct (e.g., enjoyment). Because these sub-factors are not interchangeable it was believed that formative formulation is a more appropriate model specification. Following Hair and Alamer (2022) recommendations, we evaluated the validity of the model through collinearity and the weights of formative indicators. All VIF values were below the cut-off value of 5 and only two indicators, "emotional regulation" and "emotional positivity" exceeded 3 but did not reach 5. We concluded that there is no critical issue of multicollinearity in the model and that the indicators uniquely measure their constructs. With respect to significance testing, we found that all outer weights have shown positive and significant weights with the exception of three constructs "Task boredom," "Persistence" and "Self-control" as their outer weights were found to be non-significant. We inspected their outer loadings and found that their loadings were >0.60 with $p < 0.001$. We retain these indicators and conclude that they are important but not substantially relevant. Note that the assessment of redundancy analysis (i.e., convergent validity) for the formative constructs was not possible due to the lack of a global single-item measure in our data.

TABLE 1 Descriptive statistics, normality, and correlation matrix.

	TEI	Resilience	Effort	Enjoyment	Anxiety	Boredom
TEI	4.66					
Resilience	0.28***	3.54				
Effort	0.45***	0.39***	3.76			
Enjoyment	0.39***	0.59***	0.45***	3.74		
Anxiety	−0.25***	−0.25***	−0.24***	−0.31***	3.05	
Boredom	−0.36***	−0.52***	−0.39***	−0.58***	0.53***	2.88
Skew/Kurt	0.35/−0.54	−0.43/−0.69	−0.66/1.22	−0.29/−0.75	−0.01/−0.65	−0.31/−0.56

Assessing the structural model

The assessment of the structural model presented in Figure 1 starts by reporting the SRMR fit index. The results show that the SRMR of our structural model was 0.07 (HI 95% CI = 0.03), suggesting an approximate fit to the data. We also evaluate the collinearity by inspecting the VIF values. The results showed that all values were below 3, with a maximum value was 2.40 which is way below the cut-off value of 5. We conclude that empirical evidence suggests that the constructs are distinct, thus discriminant validity is established. Next, we assessed the explanatory power of the model with a special focus on the outcome variable, “resilience.” The results of the structural model can be seen in Figure 2. In this way, the analysis showed that the model explained around 50% [95% CI: 40, 0.55] of the variance in the outcome. According to Hair and Alamer (2022), this value indicates a strong explanatory power of the model for the outcome. The results of the PLS_{predict} have been, then considered. Because the outcome variable consists of three indicators, the PLS_{predict} assesses the RMSE values on these indicators. Table 2 presents the results of the RMSE in the two models and shows that all indicators in the PLS model have generated the almost exactly same amount of prediction error. Following Hair and Alamer (2022), the results of PLS_{predict} suggest that the structural model has a moderate out-of-sample predictive power.

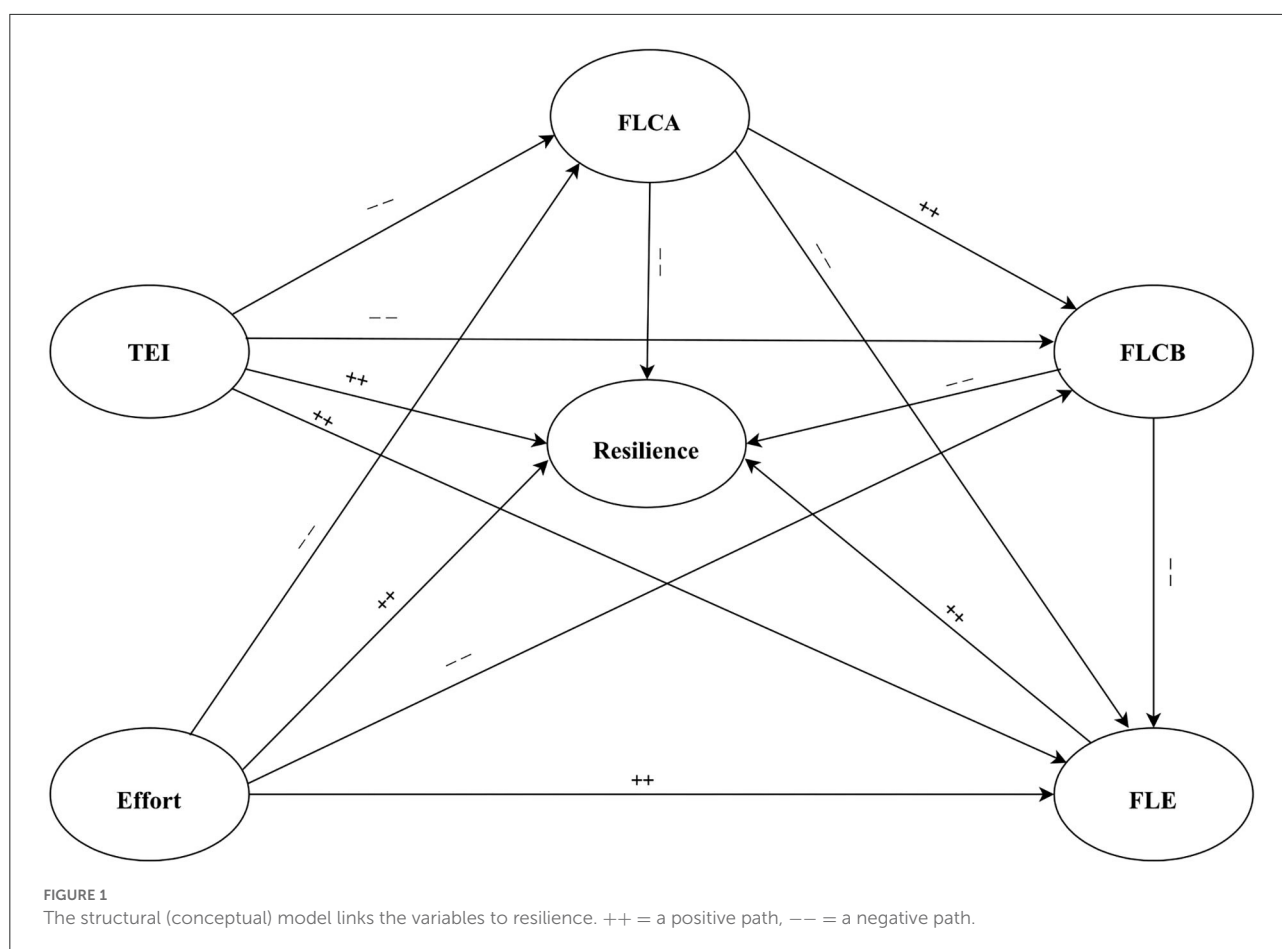
After assessing the overall model, we evaluated the path coefficients for their size and significance. Direct effects between the variables are illustrated in Figure 2 while indirect and total effects are presented in Table 3. Several pathways were found meaningful in the model, and we highlight these accordingly (see Table 3 for full details of mediation). The model indicates that TEI did not relate directly to resilience. Its effect, however, was fully mediated through learner emotions of enjoyment, anxiety, and boredom (see Table 3 for full details of mediation). Moreover, the total effect of TEI on resilience was significant but modest in size ($\beta = 0.13$, $p = 0.003$). Similarly, anxiety was not related directly to resilience or enjoyment. Its effect on the outcome ‘resilience’ was only indirectly detected *via* boredom and enjoyment. Anxiety total effect was positive but modest

in magnitude ($\beta = -0.13$, $p = 0.004$). Furthermore, ‘effort’ appears to be associated with resilience directly and indirectly through boredom and enjoyment. The total effect of ‘effort’ was moderate in size ($\beta = 0.36$, $p < 0.001$). Moreover, the effect of boredom on the outcome ‘resilience’ was direct and indirect through enjoyment. Its total effect was strong ($\beta = 0.50$, $p < 0.001$). In addition, enjoyment has a positive and strong effect on resilience ($\beta = 0.53$, $p < 0.001$), and no indirect effect was postulated with this variable.

Discussion

Previous research has established that TEI of language learners is associated with positive (Chow et al., 2018; Dewaele and Alfawzan, 2018; Saito et al., 2018) and negative emotions (Dewaele et al., 2019; Pawlak et al., 2020; Li et al., 2021) as well as their resilience in language learning (Fiorilli et al., 2020; Trigueros et al., 2020). Despite that, only a few studies have illustrated the mediational process of how character strengths might facilitate learners’ L2 emotions and eventually their L2 resilience. The primary goal of the present research was to identify and elaborate on the interrelationship among L2 learner character strengths of TEI and effort, classroom emotions, and resilience in language learning *via* a comprehensive model from a positive psychology perspective. To fully understand the complex relationship between the constructs, we applied PLS-SEM to study the structural relationships and assess their external validity through predictive assessment (Alamer, 2022b; Hair and Alamer, 2022). Our selection of PLS-SEM was also justified given the nature of the constructs involved in the assessment as they were operationalized as formative constructs. Specifically, using higher-order factor models where the lower-order factors are affecting (or causing) the higher-order factor are not directly possible in the way we tested the mode in CB-SEM. Thus, PLS-SEM was an appropriate selection for such a case (Hair et al., 2019).

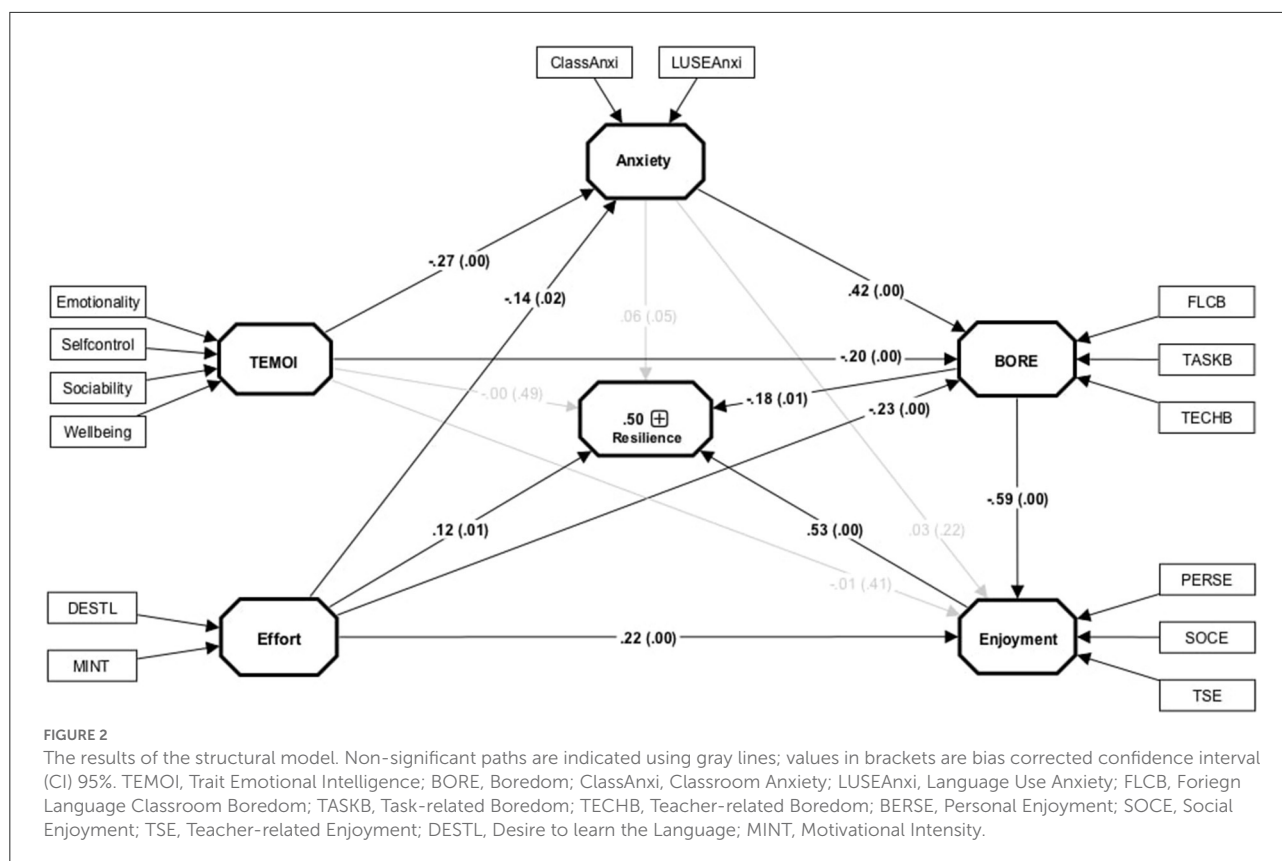
The findings of this study revealed that our model shows adequate applicability and fit the data and accounted for 50% of learner L2 resilience. Thus, the hypothesized model can



be better generalized to language learning contexts. By testing this structural model, it is clear that L2 learner TEI had no significant direct associations with learner positive emotion of enjoyment as well as with the outcome variable in the model (i.e., L2 resilience). These findings run counter to theoretical conclusions made by past research emphasizing that, in the course of language learning, learners' TEI is usually associated with their classroom enjoyment (e.g., Li and Xu, 2019; Li, 2020; Shao et al., 2020) and L2 resilience. It is noteworthy mentioning, nevertheless, that a direct negative relationship between TEI to negative emotions was depicted in this study. Our results, therefore, provide support for the direct negative associations identified by past research (e.g., Dewaele et al., 2008; Dewaele, 2013; Piniel and Csizér, 2013; Shao et al., 2013; Li and Xu, 2019; Li, 2020; Pawlak et al., 2020) between learner character strengths of TEI and negative emotions of anxiety and boredom confirming that learners with higher levels of TEI usually demonstrate lower levels of negative emotions, such as anxiety and boredom, and vice versa. While past research has well-acknowledged the bidimensional role of TEI in both regulating NEs (e.g., Dewaele et al., 2008; Dewaele, 2013; Lake, 2013) and boosting PEs (Dewaele and Mercer, 2018; Li and Xu, 2019; Li, 2020; Resnik and Dewaele, 2020), the data of our model showed

that learner TEI adopt the former rather than the latter role (i.e., regulating learner NEs rather than promoting PEs). This could emphasize the importance of self-control as a significant factor of TEI which considers emotion regulation and alleviating the negative effects of negative emotions a prerequisite step for enhancing the positive side of learner emotions. This has been established by (e.g., Dewaele et al., 2008; Dewaele, 2013; Li and Xu, 2019) who illuminated that TEI is largely about regulating emotions and that developing greater TEI helps learners in primarily regulating their NEs, and in turn, transforming their NEs into PEs, and accordingly improving learning outcomes like achievement and resilience (Oxford, 2016; Li and Xu, 2019).

While no direct relationship between TEI and the outcome variable (i.e., resilience) was established in our model, such a relationship was captured through the mediation of boredom as a negative emotion, the combined role of negative emotions of boredom and anxiety, and the combined role of a mix of negative and positive emotions (e.g., boredom and enjoyment, anxiety, boredom, and enjoyment, etc.). This emphasizes that learner emotions can hold multiple functions in determining the relationship between learner strengths and learner L2 resilience in that they are not only directly associated with their TEI but also facilitate the relationship between learner



TEI and their resilience. Such findings cooperate with those of earlier investigations where negative emotions, namely anxiety, mediated the relationship between EI and outcome variables (e.g., L2 achievement) as in the studies of Shao et al. (2013) and Li (2020). In addition, the coexistence of both negative and positive emotions as co-mediators of the relationship between TEI and L2 resilience in the present study is similar to the mediation model in the study of Li (2020) where negative and positive emotions co-mediated the relationship between TEI and L2 achievement confirming that learner negative and positive emotions could interchangeably interact not only to predict L2 learning outcomes but also to account for the relationship between L2 learner character strengths and such outcomes.

Effort, the second character strength, appeared to play a greater role in predicting L2 resilience than TEI. In this respect, the effort had a positive direct association with L2 resilience. This supports the very few theoretical conclusions available (Kim et al., 2017) about the strong connections between effort and resilience in language learning in the way that learners devoting a larger amount of effort to learning a foreign language usually tend to exhibit higher levels of resilience in learning that language. Besides its direct negative links with L2 resilience, the effort had a significant positive relationship with enjoyment as a positive emotion indicating that learners who expend greater effort in language learning display higher levels of enjoyment

in learning that language as theoretically hypothesized by earlier L2 investigations (see Pekrun and Linnenbrink-Garcia, 2014; MacIntyre et al., 2019; Shao et al., 2020). In addition, negative connections have been detected between effort and the negative emotions of anxiety and boredom in predicting learner resilience. Such negative connections support the theoretical claims in the literature that learners who devote much effort to learning an L2 usually demonstrate lower levels of anxiety (Piniel and Csizér, 2013) and boredom (Pawlak et al., 2020) while learning this language. Moreover, the effort did not only associate with negative and positive emotions in predicting L2 resilience but also significantly indirectly accounted for learner resilience *via* the mediation of learner enjoyment as a positive emotion, boredom as a negative emotion, and the mixed positive and negative effect of these two emotions. This yet again emphasizes the vital mediating role of learner emotions in accounting for the relationship between learner character strengths and L2 learning outcomes. It establishes that, in the course of language learning, learners devoting much effort to L2 learning are expected to show a higher degree of L2 resilience, dependent on those learners experiencing higher EFL and lower boredom in their EFL classes.

As anticipated in the hypothesized model, the negative emotion of boredom had a negative direct negative influence on L2 resilience revealing that the degree of resilience L2 learners

TABLE 2 The results of PLS_{predict} analysis.

Indicators of the outcome (resilience)	RMSE value	
	PLS model	LM model
Happiness	0.76	0.76
Persistence	0.85	0.85
Self-regulation	0.93	0.93

TABLE 3 Indirect and total effects of the variables on the outcome.

Indirect paths	β	p
TEI -> Boredom -> Resilience	0.04	0.03
TEI -> Boredom -> Enjoyment -> Resilience	0.06	0.001
TEI -> Anxiety -> Boredom -> Resilience	0.02	0.04
TEI -> Anxiety -> Boredom -> Enjoyment -> Resilience	0.04	0.001
Effort -> Boredom -> Resilience	0.04	0.02
Effort -> Enjoyment -> Resilience	0.12	<0.001
Effort -> Boredom -> Enjoyment -> Resilience	0.07	<0.001
Boredom -> Enjoyment -> Resilience	-0.31	<0.001
Anxiety -> Boredom -> Resilience	-0.07	0.02
Anxiety -> Boredom -> Enjoyment -> Resilience	-0.13	<0.001
Total effects		
Anxiety	-0.13	0.004
Boredom	-0.50	<0.001
Enjoyment	0.53	<0.001
Effort	0.36	<0.001
TEI	0.13	0.003

The 95% CI are based on bias corrected method.

demonstrate in L2 learning is adversely influenced by the sense of boredom they experience in language class (e.g., Shao et al., 2020). In addition, boredom had also a strong direct negative effect on learner enjoyment recognizing the conclusions made by earlier investigations that the learners' feelings of boredom in language class usually undermine their sense of enjoyment in learning an L2.

Despite that the other negative emotion (i.e., anxiety) had a positive direct influence on boredom confirming what has been hypothesized by past theoretical and empirical research that the feelings of language anxiety are usually coupled with a sense of boredom (Kruk and Zawodniak, 2018; Li and Dewaele, 2020; Pawlak et al., 2020), anxiety did not reach a significant level in directly predicting learner L2 resilience. In this study, the role of anxiety in predicting L2 resilience has not been unique but rather dependent on other learner emotions where it has only indirectly affected L2 resilience through the sole effect of boredom and the combined influence of boredom and enjoyment. Thus, it can be argued that anxiety in this study context might not be necessarily harmful to learner resilience but rather the increasing feelings of boredom and lack of enjoyment, likely resulting from anxiety, are responsible for affecting L2 resilience in this model.

The absence of a direct relationship between learner anxiety and L2 resilience in our model establishing that learner feelings of language anxiety have no direct role to play in their resilience in language learning appears an uncommon result given the vital role of anxiety in affecting different aspects of language learning (see Gregersen and MacIntyre, 2014; Jiang and Li, 2017; Dewaele et al., 2018; Saito et al., 2018; Li et al., 2020) including learner resilience (e.g., Chaffee et al., 2014). It, however, goes in line with what recent investigations such as that of Sparks and Alamer (2022) as well as Alamer and Lee (2021) found with L2 students as they reported that anxiety does not necessarily impact language learning, but rather might be a consequence of poor learning. This conclusion, however, merits further validation by future investigations in other EFL/ESL contexts.

Interestingly, the model in this study showed that L2 resilience has been directly positively influenced by learner enjoyment postulating that the more L2 learners feel enjoying L2 learning, the higher resilient they appear to be in learning this language (e.g., Frederickson et al., 2003; MacIntyre and Gregersen, 2012; Dewaele and MacIntyre, 2014; Oxford, 2016; Dewaele et al., 2018; MacIntyre et al., 2019; Shao et al., 2020). In addition to its direct influence on L2 resilience, enjoyment appeared to control the relationship between learner boredom and resilience establishing that the way that EFL learners' boredom affects their L2 resilience is determined by the degree of enjoyment those learners display. Another notable finding in the present study is that learner enjoyment as a positive emotion had the largest total effect on learner resilience as well as the strongest direct influence on this variable. This conclusion once again indicates that learners' sense of enjoyment in language classes plays a major role in accounting for their resilience to learn a foreign language. This matches the findings of a vast body of research that acknowledge the crucial role of enjoyment in language learning in general (see Dewaele and Alfawzan, 2018; Dewaele et al., 2018; Li et al., 2019; MacIntyre et al., 2019; Shao et al., 2020) and in enhancing L2 learner resilience in particular. Importantly, the total predictive power of learner enjoyment of positive emotion was larger than that of boredom as a negative emotion in our model. This emphasizes that learners' L2 positive emotions play a better role in explaining their L2 resilience (see, Gkonou et al., 2017; Kim et al., 2017; Dewaele et al., 2018; Kim and Kim, 2019) than their negative emotions. It also verifies the claims that the role of positive emotions in language learning usually outweighs that of negative emotions (see MacIntyre and Gregersen, 2012; Dewaele et al., 2018, 2019; Alamer and Lee, 2019).

Overall, the findings that emerged out of the original model in this study substantially align with the positive psychological perspectives that recognize the vital role of character strengths and learner emotions in language learning in that such strengths and emotions help enhance L2 learner resilience. Most interestingly, while the total effect of TEI on L2 resilience was significant, the paths in the structural model showed that L2 learner TEI did not significantly explain their resilience

directly but rather completely indirectly through the mediation of learner negative and positive emotions. TEI only directly correlated with the negative emotions of anxiety and boredom and influence L2 resilience through these two emotions besides enjoyment. This suggests that regulating learner emotions could be a precondition for learner TEI to have a role in their L2 resilience. Effort, the other character strength, showed a stronger role in predicting L2 resilience than TEI in that it, directly and indirectly, predicted learner resilience and, consequently, had a larger total effect on this variable than TEI; and also significantly directly correlated with learner both positive and negative emotions. Learner emotions of enjoyment and boredom scored the largest total effect on learner resilience in the model because they not only directly predicted L2 resilience, but also mediated the relationship between learner TEI and effort character strengths, the negative emotion of anxiety, and L2 resilience. This multidimensional role of enjoyment and boredom recognizes their crucial role in explaining learner resilience in language learning which deserves to attract the attention of future research in the L2 domain.

Implications

It is undeniable that learner resilience plays an important role in the foreign language learning process. However, for this concept to properly operate in such a process, it should be accompanied by effective character strengths and positive emotions. This study aimed to test the complex relationship among a variety of variables representing language learner character strengths and emotions and to unveil the role of these variables in learner resilience for foreign language learning. The primary conclusion this study came up with acknowledged the positive association between learner character strengths and learner emotions and emphasized the bi-dimensional very significant role of these variables in language learning in that they help in enhancing L2 learner resilience.

The findings of this study suggest a number of pedagogical implications. Because learners who are emotionally engaged in language learning are usually less likely to experience negative emotions, L2 educators need to build on learners' character strengths and positive emotions to ease the negative consequences of negative emotions and in turn enhance their resilience in language learning. In this respect, L2 teachers can also get the benefit of deploying EI intervention programs such as the one utilized by [Li and Xu \(2019\)](#) with their EFL learners to improve their emotional experiences of those learners. Furthermore, deploying positive emotions interventions to promote PEs (with greater emphasis on enhancing learner enjoyment) might be very useful in controlling negative emotions and endorsing resilience and persistence in language learning as established by earlier research ([MacIntyre and Gregersen, 2012](#); [Dewaele et al., 2018, 2019](#); [Kim and Kim, 2019](#)). Besides promoting learners' positive emotions, L2 teachers

should take whatever is necessary to regulate negative emotions, such as boredom and anxiety. One useful example in this regard is what [MacIntyre \(2016\)](#) recommended that teachers and learners can complement anxiety-reduction strategies with applications of character strengths, such as courage to undo the negative effects of language anxiety.

Conclusion and limitations

The present study aims to model the relationships between learner character strengths (including TEI and effort) and classroom emotions (enjoyment, anxiety, and boredom) in language resilience. We applied SEM to test the hypothesized model and the results provided us with a greater understanding of how these variables are serially connected and lead to resilience. Nonetheless, a number of limitations must be presented. First, the present study applied the hypothesized model to one socio-educational context (i.e., Saudi learners of English). Thus, although the results can be generalized to learners with similar characteristics, they cannot be assumed to be the same across distinct socio-cultural contexts. It would be plausible to replicate the findings of the present study with other types of L2 learners to see the similarities and differences. Moreover, the present article employed only a cross-sectional survey design. Thus, any claims about cause-and-effect relationships among the predictors and outcomes were not assumed. The structural model was limited in that it depicts only how the variables are linked to each other. We suggest further research to assess the confirmed effects from a longitudinal perspective to reach a better, yet not full, understanding of causation. Lastly, the present study was based on self-report data, gathered from an online questionnaire. The respondents found the length of the survey a bit too long (about 100 items altogether), which might lead to inaccurate responses which is an inevitable challenge to the "truthfulness" in the responses. Overall, it is our hope that the information presented in this research can be disseminated to language education and positive psychology communities and be known to teachers and educators to be used to further the learning and teaching practices.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author/s.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation

and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Both authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alamer, A. (2021). Grit and language learning: construct validation of L2-Grit scale and its relation to later vocabulary knowledge. *Educ. Psychol.* 41, 544–562. doi: 10.1080/01443410.2020.1867076
- Alamer, A. (2022a). Basic psychological needs, motivational orientations, effort, and vocabulary knowledge: A comprehensive model. *Stud. Second Lang. Acquisit.* 44, 164–184. doi: 10.1017/S027226312100005X
- Alamer, A. (2022b). Having a single language interest autonomously predicts L2 achievement: Addressing the predictive validity of L2 grit. *System* 108:102850. doi: 10.1016/j.system.2022.102850
- Alamer, A., and Lee, J. (2019). A motivational process model explaining L2 Saudi students' achievement of English. *System* 87:102133. doi: 10.1016/j.system.2019.102133
- Alamer, A., and Lee, J. (2021). Language achievement predicts anxiety and not the other way around: A cross-lagged panel analysis approach. *Lang. Teach. Res.* doi: 10.1177/13621688211033694
- Bielak, J., and Mystkowska-Wiertelak, A. (2020). Language teachers' interpersonal learner-directed emotion-regulation strategies. *Lang. Teach. Res.* 26, 1082–1105. doi: 10.4324/9781003019497-4
- Chaffee, K. E., Noels, K. A., and McEown, M. S. (2014). Learning from authoritarian teachers: Controlling the situation or controlling yourself can sustain motivation. *Studi. Second Lang. Learn. Teach.* 4, 355–387. doi: 10.14746/sslnt.2014.4.2.9
- Chen, X., and Padilla, A. M. (2019). Role of bilingualism and biculturalism as assets in positive psychology: Conceptual dynamic gear model. *Front. Psychol.* 10:2122. doi: 10.3389/fpsyg.2019.02122
- Chow, B. W., Chiu, H. T., and Wong, S. W. (2018). Anxiety in reading and listening English as a foreign language in Chinese undergraduate students. *Lang. Teach. Res.* 22, 719–738. doi: 10.1177/1362168817702159
- Connor, K. M., and Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depress. Anxiety* 18, 76–82. doi: 10.1002/da.10113
- Dewaele, J. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *J. Lang. Soc. Psychol.* 38, 523–535. doi: 10.1177/0261927X19864996
- Dewaele, J., and Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Stud. Second Lang. Learn. Teach.* 8, 21–45. doi: 10.14746/sslnt.2018.8.1.2
- Dewaele, J., Chen, X., Padilla, A. M., and Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Front. Psychol.* 10:2128. doi: 10.3389/fpsyg.2019.02128
- Dewaele, J., and MacIntyre, P. (2019). *The Predictive Power of Multicultural Personality Traits, Learner and Teacher Variables on Foreign Language Enjoyment and Anxiety (1st ed.)*. Routledge. doi: 10.4324/9781351190558-12
- Dewaele, J.-M. (2013). “Emotions and language learning,” in: *Routledge Encyclopedia of Language Teaching and Learning*, eds M. Byram and A. Hu (London: Routledge).
- Dewaele, J.-M., and MacIntyre, P. D. (2016). “Foreign language enjoyment and anxiety: the right and left feet of the language learner,” in *Positive Psychology in SLA*, eds T. Gregersen, P. D. MacIntyre, and S. Mercer, (Bristol: Multilingual Matters). doi: 10.21832/9781783095360-010
- Dewaele, J.-M., and Mercer, S. (2018). “Do ESL/EFL teachers' emotional intelligence, teaching experience, proficiency and gender, affect their classroom practice?” in: *Language Teacher Psychology*, eds S. Mercer and A. Kostoulas (Bristol: Multilingual Matters). doi: 10.21832/9781783099467-015
- Dewaele, J.-M., Witney, J., Saito, K., and Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Lang. Teach. Res.* 22, 676–697. doi: 10.1177/1362168817692161
- Dewaele, J. M., and Dewaele, L. (2017). The dynamic interactions in foreign language classroom anxiety and foreign language enjoyment of pupils aged 12 to 18. A pseudo-longitudinal investigation. *J. Eur. Second Lang. Assoc.* 1, 12–22. doi: 10.22599/jesla.6
- Dewaele, J. M., and Li, C. (2022). foreign language enjoyment and anxiety: associations with general and domain-specific english achievement. *Chin. J. Appl. Linguist.* 45, 32–48. doi: 10.1515/CJAL-2022-0104
- Dewaele, J. M., and MacIntyre, P. D. (2014). “The Two Faces of Janus? Anxiety and enjoyment in the foreign language classroom.” *Stud. Second Lang. Learn. Teach.* 4, 237–274. doi: 10.14746/sslnt.2014.4.2.5
- Dewaele, J. M., Petrides, K. V., and Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Lang. Learn.* 58, 911–960. doi: 10.1111/j.1467-9922.2008.00482.x
- Elahi, S., and Taherian, T., (2021). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: latent growth curve modeling. *Int. J. Bilingual Educ. Bilingual.* 24, 31–49. doi: 10.1080/13670050.2018.1441804
- Fiorilli, C., Farina, E., Buonomo, I., Costa, S., Romano, L., Larcán, R., et al. (2020). Trait emotional intelligence and school burnout: the mediating role of resilience and academic anxiety in high school. *Int. J. Environ. Res. Public Health* 17:3058. doi: 10.3390/ijerph17093058

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.956216/full#supplementary-material>

- Frederickson, B. L., Tugade, M. M., Waugh, C. E., and Larkin, G. (2003). What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11, 2001. *J. Personal. Soc. Psychol.* 84, 365–376. doi: 10.1037/0022-3514.84.2.365
- Frederickson, B. (2009). *Positivity*. New York, NY: Harmony.
- Frederickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *Am. Psychol.* 56, 218–226. doi: 10.1037/0003-066X.56.3.218
- Gardner, R. C. (2010). *Motivation and Second Language Acquisition: The Socio-Educational Model*. New York, NY: Peter Lang.
- Gkonou, C., Daubney, M., and Dewaele, J. M. (eds.). (2017). *New Insights into Language Anxiety: Theory, Research and Educational Implications*. Multilingual Matters.
- Goleman, D. (2012). *Emotional Intelligence: Why It Can Matter More Than IQ*. New York, NY: Bantam Books.
- Gregersen, T., and MacIntyre, P. D. (2014). *Capitalizing on Language Learners' Individuality*. Bristol: Multilingual Matters. doi: 10.21832/9781783091218
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., and Chechile, R. A. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: SAGE Publications.
- Hair, J., and Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Res. Methods Appl. Linguist.* 1:e100027. doi: 10.1016/j.rmal.2022.100027
- Hair, J., Sarstedt, M., and Ringle, C. (2019). Rethinking some of the rethinking of partial least squares. *Eur. J. Mark.* 53, 566–584. doi: 10.1108/EJM-10-2018-0665
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., and Wu, J. (2021). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Lang. Teach. Res.* 2021:13621688211001289. doi: 10.1177/13621688211001289
- Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quart.* 20, 559–562. doi: 10.2307/3586302
- Jiang, G., and Li, C. (2017). SLA research in the positive psychology perspective: Review and prospects. *Foreign Lang. World* 5, 32–39. Available online at: https://caod.oriprobe.com/articles/52514696/SLA_research_in_the_Positive_Psychology_perspectiv.htm
- Jiang, Y., and Dewaele, J. M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners. *System* 82, 13–25. doi: 10.1016/j.system.2019.02.017
- Jin, Y., and Zhang, L. J. (2018). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. *Int. J. Biling. Educ. Biling.* 24, 948–962. doi: 10.1080/13670050.2018.1526253
- Karabiyik, C., and Mirici, I. H. (2018). Development and validation of the foreign language learning effort scale for Turkish tertiary-level students. *Educ. Sci.* 18, 373–395.
- Khajavy, G. H., MacIntyre, P. D., and Barabadi, E. (2018). “Role of the emotions and classroom environment in willingness to communicate: applying doubly latent multilevel analysis in second language acquisition research.” *Stud. Second Lang. Acquisition* 40, 605–624. doi: 10.1017/S0272263117000304
- Kim, T., and Kim, Y. (2017). The impact of resilience on L2 learners' motivated behaviour and proficiency in L2 learning. *Educ. Stud.* 43, 1–15. doi: 10.1080/03055698.2016.1237866
- Kim, T., Kim, Y., and Kim, J. (2019). Role of resilience in (de) motivation and second language proficiency: Cases of Korean elementary school students. *J. Psycholinguist. Res.* 48, 371–389. doi: 10.1007/s10936-018-9609-0
- Kim, T. Y., Kim, Y., and Kim, J. Y. (2017). Structural relationship between L2 learning (de) motivation, resilience, and L2 proficiency among Korean college students. *Asia-Pac. Educ. Res.* 26, 397–406.
- Kruk, M., and Zawodniak, J. (2018). “Boredom in practical English language classes: Insights from interview data,” in: *Interdisciplinary Views on the English Language, Literature and Culture*, eds L. Szymański, J. Zawodniak, A. Łobodzić, and M. Smoluk (Zielona Góra: Uniwersytet Zielonogórski).
- Lake, J. (2013). “Positive L2 self: Linking positive psychology with L2 motivation,” in: *Language Learning Motivation in Japan*, eds M. T. Apple, D. Da Silva, and T. Fellner (Bristol: Multilingual Matters). doi: 10.21832/9781783090518-015
- Li, C. (2020). A positive psychology perspective on Chinese EFL students' trait emotional intelligence, foreign language enjoyment and EFL learning achievement. *J. Multiling. Multicult. Dev.* 41, 246–263. doi: 10.1080/01434632.2019.1614187
- Li, C., Dewaele, J., and Hu, Y. (2021). Foreign language learning boredom: Conceptualization and measurement. *Appl. Linguist. Rev.* doi: 10.1515/applirev-2020-0124
- Li, C., Dewaele, J., and Jiang, G. (2020). The complex relationship between classroom emotions and EFL achievement in China. *Appl. Linguist. Rev.* 11, 485–510.
- Li, C., and Dewaele, J. M. (2020). The predictive effects of trait emotional intelligence and online learning achievement perceptions on foreign language class boredom among Chinese university students. *Foreign Lang. Foreign Lang. Teach.* 5, 33–44.
- Li, C., Dewaele, J. M., and Jiang, G. (2019). The complex relationship between classroom emotions and EFL achievement in China. *Appl. Linguist. Rev.* 11, 485–510. doi: 10.1515/applirev-2018-0043
- Li, C., Jiang, G., and Dewaele, J. (2018). Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale. *System* 76, 183–196. doi: 10.1016/j.system.2018.06.004
- Li, C., and Wei, L. (2022). Anxiety, enjoyment, and boredom in language learning amongst junior secondary students in rural China: How do they contribute to L2 achievement? *Stud. Second Lang. Acquisit.* 2022, 1–16. doi: 10.1017/S0272263122000031
- Li, C., and Xu, J. (2019). Trait emotional intelligence and classroom emotions: A positive psychology investigation and intervention among Chinese EFL learners. *Front. Psychol.* 10:2453. doi: 10.3389/fpsyg.2019.02453
- MacIntyre, P., and Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Stud. Second Lang. Learn. Teach.* 2, 193–213. doi: 10.14746/ssl.2012.2.2.4
- MacIntyre, P. D. (2016). *So Far So Good: An Overview of Positive Psychology and Its Contributions to SLA. Positive Psychology Perspectives on Foreign Language Learning and Teaching*. Cham: Springer. doi: 10.1007/978-3-319-32954-3_1
- MacIntyre, P. D., Gregersen, T., and Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *Modern Lang. J.* 103, 262–274. doi: 10.1111/modl.12544
- Moskowitz, S., and Dewaele, J. (2020). Through the looking glass of student perception: How foreign language students see teacher trait emotional intelligence and why it matters. *Stud. Second Lang. Learn. Teach.* 10, 239–256. doi: 10.14746/ssl.2020.10.2.2
- Oxford, R., and Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *Modern Lang. J.* 78, 12–28. doi: 10.1111/j.1540-4781.1994.tb02011.x
- Oxford, R. L. (2016). “Toward a psychology of well-being for language learners: The ‘EMPATHICS’ Vision,” in: *Positive Psychology in SLA*, eds P. MacIntyre, T. Gregersen, and S. Mercer (Bristol: Multilingual Matters). doi: 10.21832/9781783095360-003
- Park, N., Peterson, C., and Seligman, M. E. (2004). Strengths of character and well-being. *J. Soc. Clin. Psychol.* 23, 603–619. doi: 10.1521/jsc.23.5.603.50748
- Pawlak, M., Kruk, M., Zawodniak, J., and Pasikowski, S. (2020). Investigating factors responsible for boredom in English classes: The case of advanced learners. *System* 91:102259. doi: 10.1016/j.system.2020.102259
- Pekrun, R., and Linnenbrink-Garcia, L. (2014). *Introduction to Emotions in Education*. New York, NY: Routledge.
- Petrides, K. V. (2009). “Psychometric properties of the trait emotional intelligence questionnaire (TEIQue),” in: *Assessing Emotional Intelligence*, eds J. D. A. Parker, D. H. Saklofske, and C. Stough (Cham: Springer). doi: 10.1007/978-0-387-88370-0_5
- Petrides, K. V., and Furnham, A. (2003). Trait emotional intelligence: Behavioral validation in two studies of emotion recognition and reactivity to mood induction. *Eur. J. Pers.* 17, 39–57. doi: 10.1002/per.466
- Petrides, K. V., Pita, R., and Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *Br. J. Psychol.* 98, 273–289. doi: 10.1348/000712606X120618
- Piniel, K., and Csizér, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Stud. Second Lang. Learn. Teach.* 3, 523–550. doi: 10.14746/ssl.2013.3.4.5
- Resnik, P., and Dewaele, J. M. (2020). Trait emotional intelligence, positive and negative emotions in first and foreign language classes: A mixed-methods approach. *System* 94:102324. doi: 10.1016/j.system.2020.102324
- Saito, K., Dewaele, J., Abe, M., and In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in

classroom settings: A cross-sectional and longitudinal study. *Lang. Learn.* 68, 709–743. doi: 10.1111/lang.12297

Salovey, P., Mayer, J. D., Caruso, D., and Yoo, S. H. (2009). “The positive psychology of emotional intelligence,” in: *The Oxford Handbook of Positive Psychology*, eds S. J. Lopez and C. R. Snyder (New York, NY: Oxford University Press). doi: 10.1093/oxfordhb/9780195187243.013.0022

Seligman, M. E., and Csikszentmihalyi, M. (2014). “Positive psychology: An introduction,” in: *Flow and the Foundations of Positive Psychology*, ed M. Csikszentmihalyi (Cham: Springer). doi: 10.1007/978-94-017-9088-8_18

Seligman, M. E. P., and Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *Am. Psychol.* 55, 5–14. doi: 10.1037/0003-066X.55.1.5

Shao, K., Nicholson, L. J., Kutuk, G., and Lei, F. (2020). Emotions and instructed language learning: Proposing a second language emotions and positive psychology model. *Front. Psychol.* 11:2142. doi: 10.3389/fpsyg.2020.02142

Shao, K., Yu, W., and Ji, Z. (2013). An exploration of Chinese EFL students' emotional intelligence and foreign language anxiety. *Mod. Lang. J.* 97, 917–929. Available online at: <http://www.jstor.org/stable/43651729>

Shin, W., Kim, M., and Kim, J. (2009). Developing measures of resilience for Korean adolescents and testing cross, convergent, and discriminant validity. *Stud. Korean Youth* 20, 105–131.

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., and Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *Eur. J. Mark.* 53, 2322–2347. doi: 10.1108/EJM-02-2019-0189

Sparks, R., and Alamer, A. (2022). Long-term impacts of L1 language skills on L2 anxiety: The mediating role of language aptitude and L2 achievement. *Lang. Teach. Res.* doi: 10.1177/13621688221104392

Trigueros, R., Padilla, A. M., Aguilar-Parra, J. M., Rocamora, P., Morales-Gázquez, M. J., and López-Liria, R. (2020). The influence of emotional intelligence on resilience, test anxiety, academic stress and the Mediterranean diet. A study with university students. *Int. J. Environ. Res. Public Health* 17:2071. doi: 10.3390/ijerph17062071



OPEN ACCESS

EDITED BY
Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY
Juan Andres Merino,
Universidad Internacional De La Rioja,
Spain
Verónica Alcaraz-Muñoz,
Catholic University San Antonio
of Murcia, Spain

*CORRESPONDENCE
Alberto Gómez-Mármol
alberto.gomez1@um.es

SPECIALTY SECTION
This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

RECEIVED 26 September 2022
ACCEPTED 29 September 2022
PUBLISHED 21 October 2022

CITATION
Clares-Clares E and Gómez-Mármol A
(2022) Influence of performance
in Spanish language and literature on
physical education and music grades.
Front. Educ. 7:1054024.
doi: 10.3389/feduc.2022.1054024

COPYRIGHT
© 2022 Clares-Clares and
Gómez-Mármol. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Influence of performance in Spanish language and literature on physical education and music grades

Esperanza Clares-Clares and Alberto Gómez-Mármol*

Faculty of Education, University of Murcia, Murcia, Spain

The aim of this research was to analyze the influence of performance in the subject of Spanish language and literature on the grades in the subjects of Physical Education and Music and whether there were differences by sex or not. The participating sample was made up by a total of 451 students (235 boys and 216 girls), aged between 10 and 13 years old, in the fifth and sixth grades of Primary Education. The results show, on the one hand, a positive correlation between the grades of the three subjects and, on the other hand, in terms of gender, significant differences only in the subject of Music whose grades are higher in girls compared to boys. Nevertheless, these statements should be considered taking into account that academic performance was measured only with one instrument and that the sample, besides being limited in size, was accessed with a non-probabilistic method. Based on these results, it is suggested that future studies could encompass different research methodologies and include, in turn, new variables as well as other collectives which are also concerned with this regard.

KEYWORDS

academic performance, qualification, reading skills, writing skills, student achievements

Introduction

Academic performance in the school context is a research topic of growing interest for groups such as: teachers, families, pedagogues, psychologists, sociologists (Bacon and Lord, 2021; Duric et al., 2021; Estrada et al., 2021; Chen et al., 2022). In this sense, examples of educational policies that invest in ensuring academic performance can be found in the PISA program (Programme for International Students Assessment) or in TIMSS (Trends in International Mathematics and Science Study), which periodically analyze the results obtained in educational centers at an international level. In this regard, Graus (2021) recently noted a downward trend in terms of academic performance.

According to Cupani and Lorenzo (2010), academic performance is a theoretical construct about which it is difficult to reach a universally accepted definition. Among the different definitions that have been made about academic performance, those of Álvarez and Albuérne (2001), Alcalde (2009), and Campaña Preciado (2019) could be highlighted. While in the first, authors understand academic performance as the set of educational actions that the educational system develop to value one student achievements, the second (Alcalde, 2009) defines it as the process to achieve the planned objectives and that can be registered through the execution of evaluation activities. More recently, Campaña Preciado (2019) defined it as the aptitudes that child show according to what they have learned in the different curricular areas, which will help them develop those required abilities and skills to overcome their personal challenges.

In more detail, there are many recent examples of previous research which have analyzed academic performance. However, López Guillamón and Clares-Clares (2019) remark that this research trend focuses on subjects such as Spanish language or Mathematics and, subsequently, studies about Physical Education or Music are still scarce. In this research, the focus is set in the combination of Spanish language, Music and Physical Education. Other examples of research in these disciplines are: in Mathematics and Spanish language (Santillán and Vargas, 2022), Physical Education (García-Hermoso et al., 2021) and Music (Barrientos et al., 2019). As the most notable skills linked to performance in the Language and Literature subject, Chóez (2022) pointed out reading and writing habits, as for Music, León et al. (2015) found that a greater development of critical thinking skills and intrinsic motivation contributes to higher grades in Music subject. As far as Physical Education is concerned, Prada (2022) demonstrated the positive influence of the practice of physical-sports activity on its grades. Other research pointed out how important motivation, satisfaction of psychological needs, responsibility and school social climate are so as to avoid studies' drop out, which is linked to low grades (Manzano-Sánchez et al., 2021).

The student's gender is also a variable that has been previously linked by other research with academic performance (Meinck and Brese, 2019). However, it should be noted that contradictory results have been found in this regard that do not allow rigorous conclusions to be drawn. For example, while Hyde (2014) suggests that the differences are not based on sex but on the cultural gender influence, Echavarri et al. (2007) did find higher innate verbal skills in girls and higher mathematical skills in boys. In this sense, research that has given a more holistic treatment to academic performance, analyzing these three disciplines at the same time, is still scarce and, even less frequent, if the aim is

to know the influence of performance on one over the other.

Thus, the objective of this preliminary study was, on the one hand, to test whether there is a correlation between the grades obtained in the subject of Spanish Language and Literature and those of Physical Education and Music in Primary Education and, on the other, to analyze the possible differences according to the gender.

Materials and methods

Participants

A total of 545 students participated in this research although, after eliminating the cases of those who did not provide the grades of the three subjects, the resulting sample was 451 subjects (235 boys and 216 girls) aged between 10 and 13 years (11.49 ± 0.59) from the fifth and sixth grades of Primary Education in eight educational centers in the Region of Murcia, in Spain. Access to the sample was carried out through non-probabilistic incidental sampling (Latorre et al., 2003, p. 82) "which is based on the selection of subjects to whom the researcher has access." Participation was voluntary, with no link to the grade in any subject, and anonymity was ensured at all times in the statistical treatment of the information collected.

Instruments

This is a cross-sectional, descriptive and *ex post-facto* type of research (Thomas et al., 2022) with the following study variables: gender, grades in the subjects of Spanish Language and Literature, Music and Physical Education.

Procedures

To carry out this study, firstly, its design was assessed by Ethical Committee of the University of Murcia, which approved it. Secondly, the participation of the educational centers was requested through a letter addressed to the directors. It informed the purpose of the investigation. Once consent was obtained, those tutors who voluntarily wanted to participate sent the requested information. Participants did not receive any reward.

Data analysis

For the statistical treatment of the data, the SPSS 24.0 program was used. So as to analyze the normality of the

TABLE 1 Grades in the three subjects for the whole sample and according to gender.

Spanish language and literature grades		<i>p</i>	Music grades	<i>p</i>	Physical education grades	<i>p</i>	
Sex	Boys	7.05 ± 1.97	0.231	8.59 ± 0.84	0.021	7.16 ± 1.56	0.445
	Girls	6.75 ± 2.21		8.74 ± 1.01		7.25 ± 1.63	
Total		6.90 ± 2.09		8.66 ± 0.93		7.20 ± 1.59	

TABLE 2 Correlations between the grades of the three subjects.

Subject	Music	Physical education
Spanish language and literature	0.611**	0.762**
Music	–	0.639**

**Correlation is significant at $p \leq 0.01$.

data distribution, the test used was Kolmogorov Smirnov which reported a non-parametric distribution. To calculate the differences by gender, the Mann Whitney *U* test was used, while the Spearman rank test was applied for the correlation between the grades of the different subjects.

Results

Next, Table 1 shows the descriptive statistics of the grades in the three subjects, for the total sample and distinguishing by gender, as well as the statistical significance of the differences.

In general, it is observed that the highest grade corresponds to the subject of music while, in terms of gender, statistically significant differences are only found in this same subject, with higher values among women compared to men.

On the other hand, in order to analyze the correlations between the qualifications, the results collected in Table 2 are shown:

As can be seen from the table above, there is a positive correlation between the grades of the three subjects, in such a way that a higher grade in any one is associated with higher grades in the other two. In the opposite direction, the lowest grades are also associated between the different subjects.

Discussion

In this research, the influence of performance in the subject of Spanish language and literature on the grades in the subjects of Physical Education and Music has been studied and it has been analyzed if there are differences by sex.

The results obtained in this study provide evidence to affirm, in the first place, that the grades in the subject of Music are higher than those of Spanish Language and Physical Education, in line with those reported by Oliver (2015) and Benítez Roca and Gómez-Mármol (2019). It should be noted

in this regard that various studies have already shown the positive influence of music and the cognitive skills on the academic performance of schoolchildren. Chong (2006) has pointed out that structured musical activities can facilitate basic academic skills and motivate participation in tasks in low-achieving students, while Jaschke et al. (2018) have shown a long-term positive influence of music education in cognitive abilities such as inhibition and planning. These authors have also pointed out that the research on the effects of music interventions on cognitive abilities, while promising, still needs more randomized longitudinal studies to support a positive claim, especially as the results of some studies imply opposite conclusions. Sala and Gobet (2020), for example, concluded that researchers' optimism about the benefits of music training is unwarranted and possibly stems from a misinterpretation of empirical data. In addition, the idea of the positive influence of creativity (which is inherent to music according to Serrano, 2009) in the teaching-learning process, as well as its subsequent impact on the academic performance of students, seems to have already been confirmed (Alonso et al., 2015; Bernabeu and De la Peña, 2021; López-Martínez and Lorca, 2021; Manetti et al., 2022). On the other hand, Martín López and León del Barco (2009) showed that there are positive correlations between the music grade and the understanding of emotional states, on the one hand, and, on the other, that students with low musical performance, that is, those who failed in music, also obtained the lowest scores in the expression, understanding and regulation of their emotions. Other investigations have also confirmed a positive trend or ergogenic effect of music on physical performance, both in aerobic and anaerobic activities (Gallardo-Ríos and Conesa, 2016). Moreover, it also seems to be evident that there is a positive relationship between academic performance and physical activity (Prada, 2022).

Secondly, the data analyzed in this research show that girls have higher grades in Music than boys, although, on the contrary, neither in Spanish Language nor in Physical Education are there statistically significant differences according to gender. This idea seems to be in tune with the data obtained in other studies. Cárcamo et al. (2020) have shown that girls perform better in language than boys, and that this higher performance is modulated by self-concept, performance expectations and age. However, other research supports the hypothesis of gender similarities and that the differences can be

explained socially and culturally and based on historical roles (Hyde, 2014).

Thirdly, a positive correlation has been found between the grades of the three subjects studied. In this regard, other research has found that good results in the area of music are relatively related to good results in the school environment in general, and vice versa (Serrano, 2005) with the exception of Physical and Plastic Education. However, it is necessary to reflect on the influence of receiving extracurricular classes, since it may happen that their taste for music is combined with socio-family and economic circumstances that are more favorable than the average, with a concern for their parents to participate in other complementary training activities.

With regard to limitations, a few must be considered for this study results' interpretation. First, possible external validity restrictions inherent to non-probability sampling and cross-sectional design. In this regard, the extrapolation of the data to the study population in general is compromised by the size of the sample and by sticking to a single sociocultural setting. Second, the use of a single indicator to measure a variable such as academic grades that, moreover, do not show all the knowledge included in the educational curriculum. In addition, the exclusive positioning in the positivist paradigm, excluding causal explanations that could emerge from studies that included qualitative methodologies such as interviews and discussion groups in which not only students but also their families, teachers and school-management teams were involved. Besides, the data collection included the results of the first trimester while the marks at the end of the grade could have been considered more representative of the academic performance of the whole academic year.

Finally, from the results achieved, it can be concluded that there is a positive correlation between performance in the subject of Spanish Language and Literature and grades in Physical Education and Music. Although, to the best of our knowledge, this study is a pioneer in the joint analysis of these three subjects, other studies confirm the two-to-two correlations between their grades. In addition, with regard to gender, a higher academic performance in the subject of Music by girls has been reported, with no differences in Spanish Language and Literature or Physical Education. In this sense, the literature does not allow a globally accepted position to be reached, as contradictory results have been concluded in this regard.

In this way, based on the above-mentioned, the possibility of carrying out new works that complement this scope opens up. Thus, future studies could analyze performance not only from the final grades but also from the results in specific tests linked to each subject or skills or, from a more holistic conception, involve teachers in research

in order to, from their opinion, determine if there are potential student profiles that support the existence of students with higher performance in all subjects and others with lower performance, equally, in all subjects. Another field of knowledge to keep on studying in is the potential relationship between the type of parental socialization and academic performance. Finally, to promote action-research, the design of intervention strategies that allow increasing academic performance, starting from the improvement of linguistic, musical or motor skills, is a field that still requires many more investigations.

Data availability statement

The datasets presented in this article are not readily available because this data must remain confidential. Requests to access the datasets should be directed to corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Universidad de Murcia. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

Both authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alcalde, R. (2009). *Salir de pobres: Monomarentalidad, infancia e inmigración latinoamericana en España. Tesis doctoral*. Bellaterra: Universidad Autónoma de Barcelona.
- Alonso, R. K., Martínez-Monteaudo, M. C., and Martín-Lobo, P. (2015). Creatividad, atención y rendimiento académico en alumnado de conservatorio profesional. *Int. J. Dev. Educ. Psychol.* 1, 473–482. doi: 10.17060/ijodaep.2015.n1.v1.32
- Álvarez, C., and Albuérne, F. (2001). Rendimiento académico y estilos de aprendizaje en alumnos de segundo de bachillerato. *Aula Abierta* 77, 77–85.
- Bacon, P., and Lord, R. N. (2021). The impact of physically active learning during the school day on children's physical activity levels, time on task and learning behaviours and academic outcomes. *Health Educ. Res.* 36, 362–373. doi: 10.1093/her/cyab020
- Barrientos, A., Sánchez-Cabrero, R., and Arigita, A. (2019). Rendimiento académico, musical e inteligencias múltiples en alumnos preadolescentes de un centro específico de música. *Eur. J. Child Dev. Educ. Psychopathol.* 7, 119–132. doi: 10.30552/ejpad.v7i2.113
- Benítez Roca, M. C., and Gómez-Mármol, A. (2019). ¿Influye el estilo parental en el rendimiento académico de alumnado de educación primaria? *Trances* 11, 479–490.
- Bernabeu, E., and De la Peña, C. (2021). Creatividad en educación superior: Estudio exploratorio con función ejecutiva y rendimiento académico. *Profesorado* 25, 313–330. doi: 10.30827/profesorado.v25i3.9546
- Campaña Preciado, F. (2019). *El estrés asociado al rendimiento académico en los niños de II ciclo de educación inicial. Tesis doctoral*. Tumbes: Universidad Nacional de Tumbes.
- Cárcamo, C., Moreno, A., and del Barrio, C. (2020). Diferencias de género en matemáticas y lengua: Rendimiento académico, autoconcepto y expectativas. *Suma Psicol.* 27, 27–34. doi: 10.14349/sumapsi.2020.v27.n1.4
- Chen, C. C., Cheng, S. L., Yaoying, X., Rudasill, K., Senter, R., Zhang, F., et al. (2022). Transactions between problem behaviors and academic performance in early childhood. *Int. J. Environ. Res. Public Health* 19:583. doi: 10.3390/ijerph19159583
- Chóez, R. M. (2022). *La lectoescritura y su incidencia en el rendimiento escolar de los alumnos. Tesis doctoral no publicada*. Ecuador: Universidad estatal del sur de Manabí.
- Chong, H. Y. (2006). Music cognitive skills and academic achievements in mathematics and science. *J. Learn. Cent. Curric. Instr.* 6, 381–397.
- Cupani, M., and Lorenzo, J. (2010). Evaluación de un modelo social-cognitivo del rendimiento en matemática en una población de preadolescentes argentinos. *Infancia y Aprendiz.* 33, 63–74. doi: 10.1174/021037010790317216
- Duric, S., Bogataj, S., Zovko, V., and Sember, V. (2021). Associations between physical fitness, objectively measured physical activity and academic performance. *Front. Public Health* 9:778837. doi: 10.3389/fpubh.2021.778837
- Echavarrí, M., Godoy, J., and Olaz, F. (2007). Diferencias de género en habilidades cognitivas y rendimiento académico en estudiantes universitarios. *Univ. Psychol.* 6, 319–329.
- Estrada, M., Monferrer, D., Rodríguez, A., and Moliner, M. A. (2021). Does emotional intelligence influence academic performance? The role of compassion and engagement in education for sustainable development. *Sustainability* 13:1721. doi: 10.3390/su13041721
- Gallardo-Ríos, L., and Conesa, E. (2016). “Efectos de la música sobre el rendimiento físico. Una revisión,” in *Actas IX congreso internacional de la asociación española de ciencias del deporte*, eds J. M. García García, B. Calvo Rico, and R. Mora Rodríguez (Ciudad Real: Universidad de Castilla La Mancha), 289.
- García-Hermoso, A., Ramírez-Vélez, R., Revalds, D., and Izquierdo, M. (2021). Effects of physical education interventions on cognition and academic performance outcomes in children and adolescents: A systematic review and meta-analysis. *Br. J. Sports Med.* 55, 1224–1232. doi: 10.1136/bjsports-2021-104112
- Graus, J. M. (2021). Propuesta de intervención psicopedagógica para estudiantes que presentan un bajo rendimiento escolar en un Colegio Bilingüe en Chile. *Rev. Estud. Ter. Occup.* 8, 154–188.
- Hyde, J. S. (2014). Gender similarities and differences. *Annu. Rev. Psychol.* 65, 373–398. doi: 10.1146/annurev-psych-010213-115057
- Jaschke, A. C., Honing, H., and Scherder, E. J. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Front. Neurosci.* 12:103. doi: 10.3389/fnins.2018.00103
- Latorre, A., Arnal, J., and Del Rincón, D. (2003). *Bases metodológicas de la investigación educativa*. Barcelona: Hurtado Mompeó.
- León, J., Núñez, J. L., Ruiz-Alfonso, Z., and Bordón, B. (2015). Rendimiento académico en música: Efecto de la motivación intrínseca y el pensamiento crítico. *Rev. de Psicodidáctica* 20, 377–391. doi: 10.1387/RevPsicodidact.12673
- López Guillaumon, A. M., and Clares-Clares, E. (2019). “Rendimiento académico en el aula de música en Educación Primaria,” in *Investigación e innovación educativa en el siglo XXI*, eds A. Gómez-Mármol and C. De Pro (Sevilla: Wanceulen).
- López-Martínez, O., and Lorca, A. J. (2021). Specific Domains of creativity and their relationship with intelligence: A study in primary education. *Sustainability* 13:4228. doi: 10.3390/su13084228
- Manetti, A., Lara Navarra, P., and Sánchez Navarro, J. (2022). El proceso de diseño para la generación de escenarios futuros educativos. *Comunicar* 73, 33–44. doi: 10.3916/C73-2022-03
- Manzano-Sánchez, D., Gómez-Mármol, A., Conte Marín, L., Jiménez-Parra, J. F., and Valero-Valenzuela, A. (2021). Future academic expectations and their relationship with motivation, satisfaction of psychological need, responsibility and School social climate: Gender and educational stage. *Int. J. Environ. Res. Public Health* 18:4558. doi: 10.3390/ijerph18094558
- Martín López, E., and León del Barco, B. (2009). Inteligencia emocional y rendimiento musical. *Campo Abierto* 28, 119–127.
- Meinck, S., and Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large Scale Assess. Educ.* 7:8. doi: 10.1186/s40536-019-0076-3
- Oliver, A. (2015). *Rendimiento académico, sexo y valores interpersonales. Tesis doctoral*. Almería: Universidad de Almería.
- Prada, A. (2022). *Efectos de la práctica de actividad física en el rendimiento académico de estudiantes de secundaria [Tesis de doctorado no publicada]*. Villaviciosa de Odó: Universidad Europea de Madrid.
- Sala, G., and Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Appl. Sci.* 48, 1429–1441. doi: 10.3758/s13421-020-01060-2
- Santillán, A. S., and Vargas, J. R. (2022). Decomposition of the inequality of school performance by the condition of poverty in Mexican students. *Anal. Econ.* 37, 125–141.
- Serrano, L. (2009). Armonía creativa. *Rev. Eufonía* 47, 20–32.
- Serrano, R. M. (2005). Música y desarrollo cognitivo. *Int. J. Dev. Educ. Psychol.* 3, 393–402.
- Thomas, J. R., Martin, P. E., Etnier, J. L., and Silverman, S. J. (2022). *Research methods in physical activity*. Champaign, IL: Human Kinetics.



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga,
Spain

REVIEWED BY

Hui Helen Li,
Wuhan University of Technology,
China
Tribhuvan Kumar,
Prince Sattam Bin Abdulaziz University,
Saudi Arabia

*CORRESPONDENCE

Rafael Crismán-Pérez
rafael.crisman@uca.es

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 14 August 2022

ACCEPTED 07 October 2022

PUBLISHED 31 October 2022

CITATION

Peña-Acuña B and Crismán-Pérez R (2022)
Research on Papua, a digital tool with
artificial intelligence in favor of learning and
linguistic attitudes towards the learning of
the English language in students of Spanish
language as L1.
Front. Psychol. 13:1019278.
doi: 10.3389/fpsyg.2022.1019278

COPYRIGHT

© 2022 Peña-Acuña and Crismán-Pérez.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Research on Papua, a digital tool with artificial intelligence in favor of learning and linguistic attitudes towards the learning of the English language in students of Spanish language as L1

Beatriz Peña-Acuña¹ and Rafael Crismán-Pérez^{2*}

¹Facultad de Educación, Psicología y Ciencias del Deporte, Universidad de Huelva, Huelva, Spain,

²Departamento de Filología, Facultad de Filosofía y Letras, Universidad de Cádiz, Cádiz, Spain

This study examines learners' perceptions of the linguistic and learning potential of an AI-based English language learning app called Papua. The study took a quasi-experimental approach in which 128 students of the degree in Primary Education at the University of Huelva, Spain, gained experience of using the app over a six-week period, and then answered a questionnaire. This was of a quantitative design, and included the following components: attitudes towards learning English; attitudes towards learning oral and written skills, and towards opportunities for interacting in English; and attitudes towards the motivation provided by the app. These attitudes were contrasted according to the variables memory and self-evaluation. The results of the study showed that oral skills were perceived as the most improved skill. Participants also foregrounded the enhancement of memorization of vocabulary, and positively evaluated the self-evaluation feature included in the app.

KEYWORDS

language learning, English language, study and teaching, foreign speakers, digital analog conversion, Spanish language, attitudes, inclusive education

Introduction

Among the range of foreign language learning options, Computer Assisted Language Learning (CALL) has seen good results (Jen-Jiun and Huifen, 2019; Khafaga and Alghawli, 2021). A recent development in this kind of software is the use of apps, which have become increasingly popular as a medium of instruction in general (García-Aretio, 2016), as well as specifically for learning English, with the incorporation of artificial intelligence (Jin et al., 2019). Such programs incorporate a range of approaches, such as task scaffolding and conversational interaction (Chie Fang et al., 2021), and the development of speaking (Moreno-Ibañez et al., 2016), reading and writing (Liou, 2016) skills. One of novel features of these apps currently

being researched is the inclusion of dynamic assessment (henceforth DA), by which the learner takes an active role (Rezaee et al., 2019).

Mobile Assisted Language Learning (MALL) is breaking new ground in the area of research into the process of language teaching and learning. In one study, listening comprehension skills were found to be more effectively developed when learners watched the movements of their own 3D avatar in Second Life as opposed to performing the same movements themselves or following the task without making any movements (Lan et al., 2018). This study is just one example of the multiple possibilities which have been opened up within the field of language education by these programs in recent years. Research into MALL has taken various methodological approaches and considered many different variables, among which can be found quantitative (Lin and Lin, 2019), qualitative (Luef et al., 2020) and mixed methods (Al-Ahdal and Abdullah Alharbi, 2021) studies. The majority of the studies report positive findings regarding the effectiveness of such apps (Rezaee et al., 2019).

The use of technology in education was already widely established, including Artificial Intelligence, but its full implementation has been accelerated by the circumstances of the pandemic and by the tendency to personalize Apps according to studies such as that of Parra-Sánchez (2022). It is not so much an advance in Artificial Intelligence but in the massive use of technologies in sectors of society where they had not penetrated so much before. For example, virtual glasses already existed as Nintendo's Virtual Boy released in 1995, but it is not until now that their use has become widespread. In education, in adolescents and post-adolescents, there is an overuse at times of mobile phones and social networks, so much so that, for example, it has become a good way of promoting the reading habit, as stated by Hernández Ortega et al. (2021). In this sense, Artificial Intelligence comes to take its place within the massive use of technologies, in societies that can afford it.

The app we investigated, Papua, is an artificial intelligence-based application which incorporates several functions, including voice recognition and pronunciation coaching. It is aimed at developing spoken fluency, reading and writing skills, grammar and socio-cultural skills. Its main features are the following:

- a. Instructional immediacy and ubiquitous learning. This means learners can access the app wherever they happen to be and so maximize the use of their time.
- b. Compatibility with most devices – smartphones and tablets – so that learners can easily connect and interact with the interface.
- c. Teacher supervision. The app allows teachers to supervise and evaluate their learners' quantifiable progress *via* the platform on the PAPUA website.
- d. Tools for receiving support and feedback on progress. The app scaffolds user learning when necessary, breaking it up into more manageable units (García Botero et al., 2021). Users access the website *via* a password, after which they can receive evaluative input from the teacher along with feedback on their performance through DA.

All these novelties are underpinned by the software which enables the device to give the learner immediate feedback. From the user perspective, the app is structured according to a series of scenarios following a storyline. The user takes the role of protagonist in these scenarios, and is offered a set of alternatives such that the story develops dynamically in response to the user's choices. The user can select both the degree of difficulty and the itinerary. The app dynamically updates these choices in real time to move the learner through the succession of different learning scenarios. At the end of each itinerary, the app produces a set of final results according to the user's responses along the way. These results include pronunciation (the phonetic-phonological dimension of language use), grammar (the morpho-syntactic dimension), vocabulary (the lexical and semantic dimension) and cultural issues and references (the pragmatic dimension).

Research on language learning app with IA

This study pertains to the relatively recent field of study into the effectiveness of apps. It evaluates an artificial intelligence-based app offering a series of specific features. It aims to assess the effectiveness of two aspects of the app, with the following specific objectives. First, find out users' recognition of the cognitive impact on memory. Second, find out users' perception regarding the effectiveness of the self-evaluation provided by the app. So, this paper explores users' appreciation of two of these aspects which previously have not received much attention, specifically learners' memorization of content (Al-Ahdal and Abdullah Alharbi, 2021), and management of evaluation a form of self-regulation, as understood by García Botero et al. (2021).

Both these dimensions were explored in this study of university undergraduates learning English as a foreign language. Memory turned out to be a decisive variable with respect to gender, age, and the students' level of study (Wang et al., 2021). In addition, the learning strategies based on memory were found to be more stable among the students participating in the study.

At the same time, we took self-evaluation into account as the second research dimension, among other reasons, because of the current boom in this kind of methodological approach to evaluation, and because of the implications it could have for other variables of an affective nature (García Naranjo, 2018; Andrade, 2019; García Sanz, 2020).

In monitoring these two dimensions, we aimed to cover both direct and indirect strategies involved in language learning (Oxford, 1990). Hence, with respect to direct strategies (memory strategies, cognitive strategies and compensation strategies), memory was considered the chief factor in the study, whilst in the case of the indirect strategies (social strategies, affective strategies and metacognitive strategies), we took self-evaluation into account for its affective and metacognitive connotations as a general methodological strategy in language learning.

Beyond the novelty of monitoring these two combined dimensions in university level EFL students, the main contribution of this study lies in the analysis of these strategies in relation to the use of the Papua virtual reality tool for learning English. This meant observing the strategies operating in the context of a virtual learning methodology. In the last few years, this kind of methodology has become extremely popular, to the extent that many researchers have expressed the need to further develop learning tools of this kind (Chu-Chu et al., 2013; Santamaría and Alcalde, 2021). For this reason, we took SE into consideration as a potential means of improvement in this process in comparison with other MALL applications.

Memory

As Huang et al. (2022) note, memory has traditionally been considered one of the four components of linguistic aptitude (Carroll, 1981) for learning first and second languages, alongside phonemic coding ability, grammatical sensitivity, and the ability to learn inductively. Specifically, associative memory – the ability to make connections between stimuli and stored responses – was long considered to be a major factor in language learning, but working memory has come to supplant it as the key aspect of memory in this area (Skehan, 2015). Working memory is linked to complex cognitive tasks such as understanding language, learning and reasoning. Through these tasks the system not only understands and manages stored information, but also stores and processes the information temporarily at the same time.

Working memory is thus connected to the central executive system, attention control, and three subordinate systems, namely the phonological, visual spatial information, and episodic memory. All this is what allows working memory to be able to simultaneously store information and process it in context. This processing implies that the content stored in working memory has been through a selection process, given the huge amount of sensory information the individual is constantly receiving. In this way, working memory operates with conscious information, while at the same time there is potentially manageable virtual memory. This duality between stored information and potentially processable information, along with its link to other higher cognitive tasks, turns out to be vitally important in the context of EFL/ESOL learning/acquisition (Shen and Park, 2020). As the two authors note (2020, p. 86):

According to this model (Baddeley, 2000), working memory is comprised of four components which process distinct types of input information. Auditory information is stored and processed by the phonological loop, while the visuospatial sketchpad is responsible for visual images and spatial relations. The central executive plays a role similar to a traffic cop who regulates the other components of working memory and allocates cognitive resources. The episodic buffer is assumed to be the storage component of the central executive, and a link from working memory to long-term memory. In the information processing model, long-term memory is the final stage.

Working memory

Working memory is of paramount importance, due, among other factors, to the constant feedback between the phonemic coding/decoding and the visual spatial management of alternative learning scenarios. Especially so in the case of the app in question, Papua, as said learning scenarios are virtually unpredictable, as they are the result of the prior choice of difficulty level on the part of the user, and their responses during the interactive leaching-learning process. In consequence, the attentional component plays a major role in the learning process, given the constantly changing nature of the scenarios and situations of said learning contexts.

The overlap of learning scenarios leads to the need to handle a greater quantity and wider variety of information, which can potentially be stored in long-term memory after first being selected by working memory. Consequently, the memorization of information can be increased with this kind of virtual tool for the learning/acquisition of a foreign or second language. In this regard, this study considered that the Papua virtual tool would improve these memory-related strategies. The students needed both declarative and working memory in order to complete the activities included in the app. The study by Al-Ahdal and Abdullah Alharbi (2021) provides an appropriate starting point. These researchers showed that the collaborative use of mobile devices by the experimental group improved their retention of vocabulary post-intervention, while overall group performance drastically improved, with more learners scoring closer to the mean value. The control group, by contrast, showed no significant difference in performance.

The first research question forming the basis of this study was the following: Does Papua improve users' vocabulary retention as a result of its interactivity and constant feedback on their responses? This feedback relates to the possibilities for storage of content by students and the operative possibilities for simultaneous processing of this information by the subjects of the sample. Both tasks are linked to the users' working memory.

Storytelling and memory management

At this point, we should also highlight the concept of storytelling as a thematic thread running through the interaction between the user as protagonist and the Papua app as another influential factor in memory management.

As McGregor and Holmes (1999) note, the concept of storytelling boosts long-term memory in comparison with other narrative techniques. Secondary cognitive effects are produced through storytelling techniques, and these contribute to idealization and satisfaction in relationships, which in turn converge on the quality of the foreign/s language learning/acquisition process by virtue of affect (Arnold, 2000).

In similar vein, the concept of storytelling, as adopted by Papua, draws on the so-called *enactment effect*, also known as the *self-performed task effect*. The effect basically refers to the fact that a linguistic sequence represented by an action is memorized better than a linguistic sequence without such representation, especially if the learner physically enacts the sequence virtually. This representation of action has an effect on action memory (Zhang and Zuber, 2020).

In our approach, it was expected that these novelties in Papua would contribute to an improvement in the quality of the learning process in terms of vocabulary retention as a result of the development of working memory on the part of the user, who has to simultaneously process the stored contextual information. Relevant studies in this regard include the traditional use of narratives in the primary classroom (Kirsch, 2016), and its use as a teaching strategy in language learning (Rezende, 2016). These studies found positive results for storytelling techniques applied to learning languages.

The present study builds on these previous studies with the aim of verifying the degree of success of the storytelling concept for learning English with two new variables: the sample was composed of adult university education students, and the storytelling element was carried out in a virtual interactive environment. In place of a physical storyteller, Papua uses an action framework in which users visualize themselves in a learning scenario requiring a series of actions related to learning English to be carried out.

This approach represented a step forward in internalizing content. Rather than visualizing an external agent for carrying out the action, or assimilating content in an abstract manner, users themselves carry out the activities, thus reinforcing retention.

Further, for verifying the success in learning/acquiring vocabulary, the Papua app structured strategies involving the repetition of words and expressions by the user according to a variety of sentence and sub-sentence-based schemes inserted in certain communicative situations, at the same time the same program pronounced a series of words and expressions chiefly by designation. The program, through said strategies for the learning/acquisition of EFL/TESOL, featured a function for correcting the user's pronunciation while it offered a rating as a result of a process of evaluating the quality of the pronunciation (which was also extendable to other functions and modules, such as lexis, through the evaluation of vocabulary and grammar retention).

These strategies resulted in simultaneously enhancing vocabulary retention and improving pronunciation and comprehension of English as foreign or second language (thus linking to the oral skills of speaking and listening).

This derives from the fact that working memory, as mentioned above, took on a major role in the interactive learning/acquisition process between the students and the software, which gave rise to the processes of phonetic coding and decoding being reinforced in comparison with other programs and tools for learning a foreign or second language. Previous studies of relevance include those (Han, 2015; Han et al., 2017; Wang, 2019). These studies found a correlation between the working memory and oral skills of Chinese learners of English across a range of ages, with a stronger correlation of working memory with oral production than with writing skills. Nevertheless, working memory was found not to have such significance in relation the written skills according to other researchers into English language learning with Chinese students (Yi and Luo, 2012; Yi and Ni, 2015). In these studies, aspects such as motivation and self-regulation in the

language learning process were taken into account. This led us to take into consideration, in addition to the memory component, self-evaluation as a process in monitoring one's actions and their outcomes.

Additionally, one of the unique features of the Papua tool is the possibility of real-time narrative-based interaction in the form of a story. Through this feature, users can choose from various response options which entail corresponding changes in the way the action unfolds. We thus consider that memory represents one of the fundamental element for students, as the choices they make will converge in parallel stories. This means working on the two main dimensions of memory: declarative and procedural. Likewise, the possibility of intervening in real time in the story enhances the general constructivist principle of the protagonist role of the learner during the learning process.

Evaluative feedback from the app

Papua provides feedback on user performance across a range of scales corresponding to various aspects. To this extent, the process constitutes a form of self-evaluation (SE), a common practice in educational settings in different domains. The teacher can also access the student's progress on the website platform by tracking the textual indicators providing the specific feedback. This metacognitive tool allows students to gauge their progress (Rodríguez Gómez et al., 2009; Fernández-Novell and Zaragoza Domeneq, 2014), and to gain autonomy (Fernández, 2011). SE has especially been taken up in language education (García Botero et al., 2021), and has been received positively by education graduates as a means of increasing participation (García Botero et al., 2019), although it has been criticized elsewhere.

Calatayud (2002, p. 358) defines SE as *the foremost strategy for developing responsibility and for learning to value, critically assess and reflect on the process of individual teaching and learning. It concerns the knowledge of what has, and has not, been achieved towards gaining this command, and what further steps can be taken to correct errors and improve the outcomes*. García Botero et al. (2021), p. 41 notes that learning strategies are acquired through SE: *SE is not only the evaluation of the learner's own performance or the ability to make judgements about their own performance, but rather a means of developing strategies for learning and for improving learning outcomes*.

With respect to Papua, the SE adopted falls within the framework of *Mobile-based Dynamic Assessment* (MDA) in which the student has a positive attitude to language learning. It provides users with detailed information on quantitative scales featuring two measures focusing on positive feedback and three on negative, in the following five areas:

- Pronunciation score. This function awards the learner a score for the accuracy of their pronunciation in English. 100% is a perfect score.

- Response score. This function also gives the learner a score out of 100 for the speed and fluency of their response.
- Wrong answers (content). This item gives the learner negative reinforcement according to behaviorism model.
- Wrong answers (grammar). This item gives the learner negative reinforcement according to behaviorism paradigm.
- Wrong answers (socio-cultural competence). This item gives the learner negative reinforcement according to behaviorism theory.

The effect of MDA on students' EFL fluency was investigated by Rezaee et al. (2019). Analysis of the results showed that the experimental groups – that is, those receiving MDA – performed significantly better than the control group in the post-test. The findings of the study suggested that MDA had a positive effect on the EFL learners' oral fluency. Hence, after considering the findings of these previous studies, we hypothesized that the constant feedback provided for the students by the app would improve their strategies for learning English as a result of mainly their indirect strategies being enhanced and the leading role of the learner (Oxford, 1990).

In similar research, García Botero et al. (2021) carried out a study into self-regulation and scaffolding for supporting out-of-class MALL. One of the two experimental groups received training in self-regulation and scaffolding support in preparation for their MALL experience. The results of the study suggested that this training was beneficial for student uptake of MALL, along with several other benefits. Those students who were trained in self-regulation and received temporary scaffolding demonstrated higher levels of participation and greater motivation than those who received no preparatory training. The study found a significant link between self-regulation and intrinsic motivation.

On the other hand, we think that the opportunity for self-evaluation, in addition to being motivating, might represent a significant factor in affective involvement in the learning process, as it provides their avatar in the app with real choices that the students must make. The projection of the learner from the avatar allows the student to monitor their own learning process in real time. This provides greater autonomy to it with respect to its evolution and, therefore, may have implications in factors related to the affective and cognitive dimension of the learner's attitudes by virtue of their motivation according to a possible link between autonomy and self-regulation of the process of learning/acquisition of foreign languages/s languages.

Intrinsic motivation

Intrinsic motivation, a psychological construct, refers to the personal satisfaction inherent in an activity which leads to the engagement and enjoyment of the person involved. Intrinsic motivation thus constitutes an end in itself (Deci and Ryan, 2000). This has currently been corroborated by studies with positive results into self-regulated learning (Wang, 2019; Chih-Ming et al., 2019).

Likewise, there have been various positive studies into how language learning apps can draw on intrinsic motivation using MBA.

Cahyono and Ludwig (2017) indicated that intrinsic motivation and identified regulation became essential factors in the students' motivation to participate in the activity. Jenö et al. (2020) found a credible interaction effect between the mobile app and intrinsic goal-framing for intentions and identified regulation. For the variables *effort* and *achievement*, the main effect of mobile learning was consistent with substantial effect sizes.

Chin-Hsuan et al. (2021) indicate that conscientiousness and m-learning readiness are critical antecedents of intrinsic motivation and extrinsic motivation. In addition, both types of motivation had a significant positive effect on students' intention to use this kind of app in terms of encouraging progress through the language learning process.

Promoting progress and personalized learning

According to (Pérez Ariza and Hernández Sánchez, 2014), the question of progress in learning has been the subject of detailed study within the fields of both psychology and education, with valuable contributions by internationally renowned theorists such as Thorndike (1931), Piaget (1972), Bruner (2001), Vigotski (1987) and David et al. (1991). These authors affirm that any learning process presupposes some kind of cognitive transformation in the individual, such that any process concerns its own development.

In the field of educational technology, automatized digital tools have been designed to promote learning and, in some cases, provide feedback. For example, Charles (2009) filed a patent for an automatized device for language learning which draws on specific content from a database, combining this information with images for viewing. It also promoted active listening on the part of the user. Other devices, such as the system created by Kao et al. (2009), evaluate learning progress in the following manner. The learning module plans lessons according to the progress measured by the test module. An evaluation module sets the evaluation for the lessons covered and provides feedback with the aim of evaluating user progress. The progress test questions are stored in a database along with the responses, so that these can be considered in future tests.

García Botero et al. (2019) carried out a study comparing student SEs with the assessment descriptors provided in an oral communication skills course reaching the conclusion that students require guidance in carrying out SE activities. The study also noted the need to pay special attention to the wording of the SE grid used by the students to assess their learning, given that few students used this wording in their self-reports, and instead either mixed the given wording with their own expressions, or relied entirely on their own wording. It was also notable that the perceived importance of particular areas of learning did not match the self-evaluated progress across the same range. All these tools take into account individualization and personalization in the process of learning a foreign language.

Personalized learning is an area which has been considered by educators and teaching professionals at both the theoretical and empirical levels (Fernando Calderero et al., 2014). Over the last few decades, the approach has become an established paradigm

applied across a range of educational contexts. According to the authors (Fernando Calderero et al., 2014, pp. 140–141), it is characterized by three main features: uniqueness, autonomy and openness. In the definition provided by García-Hoz (1988, p. 25), ‘personalized education is personalized to the extent that it is applied to an individual with their own unique set of characteristics, who feels engaged and committed to realizing their personal possibilities, and who, at the same time, is ennobled by the mere fact of being and doing as a person.’

Educational technology is also directing itself towards personalized learning, with various patents being filed to this effect. An example is Chao et al. (2009), which describes an intelligent, interactive learning system. The system automatically generates learning materials by capturing and restructuring information from the internet. By tracking users’ learning behavior it is able to provide personalized materials more efficiently. According to Zhou et al. (2016), among current areas of research into personalized learning can be found data-mining, massive open online courses (MOOC), the learning environment, learning analysis, ubiquitous learning, learning styles, E-learning, and teaching modes. In like fashion, research into personalized learning models has tended chiefly towards exploring theoretical frameworks, diversification, and the implementation of personalized modes and technology.

Aims

The general aim of this research paper is to evaluate ‘Papua’, an app designed to facilitate learning English, whether as a foreign (EFL) or second (ESL) language. The app downloads onto any portable device (smartphone or tablet), and is the product of a start-up called Aprendizaje Inmersivo S.L. (Immersive Learning Ltd.), based in Móstoles (Madrid). It has been developed by a multi-disciplinary, multinational team, and uses artificial intelligence to develop different communicative competences.

The specific aim is to evaluate the efficacy of the app in two dimensions: memory reinforcement and self-evaluation from the process of self-regulation of learning/acquisition of English as a foreign language.

Research questions

From all the above, this study considers four research questions:

1. What impact does the virtual tool Papua have on memory-based strategies with respect to university level learners of English as a foreign language?
2. What impact does the self-evaluation methodology of the virtual tool Papua have on the English language learning of university level students?

3. Does the intervention program through the Papua app improve the oral skills of the students comprising the sample more than their written skills?
4. Did the incorporation of the components *intrinsic motivation*, *monitoring of progress*, and *personalized learning* into the SE function of the Papua app, have a significant impact on the language learning process?

Materials and methods

Participants

We recruited a sample of students in the second year of a degree in Education in the academic year 2020–2021, who participated in a quasi-experimental study meanwhile they followed an Education Degree in the university level. In the case of these Spaniards, the students had a particular interest in participating in the study as, in order to graduate, they were required to accredit a level at B1 (as specified by the Common European Framework of Reference) in any of the official languages of the EU by means of certification through one of the many recognized examination bodies, although for the purposes of this study the language selected was English. This stipulation is a recommendation by the Council of Europe, take-up of which has not been uniform either across the EU or indeed within Spain itself, where devolved responsibilities have seen significant variation across autonomous communities in terms of implementation and required level, Andalusia (where the study took place) currently opting to implement the recommendation at B1. Given its status as a *lingua franca*, English is the predominant second language choice, as achieving the stipulated level not only meets their degree requirements, but also provides students with a necessary competence in the labor market.

The sample consisted of 128 students from two groups in their second year of the degree in Primary Education at the University of Huelva (see Tables 1, 2).

Students’ responses were stored in a database, from which each variable was analyzed in turn. The variables were as follows:

- Age; six intervals were established: 18–19; 20–21; 22–23; 24–25; 26–30; >30. The average age was 23.2.
- General level of English; three bands were established: below B1; B1 (non-certified), B1 (certified). 77.8% of the respondents were in possession of B1 certification.

TABLE 1 Age of participants.

Age band	18–19	20–21	22–23	24–25	26–30	>30
No. of participants	40	57	20	4	4	3

Six age intervals were established for classifying the participants.

TABLE 2 Gender of participants.

MALE: 37 participants	FEMALE: 91 participants	OTHER: 0 participants
-----------------------	-------------------------	-----------------------

Participants self-identified their gender.

- Gender; three responses were available: male, female, other.

Instruments and data collection

Semi-structured interview

A semi-structured interview was used, hosted in Google Forms, consisting of a series of quantitative questions drawn from the CAAAL questionnaire (Peña-Acuña and Crismán-Pérez, 2021), plus two qualitative open-ended questions which gave participants the opportunity to amplify aspects of the previous questions relating to memory and self-evaluation.

The CAAAL questionnaire had been validated in a previous study (Peña-Acuña and Crismán-Pérez, 2021; see Appendix). It consisted of 13 items, with the following Likert scale options: *not really, to a limited extent, reasonably well, quite a lot, very much so*. It was designed to measure *attitudes to learning* and *linguistic attitudes to L2*.

The two qualitative open-ended questions that were added to the semi-structured interview were the following:

1. 'Why do you think that repetition of content in Papua improves memorization of English content?'
2. 'Why do you think that self-evaluation in various aspects (average score in pronunciation, average number of correct answers, correction of pronunciation, incorrect answers in terms of content, grammar or appropriateness) in each scenario in the Papua app improves intrinsic motivation for learning English?'

Procedure

The study followed a mixed methodology within the framework of a previous six-week experimental teaching intervention with university students in the second semester. The students were provided with a password which gave them access to 9 scenarios at the basic level of the app. They were given time to explore the scenarios and test out the app at their leisure, so as to get to know it for themselves and assess how useful they might find it. At the end of this period, the participants remotely evaluated different aspects of the Papua app in terms of various characteristics and components of the software. Specifically, they completed a semi-structured interview composed of a quantitative questionnaire with two additional qualitative questions available in Google Forms. This enabled the researchers to collect data and learn the students' assessment of the experience.

The advantage of this sample was their background in teacher education, meaning they were familiar with pedagogical concepts. The results were contrasted with two basic components of the foreign language learning process: memory (Ellis, 1996) and self-evaluation, which influenced oral production skills (Kissling and O'Donnell, 2015). Nevertheless, these results should be compared with other prospective studies, both of apps as learning software and of other non-digital tools for learning a foreign language.

Data analysis

The analytical procedure took the following methodology. First, all documentation concerning the app provided by its creators was read by the researchers so as to familiarise themselves with the conceptual origins of the app and the features it offered. The researchers then followed this up with two interviews with the creators of the app. The aim was to find out more about the kind of learning its creators hoped to achieve through the app. The outcome of these interviews was the addition of the two qualitative questions to the quantitative questions in the CAAAL questionnaire in order provide a semi-structured interview in accordance with a mixed methodological approach. These open questions focused on the individual components of Papua in relation to the factors included in the quantitative section.

The same 128 participants then individually completed the two-part questionnaire on Google forms while their experience of using the app was still fresh in their minds.

The quantitative analysis

Once the responses of the sample had been collected, the quantitative data were analysed using statistical tests in the SPSS package (version 25).

After the questionnaire had been statistically fit to the sample in question, Cronbach's alpha was over 0.7, while the KMO index was 5.1, with a significance of 0.001 with respect to these two variables (attitudes to learning and linguistic attitudes to L2). It was administered online *via* Google forms at the end of the six-week intervention period during the academic year 2020–2021.

These variables were contrasted with the responses to the dimensions below, giving particular focus to those concerning memory and self-evaluation. In addition, questions 5, 8, 10, 11, 12 and 13 in the CAAAL questionnaire (Peña-Acuña and Crismán-Pérez, 2021) were considered. The full list of dimensions is as follows:

- Oral competences
- Reading competence
- Linguistic communication (interpretation and communication feedback).
- Cognitive knowledge facilitators (memory)
- Self-evaluation report (response rating for each content unit).

The qualitative component

In addition, regarding the qualitative analysis of the collected data, an initial analogical analysis of the components was carried out by the researchers. Finally, when the components for each question were available, the analysis of the qualitative data was carried out using the QDA Miner program, version 2.0.8. This program allows categorization into components, subcomponents and descriptors. The main advantage was that the program allowed the use of qualitative data alongside the quantitative frequency of the subcomponents used by the respondents, enabling a percentage hierarchy of standout terms in the sample to be established.

The advantage of incorporating a qualitative dimension in the study is that it enables the topic to be studied in greater depth, with a wide range of variables, some expected and others emergent. These data are of a different, but complementary, nature. While the quantitative data place the emphasis on the objective description of phenomena external to the individual, the qualitative perspective interprets the subjectivity of the agents involved, along with those phenomena resulting from this interaction (Castro, 1996, 64).

The qualitative study was carried out following the eight “big-tent” criteria of rigor and high quality expressed by Tracy (2021, 178), namely: methodological qualitative research tackling a worthy topic, rich rigor, sincerity, credibility, resonance, significant contribution, ethics and meaningful coherence.

Results

Quantitative results

With respect to the quantitative results, the variable *level of English* correlated positively with the variable *oral competence* ($r=0.36$; $p<0.05$; $r=0.05$ $p<0.01$). This result was corroborated by a one-factor Anova for the variable *certification of English* (B1 as minimum level) in relation to oral competence where $F(3,215)$ $p=0.25$. The Anova test also gave a result for the cognitive facilitator, *memory* $F(2,937)$ $p=0.36$. These data indicate that were H_1 , H_2 y H_3 fulfilled.

The variable *age* did not display any significant correlations with the other variables. However, the Anova test presented a significant relation $F(892)$ $p=0.44$.

The variable *gender* (feminine) displayed a significant correlation with the variable *level of linguistic communication* ($r=0.94$ $p<0.05$).

Qualitative results

First part. Cognitive factor: Memory

The starting question was: ‘Why do you think that repetition of content in Papua improves memorization of English content?’ The starting subcomponents about which it was intended to

deepen before doing the semi-structured interview were the following:

- Subcomponent: emotion favoring memory retention.
- Subcomponent: content repetition favoring memory retention.

The results of the semi-structured interview indicated three aspects worthy of note. The first of these was the perception that the app had a positive effect on users’ memorization of content. The second was that participants underlined the situational approach favored by the app, along with other positive features, while the third was an overall constructive response to the app, with a minority of negative comments.

Returning to the first – the perception of the beneficial effects of repetition on memorization – this was the predominant factor across the user sample (56.7%), providing support for with H_3 . The respondents reported that this memorization was achieved by means of understanding material delivered *via* the oral, visual, written, spoken and aural expressive manifestations (6.8%); paying attention to error correction (1%) in order to improve pronunciation (4.5%); and engaging with the metacognitive resources (0.3%) to check their own learning, in line with H_4 , H_5 and H_6 .

With regard to the second aspect – the positive features of the app – the elements most frequently identified were the incorporation of gamification and the situational approach (12.3%) by means of varied resources (activities, dialogues, laboratory, translation, etc.), and learning through storytelling (1.5%). Finally, the application received an overall positive rating (4.5%) and was considered to provide motivation for learning (5.8%).

Within this third aspect were to be found critical appraisals, from both negative and constructive perspectives, concerning certain structural aspects of the app of less significance.

Second part. Improving intrinsic motivation through self-evaluation

The starting subcomponents about the self-evaluation before passing the qualitative question of the semi-structured interview consisted of this list of items:

- Average score in pronunciation.
- Average number of correct answers.
- Incorrect answers (content errors).
- Incorrect answers (grammatical errors).
- Incorrect answers (socio-culturally inappropriate).

The starting question in the semi-structured interview was: ‘Why do you think that self-evaluation in various aspects (average score in pronunciation, average number of correct answers, correction of pronunciation, incorrect answers in terms of content, grammar or appropriateness) in each scenario in the Papua app improves intrinsic motivation for learning English?’

Three main aspects emerged from the qualitative analysis of the responses to this question. Firstly, respondents noted the positive attributes of the app. Secondly, as a constant in the results of the previous question, they highlighted the positive qualities of the app and other positive aspects. Thirdly, they expressed both constructive and negative critical appraisals.

With respect to the first of these aspects, the predominant attribute identified was motivation (54.4%) followed by learning progress (25.8%). It is important to highlight that respondents perceived these two attributes as connected, sometimes causally, sometimes correlatively. Other attributes which emerged included the ability to personalize the app (9.4%), and the self-evaluation function (2.4%).

The second aspect highlighted positive adjectives applied to the app (6.7%).

The third aspect concerns critical appraisals regarding the self-evaluation function (0.9%), including a negative one (0.3%), although these positions were in the minority.

The third component refers to more critical (0.9%) and even negative (0.3%) positions in the minority regarding the self-assessment element in the application.

Discussion and conclusion

As will be shown below, both the general and specific aims of this study were fulfilled. Regarding the first research question, concerning the perception of an improvement in learners' memorization of vocabulary using Papua, the study provided support for the findings of previous studies into MALL methodology (Han, 2015; Han et al., 2017; Wang, 2019). In short, with respect to the first research question, an improvement in vocabulary retention was perceived in the following specific respects:

- The reading activities in the app boost memorization of words.
- Taking a role involving the expression of one's emotions improves the assimilation of English vocabulary.
- Activities involving repetition of material, or a spiral syllabus, improve memory capacity.

The findings from the study are similar to those of other research in the field. Barrett et al. (2021), investigated the attitudes of 30 university students towards using app called EOPA for oral presentations in English. In this case, 37% reported a positive attitude, while 33% reported having experienced difficulties in using it.

A meta-analysis carried out by Lin and Lin (2019) of 33 quasi-experimental studies published between 2005 and 2018 found an overall positive attitude towards both mobile applications and the use of MMS/SMS for language learning. This latter modality proved to be most effective with regard to retention of words and other lexical units in L2. The meta-analysis also revealed that there remain relevant areas to be researched in the field, including the

settings in which the research takes place, the autonomy such applications confer on the user, and the duration of the experimental periods. Rosell-Aguilar (2018) demonstrated that mobile teaching applications improved L2 learning especially with regard to lexis, in terms of the memorization of words.

With regard to the third research question – the issue of whether Papua favors an improvement in oral competences above the written – our findings concurred with those of Shen and Park (2020), who studied the relationship between working memory and English language learning by Chinese students using MALL methodology over the last 20 years. Their findings showed that oral competences improved more than other dimensions involved in second language learning.

Our study of Papua likewise indicated that the app positively impacted oral competences more than it did other competences and dimensions involved in learning English (Yi and Luo, 2012; Yi and Ni, 2015). In this respect, we would maintain that the use of artificial intelligence for voice recognition is a key function of the app, and contributed to the improvement in users' pronunciation. There is, however, a need to confirm the validity of this finding through future research into the effect of such apps on the attentional component of working memory (see Tables 3, 4).

With respect to the second dimension considered by this study – the relationship between self-evaluation and users' intrinsic motivation – the results of the study confirm the second and fourth research question. That is, through the incorporation of SE, the app successfully provided intrinsic motivation, monitoring of progress and personalization of learning. The findings confirmed the relationship between gamification and student motivation (Dehghanzadeh et al., 2021), as analysis of the answers from the semi-structured interview indicated.

These results are consonant with those of a study carried out by Yeh and Chen (2019), which found that the SE component of the language learning app under consideration had most impact on students' performance and motivation, irrespective of gender. Likewise, a study into the effect of MDA on learners' spoken accuracy carried out by Rezaee et al. (2019) concluded that MDA positively influences performance. Consequently, our study can be placed alongside those such as Cahyono and Ludwig (2017), and Chin-Hsuan et al. (2021), which provide evidence for the positive benefits of MDA for English language learning.

It can be seen from the literature review that professionals in field have a broadly positive view of SE as an educational and metacognitive tool, with the monitoring of progress, attention to learning strategies, and learner autonomy all features which have been foregrounded. In this respect, García Botero et al. (2021) conclude that learners who receive training and scaffolding for self-regulation within a MALL context benefit more from using the language learning platform. Nevertheless, some possible criticisms can be noted.

In the qualitative phase of the study, the respondents foregrounded intrinsic motivation, thus providing support for the following findings corresponding to the second and fourth research question:

TABLE 3 Improving memorisation of content.

Components	Subcomponents	Descriptors
Memory retention	Repetition and learning of content 56.7%	Vocabulary, expressions, phrases, grammatical structures, helps internalize, recognize, associate, strengthen neural connections, reinforce visual and auditory memory, in written and spoken form
	Comprehension 6.8%	Content understood in relation to practice
	Pronunciation 4.5%	Improvement in pronunciation, recognising pronunciation
	From errors 1%	Dichotomy correct form/error, learning from errors
	Metacognitive resource 0.3%	Learning checks
Qualities of the app	Situational learning 12.3%	In a natural context, everyday context, learning to communicate in English; conversations set everyday situations; based on communicative situations; through various resources and activities, revises material through a choice of different ways; activities, dialogues, learning laboratory; you learn behavior through sounds, words and pictures
	Motivating 5.8%	Not monotonous, you do not get bored, it grabs your attention, it gets you hooked, easy, nice
	Gamification 5.3%	Game-based, involves games
	Positive qualifiers about the app 4.5%	Fantastic, brilliant, original, different, new, dynamic, fun, entertaining, complete, innovative, useful, interactive, effective
	Use of storytelling 1.5%	Story, narrative, episodic stories, divided into scenes; a real story in which we take the lead role
Critical attitude towards the app	Negative 0.8	It's badly thought out in this respect, as it makes us look at the same terms several times; so much repetition becomes tiresome
	Constructive 0.5%	I'd add more scenes and more varied activities at each level; I'd use visual images for memorising words; alternate different games at the end of each scene

TABLE 4 Improving intrinsic motivation through the app's self-evaluation module.

Components	Subcomponents	Descriptors
Positive functions of the app's self-evaluation module	It motivates the user 54.4%	It allows the user to keep making progress and motivates them, they do not feel the pressure of an exam or a classroom situation; there is intrinsic motivation; it's motivating; it makes you concentrate at the higher levels; your self-esteem goes up; it gives you recognition; it makes us feel proud of ourselves; it makes you demand a lot of yourself; it makes learning English like a game; it makes you push yourself; you enjoy learning
	It encourages learning progress 25.8%	Students receive different tips for improving; it makes you work harder; it helps to improve your social interaction, pronunciation, writing, reading comprehension, awareness of the other culture, knowledge of the language, the sentence structures help you gain fluency in English; it helps you to overcome embarrassment and difficulties in pronunciation
	Personalization of learning 9.4%	People who are shy or get embarrassed easily might it suits them; you can identify with it; it makes you feel secure; it takes the user's decisions into account; you can go at your own pace and level; it lets you analyse and assess your strong and weak points; you learn from your errors; it gives you autonomy and self-direction; active agent; it encourages self-reflection and feedback; it makes us focus more on where we need to improve; it lets you learn on your own; having more confidence in the things you do right
	Self-evaluation: outstanding feature 2.4%	You find out where you went wrong; it's very detailed and precise; it's even more important in this app; it gives you scores and feedback; the parameters it measures are very interesting; it focuses on what you do right and gives you recognition for this
Qualities of the app	Qualifiers about the app 6.7%	It's very original; a different learning tool; innovative, interactive; suitable and fun; useful; game-based, visual and efficient; interesting
Critical attitude towards the app	Constructive 0.9%	If it corrected you immediately or explained what the cause of the errors you made was in the test at the end of each stage; it ought to give you the chance to listen to the different opinions first and then give the answers by speaking; it should evaluate you a different way, explaining where and why you went wrong
	Negative 0.3%	The app has positive features, but this does not necessarily mean it boosts motivation; it does not seem to me to be very accurate in its pronunciation scores

- Self-evaluation in terms of average scores in pronunciation and answers for each scene boosts users' intrinsic motivation for learning English.
- Self-evaluation in terms of wrong content-based answers in each scene boosts users' intrinsic motivation for learning English.

- Self-evaluation in terms of wrong socio-cultural answers in each scene boosts users' intrinsic motivation for learning English.

Further, it is worth noting that monitoring progress and personalization were also found to be significant factors,

albeit to a lesser degree. Although monitoring progress has become an area of continued educational interest, as evidenced by the increase in the number of patents filed for related technology, there has also been some skepticism towards SE on the basis of mismatches between the descriptors and students' own perceptions (García Sanz, 2020). Similarly, a review of the literature shows that personalized learning *via* technology is highly regarded among educationalists, and is one of the most studied resources for language learning (Zhou et al., 2016), prompting a comparable deluge of patents (Chao et al., 2009). The overwhelming interest in the potential of these tools is what led us to research their effectiveness in the L2 learning process, with the aim of identifying positive common results, and complementarily, areas and aspects for improvement.

Future research

As noted above, the assessment of the Papua app carried out by this has been largely positive. Nevertheless, some drawbacks were noted in the course of the research, as well as among the findings, and these will be considered below.

First of all, as García Botero et al. (2019) notes, despite the self-directed learning capabilities of this kind of tool, the provision for training and scaffolding is nevertheless to be highly recommended. This would take the form of some kind of tutorial or partially in-person course to ensure that users get maximum benefit from its use. Otherwise, there is a high risk of learners becoming demotivated.

In addition, it would be interesting to complement the research presented in this paper with a study into the use of the app as a long-term methodological component in a language learning course, as opposed to the short-term (6-week) period which formed the basis of this study. As Shantanu Tilak et al. (2021) point out the majority of studies tend to focus on the design and evaluation of the tool under analysis itself, with far less attention paid to assessing the sustained effectiveness of such digital apps. The findings of this study should thus be treated with caution, until such a time as a longitudinal study can be carried out to explore the possibilities of this tool in particular, and the use of MALL methodology in general, as part of the everyday language learning resources.

In like manner, the current study focused on vocabulary retention by virtue of its link to memory. Nevertheless, given that the tool allows the enactment effect to be incorporated, we believe it would be useful to undertake complementary studies to explore the improvement of the user's pronunciation on the phonemic plane, as well as their grammatical knowledge. Such studies would enable us to assess the effectiveness of the tool over a range of language skills, and would allow the traditional cognitive monitoring processes applied in this tool to be compared with other tools and methodologies for language learning. In this respect, the work of Verhagen et al. (2015) and Serafini and Sanz (2015) is worthy of note, as both studies found a correlation between

phonological and grammatical memory. This kind of directed research, exploring the connections between different modules involved in the language learning process, would help Papua to be developed so as to take account of the different responses of the sample according to their intrinsic characteristics in terms of age, level of English, and gender. Likewise, another function that we have considered, taking into account the significant relationship between working memory and the memory for the oral competences and vocabulary retention, is the addition of an instrument for measuring working memory, along with an alternative methodology for increasing this kind of memory.

It is also worth noting that, as Ni (2018) and Shen and Park (2020) maintain, working memory and vocabulary storage in learning English as foreign language are correlated, something which is equally true at the initial stages of learning English as a foreign language (Li, 2004; Jiang et al., 2015). This leads us to the conclusion that it is important to consider memory as a major dimension within the field of language learning, and more specifically, the development of working memory for its interrelation with language learning and in itself.

As noted above, the study described here threw up an emergent variable in the shape of storytelling. This was identified as aiding both memory and motivation while learning English in a MALL context. This variable could fruitfully be the subject of a research question in future studies exploring its influence on the effectiveness of this kind of digital tool.

Finally, we also need to be aware that our study focused on a sample of students whose first language was Spanish. If we are to develop a broader and deeper theory of the advantages of using digital tools such as the Papua app, it will be necessary to take account of a wider range of language backgrounds. In similar vein, it would be interesting to broaden the range of languages available to be learnt on the app so as to note whether the findings regarding working memory and SE were consistent with those presented here. This would perhaps allow us to make more general statements about cognitive aspects such as the connections between memory and language learning, and at the same time establish templates for this kind of learning consistent with methodologies like SE.

We should also be aware of the absence of any measurement of affect in the study. As Arnold (2000) points out, affect is one of the key factors influencing second language learning due to its role in determining the learner's attitude towards the language and towards their perception of success, beyond any other factors involved in the process. This raises a new issue for MALL methodology, namely, how to make up for the reduced levels of affect resulting from learners working in isolation. As we suggested above, one means might be to complement this kind of learning with some degree of in-person input or tutoring, but it remains, we think, an interesting question to be tackled in future studies.

At this point, in prospective studies we also intend to examine the relationship between the self-regulation of learning/acquisition of a foreign language/L2 from a digital tool such as Papua and a

possible correlation with the motivation variable, which is linked to the plane affective behavior of students' attitudes (Arnold, 2000).

For this we must start again from the traditional EFL learning/acquisition strategies (Oxford, 1990). However, it is necessary to review these strategies according to a digital learning framework, due to the latest innovations in software for the learning/acquisition of foreign languages/L2 and the new demands and needs of students, especially noticeable since COVID-19, and the EFL virtual learning/acquisition methodology.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization and investigation, BPA and RCP; methodology and formal analysis, BPA and RCP; Writing – Original Draft, BPA and RCP; Writing – Review & Editing, BPA and RCP. All authors read and approved the final manuscript.

References

- Al-Ahdal, M. H., and Abdullah Alharbi, M. (2021). MALL in collaborative learning as a vocabulary-enhancing tool for EFL learners: A study across two universities in Saudi Arabia. *SAGE Open* 11:215824402199906. doi: 10.1177/2158244021999062
- Andrade, H. L. (2019). A critical review of research on student self-assessment. *Front. Educ.* 4:87. doi: 10.3389/feduc.2019.00087
- Arnold, J. (2000). *La dimensión afectiva en la enseñanza de idiomas*. Cambridge: Cambridge University Press.
- Baddeley, A. D. (2000). The episodic buffer. A new component of working memory? *Trends in Cognitive Sciences* 4, 417–423.
- Barrett, N. E., Liu, G. Z., and Wang, H. C. (2021). Student perceptions of a mobile learning application for English Oral presentations: the case of EOPA. *Comput. Assist. Lang. Learn.* 1–26. doi: 10.1080/09588221.2021.1881975
- Bruner, J. S. (2001). *El proceso mental en el aprendizaje*. Madrid: Ediciones NARCEA.
- Cahyono, A. N., and Ludwig, M. (2017). Examining motivation in mobile app-supported math trail environments. 10th congress of the European-society-for-research-in-mathematics-education (CERME). Dublin. February 01–05.
- Calatayud, M. A. (2002). La cultura autoevaluativa, piedra filosofal de la calidad en educación. *Educadores* 204, 357–375.
- Carroll, J. B. (1981). “Veinticinco años de investigación sobre la aptitud en lenguas extranjeras,” in *Individual differences and universals in language learning aptitude*. ed. K. C. Diller (Rowley, MA: Newbury House Publishers), 83–118.
- Castro, R. (1996). “En busca del significado: supuestos, alcances y limitaciones del análisis cualitativo,” in *Para comprender la subjetividad, Investigación cualitativa en salud reproductiva y sexualidad*. eds. I. En Szasz and S. Lerner (México: El Colegio de México).
- Chao, L., Huang, C., Huang, S., Kao, T., and Lee, H. (2009). *Interactive learning system, comprises user interface with that users can learn subject materials more efficiently, where intelligent user interface records users learning behaviors utilized for generating next personalized materials*. National Taiwan University of Science and Technology. Taipei City Taiwan Patent N° 2009-S43833. Taipei City: Taiwan (ROC).
- Charles, C. (2009). Studying language progressing method, involves selecting specific contents for learning from the contents database, providing multiple language study contents to the server, and progressing studying language through image solution by user. Patent assignee ELAMP CO LTD (ELAM-non-standard). Derwent primary accession number 2009-N78622.
- Chie Fang, J. W., Chin Yeh, H., Bo Ru, L., and Chen, N. S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL* 33, 71–87. doi: 10.1017/S0958344020000208

Funding

This study received funding from Aprendizaje inmersivo S.L. (2021). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. All authors declare no other competing interests.

Acknowledgments

Both authors thank the fellow visiting research at the University of Huelva during the second semester of 2021. This fellow visiting facilitated the construction of this article and the scientific transfer of research results between the University of Huelva and the company Aprendizaje Inmersivo S.L. (2021).

Conflict of interest

This study received funding from Aprendizaje inmersivo S.L. The funder had the following involvement with the study: funding to promote scientific research. All authors declare no other competing interests.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Chih-Ming, C., Liang-Chun, C., and Shun-Min, Y. (2019). An English vocabulary learning app with self-regulated learning mechanisms to improve learning performance and motivation. *Comput. Assist. Lang. Learn.* 32, 237–260. doi: 10.1080/09588221.2018.1485708
- Chin-Hsuan, Y., Yi-Shun, W., Yu-Min, W., and Tin-Jun, L. (2021). Drivers of mobile learning app usage: an integrated perspective of personality, readiness, and motivation. *Interact. Learn. Environ.* 1–18. doi: 10.1080/10494820.2021.1937658
- Chu-Chu, S., Papa, C., and Tsai Feng, C. (2013). Establishment of a Mobile-assisted language teaching model for English. Second international conference on robot, vision and signal processing (RVSP, December 10–12).
- David, P., Ausubel, D., Novak, J., and Hanesian, H. (1991). *Psicología Educativa. Un punto de vista cognoscitivo*. México: Editorial Trillas.
- Deci, E. L., and Ryan, R. M. (2000). The “what” and “why” of goal pursuits: human needs and the self-determination of behavior. *Psychol. Inq.* 11, 227–268. doi: 10.1207/S15327965PLI1104_01
- Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., and Noroozi, O. (2021). Using gamification to support learning English as a second language: a systematic review. *Comput. Assist. Lang. Learn.* 34, 934–957. doi: 10.1080/09588221.2019.1648298
- Ellis, N. C. (1996). Memoria de trabajo en la adquisición de vocabulario y sintaxis: poner el lenguaje en orden. *Quar. J. Exp. Psychol. Sección A*. 49, 234–250. doi: 10.1080/713755604
- Fernández, S. (2011). La autoevaluación como estrategia de aprendizaje. Marco ELE.
- Fernández-Novell, J. M., and Zaragoza Domeneq, C. (2014). “Importancia de la autoevaluación,” in *INNODOCT/14. Strategies for education in a new context*. eds. F. J. G. Simón, M. P. Ortiz, I. G. Pechuán, S. E. Miguel, J. O. M. Andrés, C. R. Armengot et al. (Valencia: Editorial Universitat Politècnica de Valencia).
- Fernando Calderero, J. F., Aguirre, A. M., Castellanos Sánchez, A., and Peris Sirvent, R. M. (2014). Una nueva aproximación al concepto de educación personalizada y su relación con las TIC. Teoría de la Educación. Educación y Cultura en la Sociedad de la Información. 15, 131–151. doi: 10.14201/eks.11890
- García Botero, G., Botero Restrepo, M. A., Zhu, C., and Questier, F. (2021). Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Comput. Assist. Lang. Learn.* 34, 1013–1039. doi: 10.1080/09588221.2019.1650780
- García Botero, G., Questier, F., and Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: do students walk the talk? *Comput. Assist. Lang. Learn.* 32, 71–97. doi: 10.1080/09588221.2018.1485707
- García Naranjo, M. A. (2018). Methodology of autoevaluation of the quality of the graduate in education. *Case study. ATENAS*. 1, 34–50.
- García Sanz, E. (2020). Assessment of the Own progress in oral communication skills vs. the importance given to self-assessment. *Criteria. Lenguaje* 48, 38–59. doi: 10.25100/lenguaje.v48i1.6571
- García-Aretio, L. (2016). El juego y otros principios pedagógicos. Supervivencia en la educación a distancia y virtual. *Revista Iberoamericana de Educación a Distancia*. 19, 9–23. doi: 10.5944/ried.19.2.16175
- García-Hoz, V. (1988). *Educación personalizada*. Madrid: Rialp.
- Han, Y. W. (2015). Gongzuo Jiyou Rongliang dui Zhongguo Yingyu Xuexizhe Kouyu Zhunquedu, Fuzadu y Liulidu de Yingxiang. [the influence of working memory capacity on oral accuracy, complexity and fluency of Chinese English learners]. *Waiyu Jiaoxue*. 36, 65–68.
- Han, Y. W., Cui, Y. Q., and Tang, Y. M. (2017). Gongzuo Jiyou Rongliang he Renwu Pinci dui Zhongguo Yingyu Xuexizhe Kouyu Chanchu de Yingxiang. [effects of working memory capacity and task frequency on Chinese English learners' oral output]. *Xiandai Waiyu*. 2, 90–98+149–150.
- Hernández Ortega, J., Sánchez García, P., and Rovira Collado, J. (2021). New transmedia ecosystems for the promotion of reading in the era of booktubers. *Investigaciones Sobre Lectura* 15, 23–45. doi: 10.24310/isl.vi15.12563
- Huang, T., Loerts, H., and Steinkrauss, R. (2022). The impact of second- and third-language learning on language aptitude and working memory. *Int. J. Biling. Educ.* 25, 522–538. doi: 10.1080/13670050.2019.1703894
- Jen-Jiun, L., and Huifen, L. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Jeno, L. M., Dettweiler, U., and Grytnes, J. A. (2020). The effects of a goal-framing and need-supportive app on undergraduates' intentions, effort, and achievement in mobile science learning. *Comput. Educ.* 159:104022. doi: 10.1016/j.compedu.2020.104022
- Jiang, Y. Z., Xie, W. Y., and Meng, X. Z. (2015). Hanyu Pudonghua Ertong Yingyu Cihui yudu de Yingxiang Yinsu. [factors affecting English vocabulary reading of the children who speak Chinese mandarin]. *Waiyu Jiaoxue yu Yanjiu*. 47, 405–416+481.
- Jin, N., Zili, C., and Wan, J. (2019). TI designing a Mobile-assisted English Reading class for college students. 18th Annual International Conference on Web-based Learning (ICWL). September 23–25, 2019
- Kao, K. C., Yen, J. Y., Yen, M. H., Kao, K., Yen, J., and Yen, M. (2009). *Evaluation learning device, has evaluation module developing evaluation according to lessons and produces learning information after user provides answer to judge learning progress of user*. Patent. Patent assignee MICRO-STAR INT CO LTD (MSIC-C). Derwent primary accession number 2009-A84391.
- Khafaga, A., and Alghawli, A. S. A. (2021). The impact of CALL software on the performance of EFL students in the Saudi university context. *Int. J. Adv. Comput. Sci. Appl.* 12, 304–312. doi: 10.14569/ijacsa.2021.0120734
- Kirsch, C. (2016). Using storytelling to teach vocabulary in language lessons: does it work? *The. Lang. Learn. J* 44, 33–51. doi: 10.1080/09571736.2012.733404
- Kissling, E. M., and O'Donnell, M. E. (2015). Increasing language awareness and self-efficacy of FL students using self-assessment and the ACTFL proficiency guidelines. *Lang. Aware*. 24, 283–302. doi: 10.1080/09658416.2015.1099659
- Lan, K. Y., Fang, W. C., Hsiao, I., and Chen, N. S. (2018). Real body versus 3D avatar: the effects of different embodied learning types on EFL listening comprehension. *Educ. Technol. Res. Dev.* 66, 709–731. doi: 10.1007/s11423-018-9569-y
- Li, H. (2004). Dier Yuyan Yuyi Tiqu zhong de Gongzuo Jiyou Xiaoying. [Working memory effect in second language semantic extraction]. *Waiyu Xuekan* 94–98.
- Lin, J. J., and Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Comput. Assist. Lang. Learn.* 32, 878–919. doi: 10.1080/09588221.2018.1541359
- Liou, H. C. (2016). “CALL tools for reading and writing,” in *The Routledge Handbook of Language Learning and Technology*. eds. F. Farr and L. Murray (London: Routledge).
- Luef, E. M., Ghebru, B., and Ilon, L. (2020). Apps for language learning: their use across different languages in a Korean context. *Interact. Learn. Environ.* 28, 1036–1047. doi: 10.1080/10494820.2018.1558255
- McGregor, I., and Holmes, J. G. (1999). How storytelling shapes memory and impressions of relationship events over time. *J. Pers. Soc. Psychol.* 76, 403–419. doi: 10.1037/0022-3514.76.3.403
- Moreno-Ibañez, A., Jordano de la Torre, M., and Vermeulen, A. (2016). VISP design and evaluation, a mobile application to practise oral competence. *Revista Iberoamericana de Educación a Distancia*. 19, 63–81. doi: 10.5944/ried.19.1.14580
- Ni, J. C. (2018). Eryu Cihui Nengli yu Cihui Gongzuo Jiyou Rongliang jiqi Baochiliang Guanxi Yanjiu. [A study on the relationship between second language vocabulary ability, vocabulary working memory capacity and its retention]. *Jiefangjun Waiguoyu Xueyuan Xuebao*. 41, 100–107.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle Publishers.
- Parra-Sánchez, J. (2022). Potencialidades de la Inteligencia Artificial en Educación Superior: Un enfoque desde la personalización. *Revista Tecnológica-Educativa Docentes 2.0* 14, 19–27. doi: 10.37843/rtd.v14i1.296
- Peña-Acuña, B., and Crismán-Pérez, R. (2021). Validación del cuestionario acerca de actitudes lingüísticas hacia una lengua extranjera debido al uso de la app Papua. *REIDOCREA*. 10, 1–13. doi: 10.30827/Digibug.69761
- Pérez Ariza, K., and Hernández Sánchez, J. E. (2014). Aprendizaje y comprensión. Una mirada desde las humanidades. *Humanidades Médicas*. 14, 699–709.
- Piaget, J. (1972). *La Construcción de lo real en el niño*. Madrid: Aguilar.
- Rezaee, A., Seyyed Mohamed, A., and Razzaghifard, P. (2019). The impact of mobile-based dynamic assessment on improving EFL oral accuracy. *Educ. Inf. Technol.* 24, 3091–3105. doi: 10.1007/s10639-019-09915-1
- Rezende, C. (2016). The role of storytelling in language learning: a literature review Working papers of the linguistics circle of the University of Victoria 26. 24–44.
- Rodríguez Gómez, G., Ibarra Sáiz, M. S., and Gómez Ruiz, M. A. (2009). Autoevaluación en la universidad: un reto para profesores y estudiantes. *Revista de Educación*. 357, 401–430. doi: 10.4438/1988-592X-RE-2010-356-045
- Rosell-Aguilar, F. (2018). Autonomous language learning through a mobile application: a user evaluation of the busuu app. *Comput. Assist. Lang. Learn.* 31, 854–881. doi: 10.1080/09588221.2018.1456465
- Santamaría, A., and Alcalde, E. (2021). Multimodal discourse in digital storytelling: an assessment tool proposal. *Comp. Assis.Lang. Learn. Elec.c J* 22, 14–25.
- Serafini, E. J., and Sanz, C. (2015). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Stud. Second. Lang. Acquis.* 38, 607–646. doi: 10.1017/S0272263115000327
- Shantanu Tilak, M. S., Kuznetcova, I., Martens, B., and Akinkuolie, B. (2021). Gamification in mobile-assisted language learning: a systematic review of Duolingo

literature from public release of 2012 to early 2020. *Comput. Assist. Lang. Learn.*, 1–38. doi: 10.1080/09588221.2021.1933540

Shen, W., and Park, H. (2020). Working memory and second language learning: A review of the past twenty years' research in China. *J. Pan-Pac. Assoc. Appl. Lingu* 24, 85–106. doi: 10.25256/PAAL.24.1.5

Skehan, P. (2015). "Working Memory and Second Language Performance: A Commentary". In *Working Memory in Second Language Acquisition and Processing*. eds. Z. Wen, M. Borges Mota and A. McNeill (Bristol: Blue Ridge Summit: Multilingual Matters), 189–202.

Thorndike, E. L. (1931). *Human learning*. New York: Appleton Century.

Tracy, S. (2021). Calidad cualitativa: ocho pilares para una investigación cualitativa de calidad. *Márgenes, Revista de Educación de la Universidad de Málaga*. 2, 173–201. doi: 10.24310/mgnmar.v2i2.12937

Verhagen, J., Leseman, P., and Messer, M. (2015). Phonological memory and the acquisition of grammar in child L2 learners. *Lang. Learn.* 65, 417–448. doi: 10.1111/lang.12101

Vigotski, L. S. (1987). *Historia del desarrollo de las funciones psíquicas superiores*. La Habana: Editorial Científico-Técnica.

Wang, H. Z. (2019). *Xuexizhe Yinsu yu Yingyu Kouyu Chengji Guanxi de Jiegou Fangcheng Moxing Yanjiu*. [A structural equation model study on the relationship between learner factors and oral English performance]. *Waiyu Jiaoxue Lilun yu Shijian*, vol. 4, 17–26.

Wang, C., Zhu, S., and Ma, B. (2021). A study on English learning strategies of university students in Hong Kong. *Asian Englishes*. 1–15. doi: 10.1080/13488678.2021.1945182

Yeh, S. W., and Chen, C. T. (2019). EFL Learners' peer negotiations and attitudes in Mobile-assisted collaborative writing. *Lang. Educ. Assess.* 2, 41–56. doi: 10.29140/lea.v2n1.100

Yi, B. S., and Luo, S. Q. (2012). Gongzuo Jiye Rongliang dui Eryu Xuexizhe Shumianyu Chanchu de Yingxiang. [the influence of working memory capacity on second language learners' written language production]. *Waiyu Jiaoxue yu Yanjiu*. 44:536–546+640.

Yi, B. S., and Ni, C. B. (2015). Renwu Zhiling he Gongzuo Jiye dui Eryu Xuexizhe Shumianyu Chanchu de Yingxiang. [the effect of task instructions and working memory on the output of written language by second language learners]. *Waiyu yu Waiyu Jiaoxue* 1, 56–61.

Zhang, X., and Zuber, S. (2020). Los efectos del lenguaje y la repetición semántica en el efecto de promulgación de la memoria de acción. *Parte delantera. Psicol.* 11:515. doi: 10.3389/fpsyg.2020.00515

Zhou, D., Xu, S., and Fan, Y. (2016). Research on hot research topics on personalized learning and its development trend in China: Based on co-word visualization analysis from CSSCI Proceedings of the 2016 International Conference on Modern Management, Education Technology, and Social Science (MMETSS 2016). *Advances in Social Science Education and Humanities Research*. 91.21–31.

Appendix

Quantitative questionnaire on language attitudes and learning attitudes towards the Papua tool.

1. Do you think that the activities included in the app boost fluency in informal communication as a result of encouraging interaction in real, contextualized social situations?

Not really To a limited extent Reasonably well Quite a lot Very much so.

2. Do you think that the activities included in the app boost oral comprehension in informal communication as a result of encouraging interaction in real, contextualized social situations?

Not really To a limited extent Reasonably well Quite a lot Very much so.

3. Do you think that the activities included in the app boost self-confidence in speaking English in informal communication as a result of encouraging interaction in real, contextualized social situations?

Not really To a limited extent Reasonably well Quite a lot Very much so.

4. Do you think that the activities included in the app improve reading comprehension in foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.

5. Do you think that the reading activities included in the app improve the memorization of words?

Not really To a limited extent Reasonably well Quite a lot Very much so.

6. Do you think that the activities included in the app improve users' morpho-syntactic knowledge of foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.

7. Do you think that the actors' emotional performances in the real-life scenes included in the storyline are influential in improving users' foreign language vocabulary?

Not really To a limited extent Reasonably well Quite a lot Very much so.

8. Do you think that taking a role which involves expressing your emotions helps you to memorize foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.

9. Do you think that taking a role which involves expressing your emotions improves your assimilation of foreign language vocabulary?

Not really To a limited extent Reasonably well Quite a lot Very much so.

10. Do you think that activities involving repetition of material or a spiral syllabus improve memory capacity?

Not really To a limited extent Reasonably well Quite a lot Very much so.

11. Do you think that self-evaluation in terms of average scores in pronunciation and answers for each scene boosts users' intrinsic motivation for learning foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.

12. Do you think that self-evaluation in terms of wrong content-based answers in each scene boosts users' intrinsic motivation for learning foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.

13. Do you think that self-evaluation in terms of wrong socio-cultural answers in each scene boosts users' intrinsic motivation for learning foreign language?

Not really To a limited extent Reasonably well Quite a lot Very much so.



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Petr Slechta,
University of Ostrava, Czechia
Paula López,
Universidad de León, Spain

*CORRESPONDENCE

São Luís Castro
slcastro@fpce.up.pt

SPECIALTY SECTION

This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 29 July 2022

ACCEPTED 11 October 2022

PUBLISHED 14 November 2022

CITATION

Nunes A, Limpo T and Castro SL (2022) Predictors of Portuguese teachers' use of Information and Communication Technologies in literacy classes. *Front. Psychol.* 13:1006713. doi: 10.3389/fpsyg.2022.1006713

COPYRIGHT

© 2022 Nunes, Limpo and Castro. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Predictors of Portuguese teachers' use of Information and Communication Technologies in literacy classes

Andreia Nunes, Teresa Limpo and São Luís Castro*

Center for Psychology at University of Porto, Faculty of Psychology and Education Sciences of the University of Porto, Porto, Portugal

In the last years, the teaching and learning of literacy has changed due to the development of Information and Communication Technologies (ICT). The use of ICT in the classroom depends largely on teachers, who are the key players in its integration. However, several factors influence teachers' decisions to use ICT in their classroom, both internal (e.g., self-efficacy) and external (e.g., school support). Indeed, despite the potential benefits of using ICT, not all teachers use them in their teaching practice. In the present study, we examined which are the main factors influencing teachers' effective use of ICT in literacy classrooms. A total of 125 teachers lecturing Portuguese Language in grades 5–12 participated in this study ($M=50.00$ years, $SD=7.88$; 89% women). Teachers filled in an online survey, comprising sociodemographic questions (*viz.*, age, gender, education, years of teaching experience, teaching level, school type, and geographical area) and four questionnaires related to ICT and teaching. Results showed that effective use of ICT was predicted by both internal (ICT' self-efficacy and constructivist conception of teaching) and external (lack of access and support, and gatekeepers) factors. These findings may help in the identification of key targets to facilitate the effective use of ICT in literacy classrooms.

KEYWORDS

literacy, education, classroom, teachers, information and communication technology, effective use

Introduction

Literacy refers to reading and writing skills and is known to greatly influence the knowledge-building process in the classroom (Iinuma, 2016). However, in the last years, the traditional scope of literacy means – including books, magazines, newspapers, or pen-and-paper writing – was expanded due to the development of different modes of communication provided by Information and Communication Technologies (ICT). The development of ICT led to major changes in the teaching and learning of literacy skills. ICT can be used in the classroom as a supportive tool to deliver traditional instruction (e.g., teachers using PowerPoint presentations) or as a

pedagogical means to enable student-centered activities, engaging them in constructive and high-order critical thinking (e.g., students using expert systems in their schoolwork; Jonassen, 1996; Sutherland et al., 2009; Ertmer and Ottenbreit-Leftwich, 2013). As students are becoming more and more comfortable with ICT, its use in the teaching of reading and writing has been capturing researchers' attention. Despite some advantages of digital methods, such as facilitating reading comprehension (e.g., Mangen et al., 2013) or word memorization (e.g., Mangen et al., 2015), digital methods are recognized to be particularly useful in the teaching of literacy. For example, recent studies found supporting evidence for the use of automated systems in the classroom to enhance reading comprehension (Wijekumar et al., 2017) and writing quality (Nunes et al., 2021).

In line with these benefits of ICT, the European Commission (EC) has recently recognized the importance of ICT in education, by publishing a digital education plan aimed to boost the use of ICT for teaching and learning in schools in the European Union (EU; European Commission, 2018). Still, in the same report, the EC identified several barriers to that end, such as lack of physical conditions (e.g., connection to Internet) in schools in the EU, as well as lack of competences and confidence from teachers to use digital tools to support their teaching. Indeed, introducing ICT into classrooms has been difficult due to several factors, such as government policy failure, inadequate funding, lack of educational vision or the apathy and resistance of teachers (Jewitt, 2006). Teachers are often uncomfortable in using ICT as it implies major changes in their current practice, thereby undermining their confidence to effectively use it in the classroom (Robertson and Dale, 2009; Triggs and Sutherland, 2009).

To know the factors influencing ICT use in education is crucial to successfully integrate ICT into literacy instruction and promote their effective use (Riasati et al., 2012). The effective use of ICT has been studied over the years, resulting in several acceptance models, such as the Technology Acceptance Model (Davis, 1989; Venkatesh and Davis, 2000) and the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2012, 2016). According to these models, the benefits of using ICT, ease of use, social influence, and technical support are key determinants that may affect the acceptance and effective use of ICT (Huang and Kao, 2015; Williams et al., 2015). Although these acceptance models provided valuable information on ICT determinants, they are based on general approaches to ICT acceptance and use across educational, organizational, and health settings, that may fail to grasp the specificities of ICT use in the practice. As a consequence, researchers have sought to identify the specific factors that may hinder teachers' integration of ICT in their classrooms. Ertmer (1999) described two types of factors: internal factors, which are rooted in teachers' underlying beliefs about teaching and learning (e.g., beliefs and attitudes about the use of ICT), and external factors, which refer to

school-related variables, extrinsic to teachers (e.g., the availability of resources).

Internal factors, namely, teachers' conceptions of teaching and learning as well as their self-efficacy and attitudes toward ICT, are often perceived as the main factors influencing teachers' ICT integration in their classroom (Drent and Meelissen, 2008; Blackwell et al., 2013). Teachers' conceptions about their preferred ways of teaching and learning can be classified in two approaches, based on distinct methods of teaching: teachers with constructivist conceptions tend to use teaching strategies focused on students as active knowledge makers, whereas those with traditionalist conceptions tend to use conventional lecture methods for teaching and perceive students as passive knowledge makers. Previous research has linked these conceptions with ICT use. It seems that teachers with constructivist conceptions tend to be more open to ICT and use it more often in their teaching practice than teachers with traditionalist views (Ertmer, 2005; Palak and Walls, 2009; Ertmer et al., 2012; Teo and Zhou, 2016). Additionally, teachers with a constructivist conception tend to use ICT in their classroom in a more innovative way (Drent and Meelissen, 2008).

Self-efficacy refers to the beliefs people have in their capabilities to complete a task (Bandura, 1997). As stated by Bandura (1995), these beliefs have a strong influence on human behavior, for example, by affecting the effort put on tasks and the level of accomplishment. The power of self-efficacy beliefs affects the use of ICT. Studies showed that teachers with stronger self-perceptions of ability in using a specific teaching tool reported more willingness to use ICT as well as using it more in the classroom (Pajares, 1992; Liaw et al., 2007; Buabeng-Andoh, 2012).

Two major sets of teachers' attitudes toward ICT involve their perceived value of ICT and their perceived preparedness for using and integrating ICT into the classroom (Blackwell et al., 2013). Negative attitudes toward the value of ICT and disbelief in its pedagogical value has been associated with an ineffective use of ICT in the classroom, even in schools with good resources (Palak and Walls, 2009; Ertmer et al., 2012). Conversely, teachers with positive and strong beliefs in the value of ICT were able to overcome external barriers, such as lack of resources, and use these pedagogical tools effectively (Ertmer et al., 2012). Teachers who believed that ICT can benefit children's learning also reported more use of ICT (Blackwell et al., 2013). Teachers' perceptions of preparedness for integrating technology into the classroom also influence ICT use (Blackwell et al., 2013). In particular, having enough time to learn how to use ICT and to successfully train it in the classroom showed up as key factors for teachers to integrate ICT into their practice (Ertmer et al., 2012; Johnson et al., 2016).

External factors influencing ICT integration in the classroom are the availability of ICT in schools and the support from key groups of people, such as school administrators and parents (Blackwell et al., 2013). Early studies of ICT integration focused on the availability of ICT tools, such as computers, in schools

(Fisher et al., 1996; Norris et al., 2003). This is a basic requirement for the effective use of ICT in education (Johnson et al., 2016), with some studies relating the frequency of computer use with the number of computers available in the classroom (Norris et al., 2003; Inan and Lowther, 2010a). Even though teachers have a central role in using technology for instructional purposes, other school players – known as gatekeepers – were found to influence teachers' integration of ICT. For example, school support from the administration was associated with innovative uses of ICT by teachers (Drent and Meelissen, 2008). Moreover, when parents perceive ICT to be challenging and beyond their skills, they seem to be less engaged in children's learning and, consequently, reduce their engagement with ICT to support learning (Osorio-Saez et al., 2021).

Overall, then, several internal and external factors seem to be relevant to successfully integrate ICT into classroom. However, the majority of evidence came from studies focused on university students and L2 teaching, and failed to address specific school subjects, such as L1 (Xu et al., 2019). The use of ICT provides a great support to mother-tongue classes, by helping teachers to center the teaching and learning process on the student rather than on themselves as well as by motivating learners to reading and writing activities (Riasati et al., 2012). For example, writing through ICT is more attractive to students than pen-and-paper writing, as ICT's writing is more colorful and flashing and can be integrated with sound and image (Matthewman, 2009).

In the present study, our key question was: Which are the main factors influencing teachers' effective use of ICT in Grades 5–12 literacy classrooms? We focused on Grades 5–12 because we wanted to specifically target teachers responsible for Portuguese Language classes. These only exist from Grade 5 onwards, since before that one teacher is responsible for all the contents (language, maths, social studies, etc.). Also, previous studies reported more use of ICT by teachers lecturing in upper than lower grades, due to the increasing complexity in the curriculum (e.g., Gorder, 2008). To answer our research question, we examined the degree to which internal and external factors predict Portuguese teachers' effective use of ICT, after controlling for sociodemographic factors, such as age, teaching experience, and gender, which have been related to ICT use (Venkatesh et al., 2003; Magsamen-Conrad et al., 2015; Parameswaran et al., 2015). In comparison with younger, less experienced teachers (the so-called “digital natives”), older teachers, with more teaching experience, are less familiar and thus less willing to use ICT in their classrooms (Prensky, 2001; Foutsitzi and Caridakis, 2019). Female teachers have also been found to have less confidence and experience than males in the use of ICT in teaching (Zhou and Xu, 2007), even though such difference was not observed in another study (Scherer and Siddiq, 2015). After controlling for these sociodemographic factors, we expected that teacher-related internal factors (*viz.*, conceptions of teaching and learning, ICT' self-efficacy, perceived value of students' education, teacher-related

constraints), and external factors (*viz.*, lack of access and support, and gatekeepers) would influence the use of ICT (Blackwell et al., 2016; Teo and Zhou, 2016).

Materials and methods

Participants

A total of 179 teachers lecturing Portuguese Language in grades 5–12 participated in this study. In the Portuguese teaching system, grades 5–6 correspond to elementary grades (10–12 years old), Grades 7–9 to middle grades (12–15 years old) and 10–12 to high-school (15–18 years old). Fifty-four participants were excluded as they did not finish the full online survey. The final sample included 125 Portuguese Language teachers aged between 25 and 65 years ($M = 50.00$ years, $SD = 7.88$; 88.80% women). As displayed in Table 1, the majority of these teachers lectured in middle-grade years (55.20% in total; 33.60% were exclusively middle-grade teachers), had a Bachelor degree (73.60%), and taught in public schools (92.80%) in the North of Portugal (55.20%). On average, they had a teaching experience of 24.72 years ($SD = 9.73$), ranging between one and 45 years. The study was approved by the Ethics Committee of the authors' institution, and informed consent was obtained from all participants.

TABLE 1 Teachers' characteristics.

Teachers' characteristics	N	%
Gender	125	100
Male	14	11.20
Female	111	88.80
Education		
Bachelor degree	92	73.60
Master degree	28	22.40
Ph.D. degree	5	4.00
Teaching level		
Elementary School (Grades 5–6)	32	25.60
Middle School (Grades 7–9)	42	33.60
High School (Grades 10–12)	24	19.20
Elementary + Middle + High School	2	1.60
Middle + High School	19	15.20
Elementary + Middle School	6	4.80
School type		
Public	116	92.80
Private	7	5.60
Public/private	2	1.60
Geographical area		
North	69	55.20
Center	33	26.40
South	21	16.80
Madeira Isle	2	1.60

Materials and procedure

The study was conducted online with the Qualtrics software. The link to participate was disseminated through social networks and email, and was available between September 2019 and December 2020. Participants were asked to respond to sociodemographic questions and five questionnaires related to ICT and teaching. Because these questionnaires were originally published in English, they were previously translated to Portuguese by two Portuguese native speakers fluent in English. After discussion, one version of each questionnaire was achieved and presented to 10 Portuguese Language teachers, who provided us the feedback to achieve the final versions (Supplementary Table S1).

Conceptions of teaching and learning

We used Teo and Zhou's (2016) two-dimensional scale of measuring traditionalist (five items; e.g., "Teaching is simply telling, presenting, or explaining the subject matter"; $\alpha=0.79$), and constructivist conceptions (e.g., "Students should be given many opportunities to express their ideas"; $\alpha=0.79$). This questionnaire aims to measure the degree to which assume to be the primary source of knowledge and students are the passive receivers (i.e., traditionalist view), and put students into the center of learning seeing them as active knowledge makers (constructivist views). Answers were given in a scale ranging from 1 (*totally disagree*) to 7 (*totally agree*).

Self-efficacy toward the use of unfamiliar ICT

We used the computer self-efficacy scale (10 items; e.g., "I could use the new technology if I had never used a product like it before"; $\alpha=0.88$; Laver et al., 2012; adapted from Compeau and Higgins, 1995). The original computer self-efficacy scale was modified to be able to be used by older people and people with disabilities, and also to cover a broader range of different ICT. Its main goal is to measure people's confidence about the use of a new and unfamiliar ICT. Teachers were asked to imagine that they received a new ICT they had never used before, intended to facilitate teaching and learning of their subject. Answers were given in a scale ranging from 1 (*not confident at all*) to 5 (*very confident*).

Attitudes toward ICT

The attitudes toward ICT measured internal and external factors regarding ICT integration (Blackwell et al., 2013). Internal factors included two dimensions: one about the value of ICT for teachers, specifically the perceived value of ICT for students' education, where items described how ICT could be useful to children's cognitive and social development (five items; e.g., "Technology can improve individualized learning"; $\alpha=0.78$); and other about internal barriers to ICT integration, specifically the perceived teacher-related constraints, where items described teachers' lack of preparation for integrating ICT into classroom, including lack of training, time to learn or comfort with ICT (five items; e.g., "Technology use is limited by insufficient or lack of training"; $\alpha=0.86$). External factors included two external barriers to ICT integration: the perceived lack

of access and support, where items described how access to ICT and school' support limit teachers' use of ICT in the classroom (three items; e.g., "Technology use is limited by insufficient or inadequate software"; $\alpha=0.74$); and the perceived gatekeepers, which described two major groups of people with power to limit ICT integration into classroom, namely parents and school administration (two items; e.g., "Technology use is limited by my school's policy" $\alpha=0.81$). Answers were given in a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Effective use of ICT in the classroom

Following the previous work conducted by Blackwell et al. (2014, 2016), teachers were presented with nine different ICT that could be used in the classrooms for instructional purposes (*viz.*, TV/DVD, laptop/desktop computer, digital camera or video recorder, interactive whiteboard, smartphone, e-reader, tablet, and assistive technologies). For each ICT, they were asked to indicate how often they used them for instructional purposes, using a scale ranging from 0 (*never*) to 6 (*daily*). The final score was computed by averaging the responses for the nine ICTs.

Data analysis strategy

First, we conducted preliminary analyses to examine descriptive statistics and correlations between all variables. Second, we conducted a stepwise multiple linear regression to examine the contribution of internal and external factors on teachers' effective use of ICT, above and beyond teachers' sociodemographic characteristics. In Step 1 we entered the main effects of age, gender, and teaching experience. In Step 2, we added the and external predictors to ICT integration.

Results

Preliminary analyses: Descriptive statistics and correlations

Table 2 presents means and standard deviations for all predictors and outcome variables, along with the bivariate correlations between them. The correlations can be organized in two main results: (1) Age is only correlated with one factor, specifically with self-efficacy toward the use of unfamiliar ICT ($r=-0.25$); and (2) both internal (i.e., constructivist conception, self-efficacy, teacher-related constraints) and external factors (i.e., lack of access and support) are correlated with teachers' effective use of ICT in the classroom.

Regression analysis: Prediction of effective use of ICT

Step 1 showed no main effects of teachers' characteristics in the effective use of ICT, $R^2=0.03$, $F(3, 121)=1.10$, $p=0.35$. The inclusion of the factors of ICT integration in Step 2 resulted in a

TABLE 2 Means, standard deviations, and correlations for all variables.

Measures	Correlations										
	1	2	3	4	5	6	7	8	9	10	11
1. Age											
2. Gender ^a	−0.01										
3. Teaching experience	0.89***	0.09									
4. Traditionalist conception	−0.04	0.07	0.02								
5. Constructivist conception	−0.06	0.07	−0.06	−0.56***							
6. Self-efficacy toward the use of unfamiliar ICT	−0.25**	0.07	−0.17	−0.13	0.23**						
7. Value for students' education	0.06	0.05	0.05	−0.27**	0.33***	0.13					
8. Teacher-related constraints	0.14	−0.05	0.07	0.28**	−0.31**	−0.50***	−0.15				
9. Lack of access and support	−0.04	−0.04	−0.05	0.25**	−0.10	−0.20*	−0.20*	0.38***			
10. Gatekeepers	−0.16	−0.05	−0.14	0.23**	−0.21*	−0.08	−0.14	0.22*	0.19*		
11. Effective use of ICT	0.002	0.16	0.03	−0.24**	0.30**	0.33***	0.10	−0.28**	−0.28**	−0.03	
<i>M</i>	50.00	0.89	24.72	2.34	6.07	3.84	4.06	2.49	3.60	2.26	1.87
<i>SD</i>	7.88	0.32	9.73	0.82	0.70	0.62	0.06	0.90	0.91	0.71	0.87

^aDummy coded, 0 = male, 1 = female.* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

TABLE 3 Final model with all main effects of age, gender, teaching experience and ICT's and teaching' related variables on teachers' effective use of ICT in their classroom.

Predictors	<i>B</i>	<i>SE</i>	β	<i>t</i>
Age	0.01	0.02	0.11	0.58
Gender	0.38	0.23	0.14	1.67
Teaching experience	0.00	0.02	−0.01	−0.03
Traditionalist conception	−0.10	0.11	−0.09	−0.89
Constructivist conception	0.26	0.13	0.21	1.99*
Self-efficacy toward the use of unfamiliar ICT	0.35	0.14	0.25	2.61*
Value of ICT for students' education	−0.09	0.13	−0.06	−0.66
Teacher-related constraints	−0.04	0.10	−0.04	−0.37
Lack of access and support	−0.20	0.09	−0.20	−2.26*
Gatekeepers	0.22	0.11	0.18	2.03*

 $R^2 = 0.26$, $F(7, 114) = 5.04$, $p < 0.001$. * $p < 0.05$.

significant increase in the prediction of effective use of ICT, $\Delta R^2 = 0.23$, $F_{\text{change}}(7, 114) = 5.04$, $p < 0.001$. Significant predictors were self-efficacy ($\beta = 0.25$, $p = 0.01$), constructivist conceptions ($\beta = 0.21$, $p = 0.05$), lack of access and support ($\beta = -0.20$, $p = 0.03$), and gatekeepers ($\beta = 0.18$, $p = 0.05$). The final model explained 26% of the variance in the effective use of ICT (cf. Table 3 for complete results).

Discussion

Over the years, ICT has been perceived as bringing value to teaching and learning practice (Iinuma, 2016). Consequently, several scholars and organizations have recommended the inclusion of ICT in literacy learning activities (Iinuma, 2016). In the present study, we aimed to determine the role of several internal and external factors influencing teachers' effective use of

ICT in Portuguese Language classes in Grades 5–12. Overall, we found that the effective use of ICT for instructional purposes in Portuguese Language classes were predicted by ICT' self-efficacy and constructivist conceptions (internal factors), as well as by lack of access and support and gatekeepers (external factors). These findings align with prior research showing that both teacher and school-related factors affect the effective use of ICT in the classroom (Drent and Meelissen, 2008; Inan and Lowther, 2010b; Ertmer et al., 2012). In what follows, we discuss in-depth each of the predictors.

As expected, we found that the endorsement of constructivist conceptions of teaching and learning, but not traditionalist ones, was associated with a greater use of ICT in the classrooms. These findings replicate past research (Palak and Walls, 2009; Teo and Zhou, 2016) and gives strength to the claim that current educational practices need to be changed for ICT to be implemented in a pedagogical way. Foutsitzi and Caridakis

(2019) argued that increasing teachers' familiarity with ICT alone is not enough. Instead, the authors proposed that training teachers on the use of ICT along with pedagogical strategies will allow the teaching and learning process to become a student-centric procedure, in which ICT is used to support students' exploration and critical thinking. On the same vein, Triggs and Sutherland (2009) argued that ICT needs to go beyond a passive use, in which students only receive information through ICT, to a more learner-centered approach, in which students autonomously manage their engagement with ICT.

In line with our hypotheses, self-efficacy to use technology also appeared as a relevant predictor of effective use of ICT. Conceptually, self-efficacy is based on beliefs rather than on skills, and these beliefs are responsible for the way people deal with adversity and for their confidence in achieving some goal (Bandura, 1997). In our study, teachers had to imagine they were using an ICT they had never used before, with the goal of facilitating the teaching and learning of their subject. Teachers who believed to be more capable of using this unfamiliar ICT reported more use in classroom. This is aligned with prior research that showed that self-efficacy is a major factor in the decision to use ICT as a teaching tool (Pajares, 1992; Liaw et al., 2007; Buabeng-Andoh, 2012). Self-efficacy as a predictor of using ICT in the classroom is not surprising as perceived self-efficacy affects thought processes, the level of motivation and affective states, and therefore affects behavior (Bandura, 1997).

Besides constructivist views and self-efficacy, only external factors were predictors of ICT use. School-related, external factors were found to predict the use of ICT to support the teaching of literacy. Specifically, the more teachers perceived schools as being a barrier to ICT integration (e.g., lack of technical support or equipment), the less often they used ICT. This finding makes sense to the extent that unless schools offer physical conditions to implement technological solutions in the classrooms, teachers will prefer to use more traditional and familiar methods (Grimes and Warschauer, 2008; Ertmer et al., 2012). Additionally, we also found that the perception that teachers had about parents and school's policy approval of ICT use predicted the use of ICT in the classroom. This finding, which is aligned with past research (Drent and Meelissen, 2008; Inan and Lowther, 2010b), reinforces the need for schools to offer the best possible conditions for teachers to overcome some obstacles to ICT integration. This includes not only to offer adequate equipment, but also to offer technical support when needed, as well as investing in teachers' professional development, providing them with ICT training (Triggs and Sutherland, 2009; Johnson et al., 2016).

Contrary to our hypotheses, internal attitudes toward ICT did not predict its use. On the one hand, the perceived value of ICT for students' education was not related to ICT use, which is not aligned with past research (Blackwell et al., 2016). The different findings may be related to methodological aspects (e.g., school levels, type of ICT, research methods) along with cultural factors. In the future, it would be worthwhile to inspect likely variables moderating the link between teachers' attitudes and ICT use. On

the other hand, teacher-related constraints was another variable that did not predict ICT use. Though contrasting with some studies (e.g., Ertmer et al., 2012), this finding is aligned with the results reported by Blackwell et al. (2013), though a preschool context. The authors suggested this lack of predictive value by teachers' attitudes may be related to a mismatch between their beliefs and the classroom reality. Teachers may use ICT that is available in the classroom, even if they feel restricted in using it due to personal constraints.

Unfortunately, in EU countries including Portugal, ICT is not yet a fully integrated resource in teaching activities (Peralta and Costa, 2007). Studies in the Portuguese context reported that the use of ICT in education is still low and that ICT did not bring a reconfiguration of the educational process, as technology is only being added to traditional teaching methods (Alves and Rodrigues, 2014; Almeida, 2018). This may be troublesome as education do not seem to be fully adjusted to students' reality. As put by Lacina and Griffith (2012): "For children of the 21st century, technology is like oxygen – a necessary component of their life" (p. 316). The more we adjust teaching to students' reality, the more likely we are to succeed in improving their educational and professional outcomes.

The findings reported here should be considered in view of at least four limitations. First, because data were obtained at a single time point and because this study is correlational in nature, causality inferences should be avoided. Additional research is needed to replicate reported results using longitudinal designs. Second, there was a larger representation of women than men in our sample. Future studies should aim to collect larger samples and to add more male teachers to the data. Third, given sample size and because some teachers lectured in several teaching levels, an analysis comparing the different teaching levels was not advisable. Future studies should aim for larger samples to provide further conclusions about the use of ICT at different educational stages. Finally, the study was designed before COVID-19 pandemic and, although data collection started before this period, part of the data collection happened during the pandemic. We do not know the extent to which this circumstance, including the massive use of online learning, affected our findings. Future research should explore the impact of this pandemic in teachers' perceptions and effective use of ICT in the teaching and learning process.

Conclusion

The use of ICT can be a major help to develop of literacy skills. However, as shown here, there are several teacher and school-related factors influencing its use. To unravel these factors is a first step to foster the use of technology in educational settings. This was the main contribution of the present study, which provided indications on key factors to target in order to improve the use of ICT in Portuguese literacy classrooms. To effectively introduce ICT into the classroom schools should invest in teachers' professional development and raise their awareness on the

potentials of technology to enhance teaching and learning. To that end, it seems important to nurture teacher's self-efficacy beliefs, promote constructivist views of teaching and learning, and assure that external conditions to support the use of ICT are in place. In sum, the teaching and learning of literacy would much benefit if teachers saw this process as embedded in the digital age.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Faculty of Psychology and Education Sciences of University of Porto. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AN, TL, and SC designed the study. AN implemented the study, analyzed the data, and wrote the first version of the manuscript, under TL and SC supervision. All authors contributed to the article and approved the submitted version.

References

- Almeida, P. (2018). Tecnologias digitais em sala de aula: O professor e a reconfiguração do processo educativo. *Da Investigação às Práticas* 8, 4–21. doi: 10.25757/invep.v8i1.124
- Alves, N., and Rodrigues, C. (2014). As tecnologias da informação e da comunicação na escola: causas de uma subutilização. *Sociologia Online* 7, 121–139.
- Bandura, A. (1995). Comments on the crusade against the causal efficacy of human thought. *J. Behav. Ther. Exp. Psychiatry* 26, 179–190. doi: 10.1016/0005-7916(95)00034-w
- Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York, NY: Freeman.
- Blackwell, C. K., Lauricella, A. R., and Wartella, E. (2014). Factors influencing digital technology use in early childhood education. *Comput. Educ.* 77, 82–90. doi: 10.1016/j.compedu.2014.04.013
- Blackwell, C. K., Lauricella, A. R., and Wartella, E. (2016). The influence of TPACK contextual factors on early childhood educators' tablet computer use. *Comput. Educ.* 98, 57–69. doi: 10.1016/j.compedu.2016.02.010
- Blackwell, C. K., Lauricella, A. R., Wartella, E., Robb, M., and Schomburg, R. (2013). Adoption and use of technology in early education. *Comput. Educ.* 69, 310–319. doi: 10.1016/j.compedu.2013.07.024
- Buabeng-Andoh, C. (2012). Factors influencing teachers' adoption and integration of information and communication technology into teaching: a review of the literature. *Int. J. Educ. Dev. Using Inf. Commun. Technol.* 8, 136–155.
- Compeau, D. R., and Higgins, C. A. (1995). Computer self-efficacy: development of a measure and initial test. *MIS Q.* 19, 189–211. doi: 10.2307/249688
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 13, 319–340. doi: 10.2307/249008
- Drent, M., and Meelissen, M. (2008). Which factors obstruct or stimulate teacher educators to use ICT innovatively? *Comput. Educ.* 51, 187–199. doi: 10.1016/j.compedu.2007.05.001
- Ertmer, P. A. (1999). Addressing first-and second-order barriers to change: strategies for technology integration. *ETR&D* 47, 47–61. doi: 10.1007/bf02299597
- Ertmer, P. A. (2005). Teacher pedagogical beliefs: the final frontier in our quest for technology integration? *Educ. Dev. Res. Dev.* 53, 25–39. doi: 10.1007/bf02504683
- Ertmer, P. A., and Ottenbreit-Leftwich, A. (2013). Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Comput. Educ.* 64, 175–182. doi: 10.1016/j.compedu.2012.10.008
- Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., and Sendurur, P. (2012). Teacher beliefs and technology integration practices: a critical relationship. *Comput. Educ.* 59, 423–435. doi: 10.1016/j.compedu.2012.02.001
- European Commission (2018). Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions on the digital education action plan. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0022&from=EN>
- Fisher, C., Dwyer, D. C., and Yocam, K. (1996). *Education and Technology: Reflections on Computing in Classrooms*. San Francisco: Jossey-Bass.
- Foutsitzi, S., and Caridakis, G. (2019). "ICT in education: benefits, challenges and new directions," in *Proceedings of the 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, Greece, 1–8.
- Gorder, L. M. (2008). A study of teacher perceptions of instructional technology integration in the classroom. *Delta Pi Epsilon J.* 50, 63–76.
- Grimes, D., and Warschauer, M. (2008). Learning with laptops: a multi-method case study. *J. Educ. Comput. Res.* 38, 305–332. doi: 10.2190/EC.38.3.d

Funding

This work was supported by the Portuguese Foundation for Science and Technology (grant UIDB/00050/2020 awarded to the Center for Psychology at University of Porto; Ph.D. studentship SFRH/BD/139195/2018 awarded to AN).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1006713/full#supplementary-material>

- Huang, C., and Kao, Y. (2015). UTAUT2 based predictions of factors influencing the technology acceptance of phablets by DNP. *Math. Probl. Eng.* 2015, 1–23. doi: 10.1155/2015/603747
- Iinuma, M. (2016). *Learning and Teaching with Technology in the Knowledge Society: New Literacy, Collaboration and Digital Content*. Singapore: Springer.
- Inan, F. A., and Lowther, D. L. (2010a). Factors affecting technology integration in K-12 classrooms: a path model. *Educ. Technol. Res. Dev.* 58, 137–154. doi: 10.1007/s11423-009-9132-y
- Inan, F. A., and Lowther, D. L. (2010b). Laptops in the K-12 classrooms: exploring factors impacting instructional use. *Comput. Educ.* 55, 937–944. doi: 10.1016/j.compedu.2010.04.004
- Jewitt, C. (2006). *Technology, Literacy and Learning: A Multimodal Approach*. London: Routledge.
- Johnson, A. M., Jacovina, M. E., Russell, D. E., and Soto, C. M. (2016). “Challenges and solutions when using technologies in the classroom,” in *Adaptive Educational Technologies for Literacy Instruction*. eds. S. A. Crossley and D. S. McNamara (New York, NY: Taylor & Francis), 13–29.
- Jonassen, D. H. (1996). *Computers in the Classroom: Mindtools for Critical Thinking*. Columbus, OH: Merrill/Prentice Hall.
- Lacina, J., and Griffith, R. (2012). Blogging as a means of crafting writing. *Read. Teach.* 66, 316–320. doi: 10.1002/trtr.01128
- Laver, K., George, S., Ratcliffe, J., and Crotty, M. (2012). Measuring technology self efficacy: reliability and construct validity of a modified computer self efficacy scale in a clinical rehabilitation setting. *Disabil. Rehabil.* 34, 220–227. doi: 10.3109/09638288.2011.593682
- Liaw, S.-S., Huang, H.-M., and Chen, G.-D. (2007). Surveying instructor and learner attitudes toward e-learning. *Comput. Educ.* 49, 1066–1080. doi: 10.1016/j.compedu.2006.01.001
- Magsamen-Conrad, K., Upadhyaya, S., Joa, C. Y., and Dowd, J. (2015). Bridging the divide: using UTAUT to predict multigenerational tablet adoption practices. *Comput. Hum. Behav.* 50, 186–196. doi: 10.1016/j.chb.2015.03.032
- Mangen, A., Anda, L. G., Oxenburgh, G. H., and Brønnick, K. (2015). Handwriting versus keyboard writing: effect on word recall. *J. Writing Res.* 7, 227–247. doi: 10.17239/jowr-2015.07.02.1
- Mangen, A., Walgermo, B. R., and Brønnick, K. (2013). Reading linear texts on paper versus computer screen: effects on reading comprehension. *Int. J. Educ. Res.* 58, 61–68. doi: 10.1016/j.ijer.2012.12.002
- Matthewman, S. (2009). “Discerning literacy,” in *Improving Classroom Learning with ICT*. eds. R. Sutherland, S. Robertson and P. John (United States: Routledge), 115–137.
- Norris, C., Sullivan, T., Poirot, J., and Soloway, E. (2003). No access, no use, no impact: snapshot surveys of educational technology in K-12. *J. Res. Technol. Educ.* 36, 15–27. doi: 10.1080/15391523.2003.10782400
- Nunes, A., Cordeiro, C., Limpo, T., and Castro, S. L. (2021). Effectiveness of automated writing evaluation systems in school settings: a systematic review of studies from 2000 to 2020. *J. Comput. Assist. Learn.* 38, 599–620. doi: 10.1111/jcal.12635
- Osorio-Saez, E. M., Eryilmaz, N., and Sandoval-Hernandez, A. (2021). Parents’ acceptance of educational technology: lessons from around the world. *Front. Psychol.* 12:719430. doi: 10.3389/fpsyg.2021.719430
- Pajares, M. F. (1992). Teachers’ beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62, 307–332. doi: 10.3102/00346543062003307
- Palak, D., and Walls, R. T. (2009). Teachers’ beliefs and technology practices: a mixed-methods approach. *J. Res. Technol. Educ.* 41, 417–441. doi: 10.1080/15391523.2009.10782537
- Parameswaran, S., Kishore, R., and Li, P. (2015). Within-study measurement invariance of the UTAUT instrument: an assessment with user technology engagement variables. *Inf. Manag.* 52, 317–336. doi: 10.1016/j.im.2014.12.007
- Peralta, H., and Costa, F. A. (2007). Teachers’ competence and confidence regarding the use of ICT. *Sisifo. Educ. Sci. J.* 3, 75–84.
- Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon* 9, 1–6. doi: 10.1108/10748120110424816
- Riasati, M. J., Allahyar, N., and Tan, K. (2012). Technology in language education: benefits and barriers. *J. Educ. Pract.* 3, 25–30.
- Robertson, S., and Dale, R. (2009). “‘Aliens in the classroom 2’: when technology meets classroom life,” in *Improving Classroom Learning with ICT*, eds. R. Sutherland, S. Robertson and P. John. (United States: Routledge), 138–157.
- Scherer, R., and Siddiq, F. (2015). Revisiting teachers’ computer self-efficacy: a differentiated view on gender differences. *Comput. Hum. Behav.* 53, 48–57. doi: 10.1016/j.chb.2015.06.038
- Sutherland, R., Robertson, S., and John, P. (2009). *Improving Classroom Learning with ICT*. United States: Routledge.
- Teo, T., and Zhou, M. (2016). The influence of teachers’ conceptions of teaching and learning on their technology acceptance. *Interact. Learn. Environ.* 25, 513–527. doi: 10.1080/10494820.2016.1143844
- Triggs, P., and Sutherland, R. (2009). “A holistic approach to understanding teaching and learning with ICT,” in *Improving Classroom Learning with ICT*. eds. R. Sutherland, S. Robertson and P. John (United States: Routledge).
- Venkatesh, V., and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal studies. *Manag. Sci.* 46, 186–204. doi: 10.1287/mnsc.46.2.186.11926
- Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: toward unified view. *MIS Q.* 27, 425–478. doi: 10.2307/30036540
- Venkatesh, V., Thong, J. Y. L., and Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q.* 36, 157–178. doi: 10.2307/41410412
- Venkatesh, V., Thong, J. L., and Xu, S. (2016). Unified theory of acceptance and use of technology: a synthesis and the road ahead. *J. Assoc. Inf. Syst.* 17, 328–376. doi: 10.17705/1jais.00428
- Wijekumar, K., Meyer, B. J. F., and Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders’ content area reading comprehension. *J. Educ. Psychol.* 109, 741–760. doi: 10.1037/edu0000168
- Williams, M. D., Rana, N. P., and Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): a literature review. *J. Enterp. Inf. Manag.* 28, 443–488. doi: 10.1108/JEIM-09-2014-0088
- Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., Wijekumar, K., Xu, Z., et al. (2019). The effectiveness of educational technology applications on adult English language learners’ writing quality: a meta-analysis. *Comput. Assist. Lang. Learn.* 32, 132–162. doi: 10.1080/09588221.2018.1501069
- Zhou, G., and Xu, J. (2007). Adoption of educational technology: how does gender matter? *Int. J. Teach. Learn. Higher Educ.* 19, 140–153.



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga,
Spain

REVIEWED BY

Misrah Mohamed,
University of West London, United Kingdom
Macarena Becerra,
University of Malaga,
Spain

*CORRESPONDENCE

Junju Wang
wangjunju@sdu.edu.cn

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 05 August 2022

ACCEPTED 27 October 2022

PUBLISHED 23 November 2022

CITATION

Zhou J, Wang S and Wang J (2022)
Investigating high schoolers' L2 writing
anxiety, L2 writing self-efficacy, L2 writing
self-regulated strategies, and L2 writing
engagement: Relationships and mediator.
Front. Psychol. 13:1012407.
doi: 10.3389/fpsyg.2022.1012407

COPYRIGHT

© 2022 Zhou, Wang and Wang. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Investigating high schoolers' L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement: Relationships and mediator

Jie Zhou^{1,2}, Shusheng Wang¹ and Junju Wang^{1*}

¹School of Foreign Languages and Literature, Shandong University, Jinan, China, ²School of Foreign Languages and Literature, Hunan University of Technology, Zhuzhou, China

This study used a structural equation modeling approach to investigate the relationships among L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies and L2 writing engagement, and possible mediators that regulate the effect of individual factors. A questionnaire was administered to 340 Chinese high school students from different parts of the country. The results of the study suggested a negative relationship between L2 writing anxiety and L2 writing self-efficacy, and a direct effect of both on L2 writing engagement. While a higher level of L2 writing self-efficacy indicated a lower level of L2 writing anxiety and more students' L2 writing engagement, L2 writing efficacy had a much stronger direct effect on L2 writing engagement than L2 writing anxiety. In addition, L2 writing self-regulated strategies were found to play a mediating role between L2 writing self-efficacy and L2 writing engagement, but not between L2 writing anxiety and L2 writing engagement. This study helps to understand the interplay of individual factors related to L2 writing and sheds light on promoting English writing abilities of Chinese high school students.

KEYWORDS

high school students, L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, L2 writing engagement

Introduction

Writing is a complex and versatile skill (Han and Hiver, 2018). Given its benefits for individuals to describe events, present information, and exchange opinions (Hayes, 2012), writing plays an important role in facilitating one's success of study, job, and daily life (Graham et al., 2018). In pedagogical contexts, the development of writing ability is crucial for L2 learners since it is a key indicator that reflects on their overall language proficiency level (Teng et al., 2022).

Despite its increasing importance, L2 writing is viewed as a challenge for EFL learners due to its complicated and multifaceted nature, often influenced by the interplay of

individual factors (Teng and Zhang, 2018). When writing, L2 learners often encounter various difficulties involving linguistic knowledge and writing skills which may result in writing anxiety, a lack of self-confidence (Cheng, 2004), and fear of negative evaluation and time limits (Kırmızı and Kırmızı, 2015). To a large extent, learner factors like anxiety and self-efficacy are determinant to students' success in writing activities (Han and Hiver, 2018).

Apart from the negative influence of L2 writing anxiety, L2 writing is also affected by L2 writing self-efficacy in a positive way (Woodrow, 2011), and the interaction between writing anxiety and self-efficacy has an impact on the degree of L2 writing engagement (Zhou and Hiver, 2022). Given the need for self-regulation as a major aspect of self-efficacy (Bandura, 1997), learners who have a good perception of L2 writing self-regulated strategies tend to have a stronger sense of writing self-efficacy and demonstrate better writing performance (Teng and Zhang, 2018). In this sense, L2 writing self-regulated strategies are thought to play a crucial role in regulating emotions and achieving success in L2 writing (Teng et al., 2022).

In the Chinese context, high school students are studying in a fierce competitive environment, under the pressure of passing the National Entrance Examination for colleges and universities (Liu and Wang, 2021). For the English test where more weight is putting on writing performance, students have to work hard to pursue high marks in writing tasks. However, they are often confronted with various difficulties due to their poor knowledge of genres and undesirable language proficiency (Li, 2017), which consequently leads to their negative feelings and lack of self-efficacy. It is thus conducive to investigate the individual factors that affect Chinese high schoolers' L2 writing and explore how these factors are related to one another and what mediates their effect on L2 writing engagement.

Previous studies on how L2 writing anxiety and L2 writing self-efficacy interact with each other were mostly focused on university students (e.g., Pajares, 2003; Woodrow, 2011; Kırmızı and Kırmızı, 2015; Abolhasani et al., 2022). Few studies have examined the interactions among L2 writing anxiety, L2 writing self-efficacy and L2 writing self-regulated strategies, and their influences on L2 writing engagement in the Chinese context, particularly in Chinese high schools. In view of this, this study takes Chinese high school students into consideration and examines the relationships among their L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement. It is hoped that this study provides insights for understanding the individual factors involved in Chinese high schoolers' L2 writing, and help students further improve their well-being of L2 writing.

Literature review

L2 writing anxiety

L2 writing anxiety has been found to be significantly related to L2 writing performance (Cheng, 2002; Tahmouresi and Papi, 2021;

Abolhasani et al., 2022). Consisting of three dimensions of cognitive anxiety, somatic anxiety and avoidance behavior, it refers to "a relatively stable anxiety disposition associated with writing, which involves a variety of dysfunctional thoughts, increased physiological arousal, and maladaptive behaviors" (Cheng, 2004, p.319). In addition, it is regarded as a situation-specific anxiety, since L2 writers might experience anxiety in various contexts, like in exams, in the class, at home, at work or in the community (Woodrow, 2011).

A large body of studies have found that L2 writing anxiety was negatively related to L2 writing self-efficacy (e.g., Pajares, 2003; Woodrow, 2011; Piniel and Csizér, 2013; Kırmızı and Kırmızı, 2015; Abolhasani et al., 2022). For example, a high level of writing anxiety may result in a decrease in writing self-efficacy (e.g., Kırmızı and Kırmızı, 2015), and students with a higher level writing self-efficacy experience a lower level of writing anxiety and have fewer writing avoidance behaviors (e.g., Pajares, 2003).

L2 writing anxiety is closely related to the use of L2 writing self-regulated strategies. Many studies demonstrated that students with low L2 writing anxiety could use more L2 writing self-regulated strategies (e.g., Asmari, 2013; Machida and Dalsky, 2014; Abolhasani et al., 2022; Bailey and Almusharraf, 2022). For example, Machida and Dalsky (2014) found that L2 writing strategies could help less anxious students more than those more anxious students.

In terms of the relationship between L2 writing anxiety and L2 writing engagement, there have been discrepancies among researchers. Whereas some of them (e.g., Tsao et al., 2017) thought that a high level of anxiety would interfere with students' engagement in the writing process, others (e.g., Astrid et al., 2017) argued that many students with low writing anxiety had positive behavior engagement and some students with high writing anxiety had positive emotional engagement in writing activities.

L2 writing self-efficacy

L2 writing self-efficacy refers to L2 writers' confidence and beliefs in their abilities to successfully perform writing tasks (Han and Hiver, 2018). Considered as the most reliable predictor of students' writing performance (Bandura, 1997), it is believed to have a recursive relationship with L2 writing performance (Robinson et al., 2020), which means that L2 writing self-efficacy enhances L2 writing performance, and in turn, nurtures the development of L2 writing self-efficacy (Sun et al., 2021).

Inconsistent research findings exist concerning the relationship between L2 writing self-efficacy and L2 writing self-regulated strategies. Whereas many studies suggested a positive influence of L2 writing self-efficacy on the use of L2 writing self-regulated strategies (e.g., Zimmerman and Risemberg, 1997; Pajares, 2003; Bruning et al., 2013; Teng and Huang, 2019; Sun and Wang, 2020), some found that there is no significant relationship between L2 writing self-efficacy and L2 writing self-regulated strategies (e.g., Graham et al., 2005). This might be associated with the inaccuracies of assessing the capabilities for young children by themselves.

In the meantime, some studies found that L2 writing self-efficacy was positively related to writing engagement (e.g., Usher and Pajares, 2008; Han and Hiver, 2018; Tsao, 2021). For example, Usher and Pajares (2008) found that L2 writing self-efficacy could enhance students' engagement in the writing process. Han and Hiver (2018) pointed out that L2 writing self-efficacy could moderate attention and cognitive engagement, and determine the level of effort that students would put into L2 writing activities. Tsao (2021) also reported that L2 writing self-efficacy was a predictive power to motivate students to engage in different types of written corrective feedback.

L2 writing self-regulated strategies

According to Teng and Zhang (2016), L2 writing self-regulated strategies are students' "deliberate, goal-directed attempts to make writing enjoyable, less challenging, and more effective" (p. 7). In the multidimensional model that they conceptualized, L2 writing self-regulated strategies included cognitive strategies, metacognitive strategies, social-behavioral strategies, and motivational regulation strategies.

Previous studies reported that L2 writing self-regulated strategies play a facilitative role for successful L2 writing (e.g., Teng and Zhang, 2016; Teng and Reynolds, 2019; Teng et al., 2022), and are closely linked to how L2 writers monitor their performance and adjust their tasks to achieve success (Zimmerman, 2001; Han and Hiver, 2018). A meta-analysis by Santangelo et al. (2016) also demonstrated that the use of writing self-regulated strategies contributed significantly to the improvement of students' writing performance.

Previous studies also suggested that L2 writing self-regulated strategies was positively related to L2 writing self-efficacy (Zimmerman and Risemberg, 1997; Ekholm et al., 2015), and was positively associated with L2 writing engagement (Zhou and Hiver, 2022). Csizér and Tankó (2017), for example, found that self-regulated strategies could increase writing self-efficacy and decrease writing anxiety for L2 writers. Teng and Zhang's (2018) study also revealed that L2 writing self-regulated strategies mediated the effect of motivational regulation strategies on L2 writing performance.

L2 writing engagement

L2 writing engagement refers to students' active and productive involvement in writing activities (Reeve et al., 2020). It is conceptualized as a multidimensional construct comprising behavioral, cognitive, emotional facets (Fredricks et al., 2004) and agentive engagement (Reeve and Tseng, 2011).

Some studies demonstrated that L2 writing engagement was positively linked to students' writing performance (e.g., Qi and Lapkin, 2001; Fredricks et al., 2004; Rahimi and Zhang, 2021). Qi and Lapkin (2001), for instance, found that the extensiveness of students' engagement and the quality of notice may lead to improved

writing performance. Rahimi and Zhang's (2021) study suggested that students' experiences with the engaging process-genre approach to writing were found to assist each other in sustaining engagement and achievements in and beyond the classroom.

Few studies have suggested that L2 writing engagement was related to some other psychological factors (e.g., Han and Hiver, 2018; Zhou and Hiver, 2022). It was noted that L2 writing self-regulated strategies functioned as an important predictor to students' L2 writing engagement in the writing class (Zhou and Hiver, 2022).

Research hypotheses

In view of the aforementioned studies, it is quite clear that most of them attempted to investigate the relationship between a certain factor and writing performance, and very few have dealt with two or three individual factors, with scarce attention paid to L2 writing engagement. Despite the abundant findings on L2 writing anxiety, L2 writing self-efficacy, and L2 writing self-regulated strategies, inconsistent findings have been yielded and the relationships among the three factors were largely underexplored. Besides, previous studies have predominantly focused on university students in other countries than the Chinese context, little is known about Chinese students, particularly Chinese high school students.

Given the above limitations, this study aims to focus on Chinese high school students and investigate the relationships among L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement. To achieve such a goal, a hypothesized model was developed (see Figure 1), in which L2 writing anxiety (L2WA) and L2 writing self-efficacy (L2WSE) were the predictors, L2 writing self-regulated strategies (L2WSRSS) a mediator, and L2 writing engagement (L2WE) the dependent variable. It was hypothesized that L2 writing self-regulated strategies can mediate the effects of L2 writing anxiety and L2 writing self-efficacy on L2 writing engagement.

Based on the above model, the following hypotheses are to be tested. H1: L2 writing anxiety is negatively associated with L2 writing self-efficacy; both L2 writing anxiety and L2 writing self-efficacy have a positive direct effect on L2 writing engagement. H2: L2 writing self-efficacy is positively associated with L2 writing self-regulated strategies and L2 writing engagement. H3: L2 writing self-regulated strategies are positively related to L2 writing engagement and mediate the effects of L2 writing anxiety and L2 writing self-efficacy on L2 writing engagement.

Materials and methods

Participants

A total of 340 students from 20 Chinese high schools participated in this study. These students came from both key and regular high schools in eight provinces across the country,

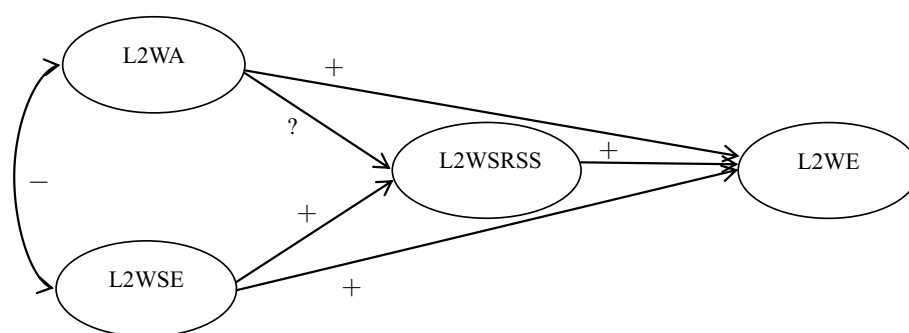


FIGURE 1
A hypothesized model of the relationship among the four variables.

including the eastern, western, southern, and northern regions of China. Among them, 128 were from key high schools and 212 came from regular high schools. Aged from 14 to 20 years old ($M = 16.06$, $SD = 1.45$), this group of students included 157 boys and 183 girls, with 226 of them in their first year, 33 in their second year, and 81 were in their third year. When the survey was conducted, these students had been attending high school English classes from 6 months to 30 months. As the overwhelming majority of their peers, their English studies had been following the same curriculum prescribed by the Chinese Ministry of Education. Each week, they had six class hours for English learning. Quizzes were regularly administered on weekly and monthly basis. Typically, they took one mid-term and one final-term English tests which basically followed the norms and requirements of the National College Entrance Examination. They were all Chinese native speakers and had been learning English as their compulsory subject since their primary education. English was their only foreign language and none of them had studied overseas or traveled to other countries.

Instruments

A composite questionnaire (See the Appendix) was used in this study to investigate Chinese high school students' L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement. The first part of the questionnaire collected the demographic information of the participants. The second part of the questionnaire, consisting of 19 items, was made up of four subscales rated on a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

The L2 Writing Anxiety Scale (L2WAS) was adapted from Cheng's (2004) Second Language Writing Anxiety Inventory to measure L2 learners' writing anxiety. It contained four items covering cognitive anxiety, somatic anxiety, and avoidance behavior anxiety. A sample item is "I am afraid that my English composition will be chosen for class discussion or evaluation." The internal consistency (Cronbach's α coefficient) of the subscale was 0.817.

The L2 Writing Self-Efficacy Scale (L2WSES) was adapted from Wang and Bai's (2017) Questionnaire of English Self-Efficacy and Sun and Wang's (2020) Questionnaire of English Writing Self-Efficacy. The subscale consisted of five items in five categories, namely organization, use of English writing, ideation, grammar and spelling, and self-efficacy for self-regulation. A sample item is "I can organize sentences into a paragraph to express an idea." The internal consistency (Cronbach's α coefficient) of the subscale was 0.890.

The L2 Writing Self-Regulated Strategy Scale (L2WSRSS) was adapted from the Self-Regulatory Writing Strategy Questionnaire by Teng et al. (2022) to measure the use of writing self-regulated strategies. This subscale consisted of four items covering cognitive strategy, social-behavioral strategy, motivational regulation strategy, and metacognitive strategy, following Teng and Zhang (2016) model. A sample item is "I believe that studying writing strategies will lead to better writing performance." The internal consistency (Cronbach's α coefficient) of the subscale was 0.897.

The L2 Writing Engagement Scale (L2WES) was adapted from Martin's (2012) English Writing Motivation and Engagement Scale to assess students' L2 writing motivation and engagement. This subscale consisted of six items covering emotional (2 items), cognitive (2 items), and behavioral engagement (2 items). A sample item is "When I encounter difficulties in the writing process, I try to overcome them." The internal consistency (Cronbach's α coefficient) of this subscale was 0.837.

Procedure and data analysis

Data collection of this study was done in January 2022 with the help of the headteachers of the participants. As most participants were teenagers, a consent letter was sent to their parents or legal guardians *via* text messages or WeChat messages. After getting their parents' or legal guardians' consents, the headteachers helped administer the questionnaire to the participants *via* an online link sent to students' WeChat or QQ class groups. Prior to that, the participants were informed of the purposes of the study, their roles in data collection, and the confidential and voluntary nature of the study. They were asked to

sign on the questionnaire to ensure their willingness to join the study. Three weeks later, 365 copies of responded questionnaire were collected and 340 of them were found valid, with a 93.15% response rate.

When analyzing the collected data, we adopted AMOS 25.0 and SPSS 26.0 to test the hypotheses by conducting confirmatory factor analysis (CFA) on the validity and examining the reliability of the subscales. Descriptive analysis was done for frequencies, means, and standard deviations. Pearson's correlation coefficients were calculated for the relationship among L2 writing anxiety, L2 writing self-efficacy, and L2 writing engagement. Structural equation modeling was for the analysis of the path relationship among variables. Bootstrapping was utilized to assess the mediation effect of L2 writing self-regulated strategies.

The goodness-of-fit statistics included chi-square statistics (χ^2), degrees of freedom (df), p -value, root-mean-square error of approximation (RMSEA), comparative fit index (CFI), and Tracker-Lewis index (TLI). In terms of model fitness judgment, it is generally considered that the model fitness is reasonable when the evaluation indexes of SEM fitness meet the following criteria: CFI > 0.90, TLI > 0.90, RMSEA < 0.10 (Byrne, 2016). A significant p -value indicates that the model may be appropriate. The guidelines of Gignac and Szodorai (2016) were adopted to interpret the effect size, by which small = 0.10–0.20, medium = 0.20–0.30, and large = \geq 0.30.

Results

Validity and reliability

Results of the study indicate that the four subscales had high structural validity and reliability and all four subscales offered an acceptable fit to the data.

On the basis of Maximum likelihood estimation, CFA was conducted through AMOS 25.0 to assess the overall fitting degree of the four subscales (L2WAS, L2WSES, L2WSRSS, and L2WES) before the hypothesized model was examined. As long as most indexes reach the standard, the data and model fitting can be identified. The fitting indexes of the model all reached the standard of good model fitness ($\chi^2 = 362.84$, $df = 146$, $p < 0.001$, CFI = 0.94, TLI = 0.93, RMSEA = 0.066).

As presented in Table 1, it is identified that the four-factor model of L2WAS fit well ($\chi^2 = 8.29$, $df = 2$, $p < 0.001$, TLI = 0.96, CFI = 0.98, RMSEA = 0.096) based on the fit indices criteria (Byrne, 2006); the five-factor model of L2WSES fit adequately ($\chi^2 = 31.48$, $df = 5$, $p < 0.001$, TLI = 0.94, CFI = 0.97, RMSEA = 0.125); the four-factor model of L2WSRSS fit as well ($\chi^2 = 6.95$, $df = 2$, $p < 0.001$, TLI = 0.98, CFI = 0.99, RMSEA = 0.085); the six-factor model of L2WES also fit greatly ($\chi^2 = 17.21$, $df = 9$, $p < 0.001$, TLI = 0.98, CFI = 0.99, RMSEA = 0.052). The factor loadings of the four subscales ranged from 0.64 to 0.80 (L2 writing anxiety), 0.67 to 0.84 (L2 writing self-efficacy), 0.77 to 0.87 (L2 writing self-regulated strategies), and 0.56 to 0.82 (L2 writing engagement).

In addition, the discriminant validity was tested by comparing the square root of AVE for each subscale and correlation coefficients between each pair of subscales. As Table 1 shows, it is evident that for each subscale the square root of AVE is larger than the correlation coefficients, showing good discriminant validity.

Descriptive statistics and correlations

Results of descriptive analysis (see Table 2) show that of the four variables, the mean value of L2 writing engagement was the highest ($M = 7.66$, $SD = 3.82$), followed by L2 writing self-efficacy ($M = 15.12$, $SD = 4.57$), then L2 writing anxiety ($M = 3.26$, $SD = 4.26$), and finally L2 writing self-regulated strategies ($M = 12.02$, $SD = 3.90$).

Results of correlation analysis revealed a negative relationship between L2 writing anxiety and L2 writing self-regulated strategies ($r = -0.06$, $p > 0.05$) and a positive relationship between L2 writing anxiety and L2 writing engagement ($r = 0.01$, $p > 0.05$), despite not significant. In the meantime, it is identified that L2 writing self-efficacy was found to have a strong positive relationship with L2 writing self-regulated strategies ($r = 0.75$, $p < 0.001$), and L2 writing engagement ($r = 0.66$, $p < 0.001$) and that L2 writing self-regulated strategies had a highly positive relationship with L2 writing engagement ($r = 0.71$, $p < 0.001$).

Structural equation modeling analysis

Figure 2 shows the coefficients of the final model and its standardized path. Although the relationship between L2 writing anxiety and L2 writing self-efficacy was not significant, L2 writing anxiety was found negatively associated with L2 writing self-regulated strategies ($\beta = -0.01$, $p < 0.05$) and positively related to L2

TABLE 1 The fitting indexes and coefficients of the model.

	χ^2	df	p	TLI	CFI	RMSEA	Coefficient α
L2WAS	8.29	2	0.00	0.96	0.98	0.096	0.82
L2WSES	31.48	5	0.00	0.94	0.97	0.125	0.89
L2WSRSS	6.95	2	0.00	0.98	0.99	0.085	0.90
L2WES	17.21	9	0.00	0.98	0.99	0.052	0.84

TABLE 2 Descriptive statistics, correlations, and reliability ($N = 340$).

	L2WAS	L2WSES	L2WSRSS	L2WES	Factor Loadings
L2WAS	0.82				0.64–0.80
L2WSES	−0.07	0.89			0.67–0.84
L2WSRSS	−0.06	0.75**	0.90		0.77–0.87
L2WES	0.01	0.66**	0.71**	0.84	0.56–0.82
Mean	13.26	15.12	12.02	17.66	
Standard Deviation	4.26	4.57	3.90	3.82	

** $p < 0.01$; Cronbach's α coefficients are in bold on diagonals.

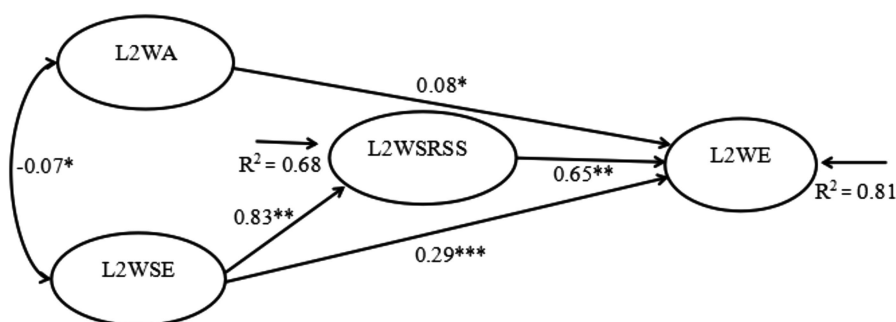


FIGURE 2

Results of SEM model ($N=340$). *** $p<0.001$, * $p<0.05$; goodness-of-fit indices: $\chi^2=362.84$, $df=146$, $p<0.001$, CFI=0.94, TLI=0.93, RMSEA =0.066; non-statistically significantly paths ($p \geq 0.05$) are not reported.

writing engagement ($\beta=0.08$, $p<0.05$), supporting Hypothesis 1. Besides, it was found that L2 writing self-efficacy was positively related to L2 writing self-regulated strategies ($\beta=0.83$, $p<0.001$) and L2 writing engagement ($\beta=0.29$, $p<0.001$) and that L2 writing self-regulated strategies had a positive relationship with L2 writing engagement ($\beta=0.65$, $p<0.001$), supporting Hypothesis 2.

Mediation analysis

Based on Preacher and Hayes (2008), bootstrapping was utilized on the basis of 5,000 samples to examine the mediation effect of L2 writing self-regulated strategies. If the 95% confidence interval between the lower and upper bound did not include zero, indirect effects would be significant (Hayes, 2009).

The mediation effect of L2 writing self-regulated strategies between L2 writing anxiety and L2 writing engagement showed very small effect size ($<|0.10|$) whereas the mediation effect of L2 writing self-regulated strategies between L2 writing self-efficacy and L2 writing engagement showed large effect size ($\geq|0.30|$; see Gignac and Szodorai, 2016). As Figure 2 shows, the 95% confidence interval of indirect effect between L2 writing self-efficacy and L2 writing engagement did not include zero [0.42, 0.69], indicating that L2 writing self-regulated strategies significantly mediated the relationship between L2 writing self-efficacy and L2 writing engagement ($\beta=0.53$, $p<0.001$). However, the 95% confidence interval of indirect effect between L2 writing anxiety and L2 writing engagement included zero $[-0.074, -0.04]$, suggesting that L2 writing self-regulated strategies did not mediate the relationship between L2 writing anxiety and L2 writing engagement. Thus, Hypothesis 3 was rejected.

Discussion

Direct effect

Results of SEM analysis showed that there was a significant negative relationship between L2 writing anxiety and L2 writing

self-efficacy, but the magnitude of the correlation was small. The significant negative relationship between anxiety and self-efficacy in L2 writing was consistent with previous findings (e.g., Woodrow, 2011; Piniel and Csizér, 2013; Kırmızı and Kırmızı, 2015), which suggested that learners with high levels of writing self-efficacy experienced low levels of L2 writing anxiety, and vice versa. Since writing is usually the most challenging part of English exams, it can put more pressure on high school students and hence make them more likely to be overwhelmed by negative feelings (Liu and Wang, 2021). Possibly, the high levels of L2 writing anxiety that the participants of this study experienced was caused by the fear of getting a low grade or negative evaluation. These stressful thoughts were unrelated to the actual writing tasks, and excessively occupied their limited cognitive resources, thus negatively affecting their writing self-efficacy.

Different from previous studies that focused on the self-efficacy in the development of L2 writing, this study focused on high school students' self-efficacy in their writing tasks based on the contents of the questionnaire. Based on the results of this study, it is clear that students' levels of writing self-efficacy were higher than their levels of writing anxiety. Such a discrepancy was possibly associated with the specific context of Chinese high school students' L2 writing tasks. In many cases, they wrote for exams and experienced various types of pressure, particularly that from the highly competitive National College Entrance Examination (Kirkpatrick and Zang, 2011). In addition, mostly lacking systematic training of L2 writing, they tended to rely more on memorizing certain knowledge and materials (Teng et al., 2022).

The study also found L2 writing self-efficacy had a significant positive effect on L2 writing engagement, suggesting that students with higher L2 writing self-efficacy would have high degree of L2 writing engagement. Similarly, Hetthong and Teo (2013) found students with higher L2 writing self-efficacy would regard difficulties as tasks to be mastered, form commitment, develop their interests, and generate more efforts to enhance their L2 writing engagement. This result confirmed the positive effect of L2 writing self-efficacy on writing performance and writing motivation (Sun et al., 2021), which means that higher levels of writing self-efficacy could stimulate students' achievement

motivation, guide them to actively engage in the writing process, and result in better writing performance (Teng and Zhang, 2018).

Additionally, the above findings may be attributed to the newly adopted teaching approaches in Chinese schools such as genre-oriented approach, process-oriented approach and collaborative writing (Yu and Lee, 2016). New teaching methods put more emphasis on learner autonomy and cooperative learning and may have influence on students' self-efficacy and promote their engagement in L2 writing. According to Bandura's (1997) social cognitive theory, students with higher L2 writing self-efficacy tended to make use of L2 writing self-regulated strategies to engage in the writing activities.

The mediating effect

Results of SEM analysis demonstrated that L2 writing self-regulated strategies played a significant mediating role between L2 writing self-efficacy and L2 writing engagement. The low-level direct effect between self-efficacy and engagement in L2 writing indicated that L2 writing self-efficacy could not completely regulate the degree of L2 writing engagement. However, L2 writing self-regulated strategies did not mediate between L2 writing anxiety and L2 writing engagement.

Of all these factors, L2 writing self-efficacy had the strongest positive effect on the L2 writing self-regulated strategies, and L2 writing self-regulated strategies had the second strongest positive effect on the L2 writing engagement. It further confirmed that L2 writing self-regulated strategy was considered as an important indicator of students' L2 writing engagement (Fredricks et al., 2004). The more self-efficacy students have, the more L2 writing self-regulated strategies they adopt, and the higher degree of writing engagement they involve.

The path coefficient suggested that L2 writing self-efficacy was highly positively related to L2 writing self-regulated strategies. This supported the findings of previous studies (Yu et al., 2019; Sun and Wang, 2020), which implied that students with higher writing self-efficacy are usually motivated to adopt various writing self-regulated strategies to engage actively in L2 writing. However, the reduced effect size of L2 writing self-regulated strategies on L2 writing engagement indicated that Chinese high school students were not very strong in adopting more self-regulated strategies to enhance their engagement in L2 writing activities. This can be attributed to the lack of a systematic design of L2 writing courses, sufficient time on writing instruction, and a sustainable L2 writing assessment system for Chinese high school students (Ai, 2015). In addition, Chinese high school students might focus more on linguistic knowledge instead of practical use of writing under the exam-oriented educational system (Sun and Wang, 2020).

As indicated by Teng and Zhang (2018), students tended to feel more self-efficacious to engage in the L2 writing activities with an increase in using L2 writing strategies. In this sense, L2 writing self-regulated strategies could predict L2 writing engagement for

enhancing writing performance since self-regulated students can more easily reflect on their work, process teachers or peers feedback, and set goals to manage and evaluate their writing performance (Teng and Huang, 2019). According to the cognitive process theory (Hayes, 2000), writing is a process-oriented communicative activity involving affective, cognitive, physical, and social conditions and factors such as L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement are all closely linked to the success of L2 writing. This explains why the combined use of cognitive, metacognitive, motivational, and social behavioral strategies can promote the quantity and quality of their engagement in the writing process.

Conclusion

In this study, we explored the relationships among L2 writing anxiety, L2 writing self-efficacy, L2 writing self-regulated strategies, and L2 writing engagement in the Chinese high school context. It has been found that there was a negative relationship between L2 writing anxiety and L2 writing self-efficacy, and the higher level of L2 writing self-efficacy, the lower level of L2 writing anxiety. Both L2 writing anxiety and L2 writing self-efficacy had a direct effect on students' L2 writing engagement, with the direct effect of L2 writing efficacy being much higher than that of L2 writing anxiety. L2 writing self-regulated strategies played a mediating role between L2 writing self-efficacy and L2 writing engagement, and this mediating effect compensated for the insufficient effect of L2 writing self-efficacy on L2 writing engagement for high school students.

Implications and limitations

The findings of this study suggest some pedagogical implications. Given the negative relationship between L2 writing anxiety and L2 writing self-efficacy, as found in this study, teachers are suggested take various measures to provide more positive feedback on L2 writing, offer more encouragement to foster students' agency, and enhance students' L2 writing self-efficacy to alleviate the negative impact of L2 writing anxiety. Given the mediating role of L2 writing self-regulated strategies, it is desirable for teachers to strengthen the training of students' L2 writing self-regulated strategies, enhance their writing engagement, and improve their L2 writing techniques to cope with various difficult challenges in L2 writing. Besides, teachers are also encouraged to help students establish specific L2 writing goals and guide them engage actively and agentively in writing activities so as to improve their writing performance.

There are still some limitations of this study. First, the present study is independent of the specific writing tasks, so it is difficult to demonstrate the interaction between L2 writing anxiety and L2 writing self-efficacy. Second, although the study revealed the

mediating role of L2 writing self-regulated strategies, it did not clarify what specific L2 writing self-regulated strategies could play such a role. Finally, given the huge number of Chinese high schoolers, the sample size of this study is not large enough to reflect the overall situation of Chinese high school students. In view of this, future studies could enlarge samples of learners to investigate students' L2 writing self-efficacy and emotion-regulation in specific writing tasks. Qualitative methods such as interview and reflective journal could be integrated with questionnaire survey to explore the implementation of L2 writing self-regulated strategies in the writing process.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the School of Foreign Languages and Literature, Shandong University, China. Written informed consent to participate in this study was provided by the participants and their legal guardians/next of kin.

Author contributions

JZ collected the data and wrote the first draft of the manuscript. SW analyzed the data. JZ and JW revised the

manuscript. All the authors contributed to the final round of revision and approved the submitted version.

Funding

This work was supported as a Major Project of the Chinese National Social Science Foundation (grant no. 17AYY022).

Acknowledgments

We would like to thank the journal editors and reviewers for their insightful comments and feedback. We are grateful for all of students for involving with this project, as well as their headteachers for helping us to collect data in their schools.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abolhasani, H., Golparvar, S. E., and Robatjazi, M. A. (2022). Modeling the role of L2 writing anxiety in graph-based composing performance and strategy use. *J. Psycholinguist. Res.* 51, 417–435. doi: 10.1007/s10936-022-09858-8
- Ai, B. (2015). A study of the EFL writing of Chinese learners: a critical narrative. *Chang. Engl.* 22, 294–309. doi: 10.1080/1358684X.2015
- Asmari, A. A. (2013). Investigation of writing strategies, writing apprehension, and writing achievement among Saudi EFL-major students. *Int. Educ. Stud.* 6, 130–143. doi: 10.5539/ies.v6n1p130
- Astrid, A., Rukmini, D., Sofwan, A., and Fitriati, S. W. (2017). The analysis of students' engagement to writing feedback activities viewed from students' writing anxiety. *Int. J. English Educ.* 6, 86–107. Available at: <http://repository.radenfatah.ac.id/id/eprint/4726>
- Bailey, D. R., and Almusharraf, N. (2022). A structural equation model of second language writing strategies and their influence on anxiety, proficiency, and perceived benefits with online writing. *Educ. Inf. Technol.* 27, 1–20. doi: 10.1007/s10639-022-11045-0
- Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: W. H Freeman and Co.
- Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., and Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *J. Educ. Psychol.* 105, 25–38. doi: 10.1037/a0029692
- Byrne, B. M. (2006). *Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming*. 2nd Edn. Mahwah, NJ: Erlbaum.
- Byrne, B. M. (2016). *Structural Equation Modeling with AMOS*. 3rd Edn. New York: Routledge.
- Cheng, Y. S. (2002). Factors associated with foreign language writing anxiety. *Foreign Lang. Ann.* 35, 647–656. doi: 10.1111/j.1944-9720.2002.tb01903.x
- Cheng, Y. S. (2004). A measure of second language writing anxiety: scale development and preliminary validation. *J. Second. Lang. Writ.* 13, 313–335. doi: 10.1016/j.jslw.2004.07.001
- Csizér, K., and Tankó, G. (2017). English majors' self-regulatory control strategy use in academic writing and its relation to L2 motivation. *Appl. Linguis.* 38, 386–404. doi: 10.1093/applin/amv033
- Ekholm, E., Zumbrunn, S., and Conklin, S. (2015). The relation of college student self-efficacy toward writing and writing self-regulation aptitude: writing feedback perceptions as a mediating variable. *Teach. High. Educ.* 20, 197–207. doi: 10.1080/13562517.2014.974026
- Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059
- Gignac, G. E., and Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Pers. Individ. Differ.* 102, 74–78. doi: 10.1016/j.paid.2016.06.069
- Graham, S., Harris, K. R., and Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: the effects of

- self-regulated strategy development. *Contemp. Educ. Psychol.* 30, 207–241. doi: 10.1016/j.cedpsych.2004.08.001
- Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., et al. (2018). Reading for writing: a meta-analysis of the impact of Reading interventions on writing. *Rev. Educ. Res.* 88, 243–284. doi: 10.3102/0034654317746927
- Han, J., and Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: the dynamics of change. *J. Second. Lang. Writ.* 40, 44–59. doi: 10.1016/j.jslw.2018.03.001
- Hayes, J. R. (2000). “A new framework for understanding cognition and affect in writing,” in *Perspectives on Writing: Research, Theory, and Practice*. eds. R. Indrisano and J. R. Squire (Newark, DE: International Reading Association), 6–44.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Commun. Monogr.* 76, 408–420. doi: 10.1080/03637750903310360
- Hayes, J. R. (2012). Modeling and remodeling writing. *Writ. Commun.* 29, 369–388. doi: 10.1177/0741088312451260
- Hetthong, R., and Teo, A. (2013). Does writing self-efficacy correlate with and predict writing performance? *Int. J. Appl. Ling. English Lit.* 2, 157–167. doi: 10.7575/ijalel.v2n.1p.157
- Kirkpatrick, R., and Zang, Y. (2011). The negative influences of exam-oriented education on Chinese high school students: backwash from classroom to child. *Lang. Test. Asia* 1, 1–10. doi: 10.1186/2229-0443-1-4-1
- Kırmızı, Ö., and Kırmızı, G. D. (2015). An investigation of L2 learners' writing self-efficacy, writing anxiety and its causes at higher education in Turkey. *Int. J. Higher Educ.* 4, 57–66. doi: 10.5430/ijhe.v4n2p57
- Li, H. (2017). The ‘secrets’ of Chinese students' academic success: academic resilience among students from highly competitive academic environments. *Educ. Psychol.* 37, 1001–1014. doi: 10.1080/01443410.2017.1322179
- Liu, E., and Wang, J. (2021). Examining the relationship between grit and foreign language performance: enjoyment and anxiety as mediators. *Front. Psychol.* 12:666892. doi: 10.3389/fpsyg.2021.6668
- Machida, N., and Dalsky, D. J. (2014). The effect of concept mapping on L2 writing performance: examining possible effects of trait-level writing anxiety. *Engl. Lang. Teach.* 7, 28–35. doi: 10.5539/elt.v7n9p28
- Martin, A. (2012). “Motivation and engagement: conceptual, operational, and empirical clarity,” in *Handbook of Research on Student Engagement*. eds. S. Christenson, A. Reschly and C. Wylie (New York: Springer), 303–311.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Read. Writ. Q.* 19, 139–158. doi: 10.1080/10573560308222
- Piniel, K., and Csizér, K. (2013). L2 motivation, anxiety and self-efficacy: the interrelationship of individual variables in the secondary school context. *Stud. Second Lang. Learn. Teach.* 3, 523–550. doi: 10.14746/ssl.2013.3.4.5
- Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/BRM.40.3.879
- Qi, D. S., and Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *J. Second. Lang. Writ.* 10, 277–303. doi: 10.1016/S1016-3743(01)00046-7
- Rahimi, M., and Zhang, L. J. (2021). Effects of an engaging process-genre approach on student engagement and writing achievements. *Read. Writ. Q.* 38, 1–17. doi: 10.1080/10573569.2021.1982431
- Reeve, J., Cheon, S. H., and Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemp. Educ. Psychol.* 62:101899. doi: 10.1016/j.cedpsych.2020.101899
- Reeve, J., and Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemp. Educ. Psychol.* 36, 257–267. doi: 10.1016/j.cedpsych.2011.05.002
- Robinson, K. A., Perez, T., White-Levatch, A., and Linnenbrink-Garcia, L. (2020). Gender differences and roles of two science self-efficacy beliefs in predicting post-college outcomes. *J. Exp. Educ.* doi: 10.1080/00220973.2020.1808944
- Santangelo, T., Harris, K. R., and Granham, S. (2016). “Self-regulation and writing: an overview and meta-analysis,” in *Handbook of Writing Research*. eds. C. MacArthur, S. Gramham and J. Fitzgerald, Vol. 2. (New York: Guilford), 174–193.
- Sun, T., and Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System* 90:102221. doi: 10.1016/j.system.2020.102221
- Sun, T., Wang, C., Lambert, R. G., and Liu, L. (2021). Relationship between second language English writing self-efficacy and achievement: a meta-regression analysis. *J. Second. Lang. Writ.* 53, 1–14. doi: 10.1016/j.jslw.2021.100817
- Tahmouresi, S., and Papi, M. (2021). Future selves, enjoyment and anxiety as predictors of L2 writing achievement. *J. Second. Lang. Writ.* 53:100837. doi: 10.1016/j.jslw.2021.100837
- Teng, F., and Huang, J. (2019). Predictive effects of writing strategies for self-regulated learning on secondary school learners' EFL writing proficiency. *TESOL Q.* 53, 232–247. doi: 10.1002/tesq.462
- Teng, F., and Reynolds, B. L. (2019). “English foreign and secondary language literacy development across Chinese speaking Asia: what do we know?” in *English Literacy for Chinese Speakers*. eds. B. L. Reynolds and F. Teng (Singapore: Palgrave Macmillan), 3–13.
- Teng, M. F., Wang, C., and Zhang, L. J. (2022). Assessing self-regulatory writing strategies and their predictive effects on young EFL learners' writing performance. *Assess. Writ.* 51:100573. doi: 10.1016/j.asw.2021.100573
- Teng, L. S., and Zhang, L. J. (2016). A questionnaire-based validation of multidimensional models of self-regulated learning strategies. *Mod. Lang. J.* 100, 674–701. doi: 10.1111/modl.12339
- Teng, L. S., and Zhang, L. J. (2018). Effects of motivational regulation strategies on writing performance: a mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacogn. Learn.* 13, 213–240. doi: 10.1007/s11409-017-9171-4
- Tsao, J. J. (2021). Effects of EFL learners' L2 writing self-efficacy on engagement with written corrective feedback. *Asia Pac. Educ. Res.* 30, 575–584. doi: 10.1007/s40299-021-00591-9
- Tsao, J. J., Tseng, W. T., and Wang, C. (2017). The effects of writing anxiety and motivation on EFL college students' self-evaluative judgments of corrective feedback. *Psychol. Rep.* 120, 219–241. doi: 10.1177/0033294116687123
- Usher, E. L., and Pajares, F. (2008). Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78, 751–796. doi: 10.3102/0034654308321456
- Wang, C., and Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Q.* 51, 931–947. doi: 10.1002/tesq.355
- Woodrow, L. (2011). College English writing affect: self-efficacy and anxiety. *System* 39, 510–522. doi: 10.1016/j.system.2011.10.017
- Yu, S., and Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Lang. Teach.* 49, 461–493. doi: 10.1017/S0261444816000161
- Yu, S., Zhou, N., Zheng, Y., Zhang, L., Cao, H., and Li, X. (2019). Evaluating student motivation and engagement in Chinese EFL writing context. *Stud. Educ. Eval.* 62, 129–141. doi: 10.1016/j.stueduc.2019.06.002
- Zhou, S. A., and Hiver, P. (2022). The effect of self-regulated writing strategies on students' L2 writing engagement and disengagement behaviors. *System* 106:102768. doi: 10.1016/j.system.2022.102768
- Zimmerman, B. J. (2001). “Theories of self-regulated learning and academic achievement: an overview and analysis,” in *Self-regulated Learning and Academic Achievement: Theoretical Perspectives*. eds. B. J. Zimmerman and D. H. Schunk, 2nd Edn. (New York: Lawrence Erlbaum), 1–36.
- Zimmerman, B. J., and Risemberg, R. (1997). Becoming a self-regulated writer: a social cognitive perspective. *Contemp. Educ. Psychol.* 22, 73–101. doi: 10.1006/ceps.1997.0919



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga, Spain

REVIEWED BY

Nieves Gutiérrez Ángel,
University of Almería, Spain
Hui Helen Li,
Wuhan University of Technology,
China

*CORRESPONDENCE

Andréanne Gagné
andreanne.gagne@fse.ulaval.ca

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

RECEIVED 22 September 2022

ACCEPTED 17 November 2022

PUBLISHED 01 December 2022

CITATION

Gagné A, Bigras N, Charron A and
Lemire J (2022) The integration
of affective characteristics of the
family environment for a more
comprehensive explanatory model
of reading abilities.
Front. Educ. 7:1051261.
doi: 10.3389/feduc.2022.1051261

COPYRIGHT

© 2022 Gagné, Bigras, Charron and
Lemire. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

The integration of affective characteristics of the family environment for a more comprehensive explanatory model of reading abilities

Andréanne Gagné^{1,2,3*}, Nathalie Bigras^{2,4}, Annie Charron^{2,4}
and Julie Lemire²

¹Department of Teaching and Learning, Université Laval, Québec City, QC, Canada, ²Equipe de Recherche Qualité des Contextes Éducatifs de la Petite Enfance, Université du Québec à Montréal, Montréal, QC, Canada, ³Department of Language Teaching, Université du Québec à Montréal, Montréal, QC, Canada, ⁴Department of Didactics, Université du Québec à Montréal, Montréal, QC, Canada

Introduction: This research focuses on the influence of familial affective characteristics on family literacy practices and children's reading abilities. Parenting stress and educational practices were two affective characteristics of interest. Parenting stress is defined as a state of psychological discomfort specifically associated with the education of a child whereas educational practices are defined as various means the parent uses to educate and socialize the child.

Methods: A sample of 154 grade 1 children allowed for a correlational analysis between parenting stress, educational practices, the frequency of family reading activities, the diversity of literacy material available and the type of child-parent exchange (alphabet-focus or story-focus). Regression analyses were conducted to develop a model predicting reading abilities.

Results: Three result outcomes are of interest for the field of reading development. First, our study establishes relations between educational practices and certain aspects of family literacy: diversity of supports, frequency of exchanges, and type of child-parent exchange and it suggests that parental engagement plays a significant role in various aspects related to at-home discussions about books. Second, our regression analysis highlights evidence that parenting stress is an explanatory factor directly linked to the child's reading abilities. Therefore, our findings add reading abilities to the list of developmental aspects that is affected by parenting stress. Finally, the results show that, when parenting stress and educational practices are integrated in the predictive model, the story-focus exchanges remain predictive of reading abilities but not the alphabet-focus exchanges.

Discussion: Our findings confirm that the benefit of parent-child exchange on reading abilities is dependent of conditions of the family environment in which these activities occur. These findings also lead us to question the

value of making alphabet-focus exchanges, the cornerstone of some literacy programs in family settings. Our findings call for caution when implementing such programs in family context. In fact, activities involving conversation about the meaning of a text or the links between the text and the child's everyday reality represent the only activities, in our study, that had a beneficial effect on reading abilities while remaining permeable to parenting stress.

KEYWORDS

literacy, parenting stress, educational practices, family literacy practices, reading abilities, shared book reading, parent–child interaction

Introduction

The benefits of reading activities in the family environment for children's development of language and writing skills are well known (Justice et al., 2005; Evans and Shaw, 2008; Anderson et al., 2018). However, the majority of studies of this topic have investigated social, demographic or cultural factors to define how, when, how often or in what language these family reading times are experienced (Justice and Ezell, 2000; Nichols, 2000; Justice and Pullen, 2003; Moody et al., 2020). Yet familial affective characteristics are indissociable from parent–child reading times. Therefore, we propose to examine the effect of familial affective characteristics, such as family stress and parental educational practices, on children's reading abilities at the end of first grade, the nature and frequency of literacy activities organized in a household, as well as the diversity of literacy materials available at home.

Theoretical framework

It has long been recognized that a family environment that exposes children to written work is conducive to young children's development of language, reading and writing (Justice and Ezell, 2000; Sénéchal and LeFevre, 2002, 2014). Among children's different experiences with writing at home, shared reading times with parents present benefits related to the motivation to learn to read and emerging literacy skills (e.g., phonological consciousness and understanding alphabetical principles), as well as reading abilities (i.e., decoding and understanding) (Bus et al., 1995; De Jong and Leseman, 2001; Bennett et al., 2002).

At a young age, shared reading familiarizes children with the written word and its prerequisites, helps them understand how a book “works” (e.g., holding it and turning the pages), recognize letters, and develop phonological awareness (Mason and Allen, 1986; Snow and Ninio, 1986). Later, this practice contributes closely to reading performance (Scarborough and Dobrich, 1994; Bus et al., 1995; Mol and Bus, 2011). Moreover,

the influence of shared reading extends beyond literacy or emergent literacy behaviors as it maintains a strong link to the oral modality of language (Frijters et al., 2000; Fletcher and Reese, 2005; Hutton et al., 2015; Niklas et al., 2015). Indeed, shared reading would stimulate oral language skills more than written language skills in child learners.

Another component of the home literacy practices is the quality of parent–child exchanges. The quality of exchanges can be defined as instructional or emotional quality (Cline and Edwards, 2013, 2017). An exchange with favorable instructional quality will include extratextual conversation when the parent guides the child to interrupt the reading and think about it, asking open-ended questions to incite discussion and reflection. On the other hand, reading during which the child remains passive is considered to be of lesser instructional quality. The emotional quality of an exchange resides in the presence of a myriad of parental behaviors, such as empathy, warmth and positive attitude (Van Ijzendoorn et al., 1987; Beaudoin, 2002; Boudreau et al., 2018). This tie between emotions and reading skills has been observed in young adults (Jiménez-Pérez et al., 2020) as well as in young children (Waters et al., 2019). The literature emphasizes that mothers have a crucial role in mediating positive emotions and their benefits to the child's cognitive and language competence (Jiménez-Pérez et al., 2019), fostering the development of literacy abilities in all families where it is observed. On the other hand, the level of instructional quality varies according to socio-familial characteristics. Consequently, the benefits of talking about books for the development of reading abilities may differ from one family setting to another (Cline and Edwards, 2013). This distinction suggests that family characteristics (socio-demographic and cultural characteristics) can impact FLAs' beneficial effects on reading abilities.

In addition, the nature of the emotional ties that the child has with the adult influences the benefits that the child will derive from shared reading times. For example, the quality of the attachment between parent and child is intimately linked to the quality of the exchanges during reading time (Bus and Van Ijzendoorn, 1988). Emotions are more and more in the

researchers' sights. Strangely, even though they are involved in all learning, they have long been the great absentees of learning models in language didactics (Swain, 2013). Yet, it is natural to assume that a negative emotion (such as anxiety) may be associated with the absence of learning, and that a positive emotion (such as joy) may be linked to sustained engagement that promotes learning. This is precisely what Fredrickson and Losada (2013) stipulated in the expansion and construction theory.

In more detail, Fredrickson and Losada theorized the major role of positive short and long-term emotions in the learning process. Their theory, as well as other approaches emphasizing the importance of affectivity in learning (e.g., Arnold, 2006; Armand et al., 2013; Swain, 2013; Pramanik and Dhir, 2020), have been translated to education in several ways. Think of the work on the importance of student well-being in the reception classroom, the supportive role of affective relationships (closeness, friendship) within the classroom and their impact on learning (Rodríguez et al., 1996), the importance of affect and emotion in second language learning (Zuniga and Rueb, 2018), or the role of feelings in vocabulary learning (Tremblay, 2018). In fact, all evidence suggests that positive emotions, fostered by an open and caring climate, catalyze language learning in school and in home settings.

A series of publications by Sénéchal and LeFevre (2002, 2014) established the foundations of the Home Literacy Model (HLM), which made it possible to clarify the benefits of family literacy activities (FLAs). FLAs were then categorized into formal and informal activities. Formal literacy activities were defined as activities involving exchanges between the parent and child that explicitly present or teach about print (e.g., naming letters, recognizing words) (Bus, 2001), whereas informal literacy activities refer to exposure to print (letters, sounds of letters, etc.) without any formal teaching of letters or sounds. During informal literacy activities, exchanges between the parent and child mostly deal with the meaning of the text or linkages between the text and the child's everyday life (Justice et al., 2006). As a result of research conducted on HLM, both formal and informal literacy activities were recognized as contributing significantly to the development of reading abilities (Sénéchal and LeFevre, 2002, 2014). Nonetheless, the preponderance of formal and informal FLAs remains uneven in family routines. A thorough examination of parent-child dialogs has revealed that the vast majority of parent-child exchanges when reading books are informal and that very few discussions of the formal aspects of literacy can be observed in natural, spontaneous contexts (Evans and Saint-Aubin, 2005; Price et al., 2009; Hindman et al., 2014).

Beyond the nature of exchanges, which can be formal or informal, the literacy experience inside a family unit varies according to the household literacy environment. For example, the supports provided in written forms (e.g., magnetic letters, picture books, books, digital tablets) and frequency of

conversations related to literacy may differ greatly from one family to another (Grieshaber et al., 2012). A child who has limited supports in written form or negligible time to converse with parents will not experience the same diversification of FLAs as one who benefits from a rich family literacy environment with ample time to talk with parents (Crosnoe et al., 2010; Marcella et al., 2013; Provencher, 2014).

In this study, we were interested in investigating family characteristics, other than socio-demographic and cultural, that can affect the benefits of FLAs for children's reading abilities. Parenting stress and educational practices were two affective characteristics of interest, especially in the pandemic context we have experienced in the past 2 years. Parenting stress is defined as a state of psychological discomfort specifically associated with the education of a child (Lacharité et al., 1992). The component of this stress that is related to the parent-child relationship (Abidin, 1995) has been identified as a predictor of the presence, quality and impact of FLAs (Deniz-Can and Ginsburg-Block, 2016), confirming past studies by Bus and Van Ijzendoorn (1988) and with Bus and Van Ijzendoorn (1995) that also associated the quality of the parent-child bond with the quality of shared reading times. Other familial affective factors have also been linked to children's reading abilities. One particularly interesting factor is educational practices, defined as various means the parent uses to educate and socialize the child (Hamel, 2001). Positive educational practices are generally observed in the parent's engagement with the child, usage of positive educational practices and a sense of efficacy with discipline (Abidin, 1995). Favorable educational practices and a positive parenting style have been associated with higher-quality interactions during reading times and have resulted in enhanced positive effects on the child's language development (Dexter and Stacks, 2014). These benefits have been observed consistently, in populations of different genders, ages and ethnic origins, and educational institutions (Dornbusch et al., 1987; Ferguson, 1987; Christenson et al., 1992).

Our study proposes a predictive model of FLAs' on reading abilities that integrates parenting stress and educational practices. We posited that (1) *the diversity of supports in written forms is independent of educational practices and parenting stress* since it is closely linked to parents' financial resources and education level. However, (2) *parenting stress should have a negative effect on the frequency of parent-child literacy-related talk*. We anticipated that (3) *favorable educational practices and low parenting stress should both impact positively reading abilities*. Moreover, since informal literacy activities are more conducive to diverse extratextual discussions and represent a larger percentage of exchanges than formal ones, we argued that (4) *the influence of educational practices and parenting stress on reading abilities will be mostly explained by its impact on informal literacy activities*.

Figure 1 illustrates these hypotheses concerning the anticipated links between familial affective characteristics

(educational practices and parenting stress), the literacy environment (diversity and frequency), FLAs (formal and informal) and reading abilities. The hypothesis pertaining to these linkages, presumed to be more important, is illustrated by the width of the arrow connecting parenting stress, informal activities and reading abilities.

Materials and methods

Participants

Following the institutional ethical approval and the authorization granted by the Commission d'Accès à l'information, the research team acquired by way of the Régie de l'assurance maladie du Québec, a list of 4,575 children turning 7-year-olds between 1 October 2015 and 30 September 2016. The children were chosen randomly from the territories serviced by les Centres de santé et de services sociaux de Montréal, where more than 30% of children aged five or younger live below the poverty line. Letters were sent to 4,575 families, followed by telephone calls to close to 3,000 of the families we were able to trace. From those families contacted, 1,575 children met the selection criteria, which included their health (Apgar > 7, weight > 2500 grams at birth and minimum 37 weeks gestation) and their typical global development, notably their language development. Their mother tongue needed to be French. Of the 1,425 remaining families, 255 families who met the criteria agreed to participate and 796 refused. Participants signed a consent form explaining the purpose of the study, its risks, and benefits. The data presented in this article are drawn from the 154 children (82 girls, 72 boys) that were evaluated during their first year of elementary school. At the time of data collection, 75.9% of the families had an income above the low-income cut-off based on Statistics Canada's criteria (Statistics Canada, 2011), 84.1% were two-parent families, and the majority of parents had a university degree (65.9% of mothers, 60.0% of fathers).

Data collection instruments

Formal literacy activities

To measure formal literacy activities, we used an adapted section of a questionnaire on parental literacy practices inspired by the work of Martini and Sénéchal (2012), which shows excellent internal validity (Cronbach's $\alpha = 0.91$). The selected section contains six questions to which the parent is asked to provide answers on a scale of 1–5. These questions assess the frequency, on a scale of 1–5, of the mother showing her child (1) the names of letters, (2) the sounds of letters, (3) how to write letters, (4) how to write the child's first name, (5) how to write words, and (6) how to read words.

Informal literacy activities

For the measurement of informal literacy activities, we used a list of titles and authors of children's books published in French (Charron et al., 2020), adapted from Sénéchal and LeFevre (2002, 2014). This indirect measure proved to be a suitable indicator of informal literacy activities. First, parents were asked to indicate on the questionnaire how familiar they were with the titles of publications for children, from a list of 25 items, of which 17 were real titles and 8 were fictitious titles. Then, parents were given a list of 25 authors of children's literature and asked to indicate which ones were known to them; similar to the list of titles, 17 were real authors and 8 were fictitious authors. The measure of informal literacy activities consisted of the sum of real titles and authors, minus fictitious titles and authors selected by the parent. A complete description of the adaptation protocol is described in Charron et al. (2020).

Diversity of supports

The diversity of supports was also measured by a parental questionnaire in which the parent was required to specify objects used during literacy activities, among 18 everyday life objects; for example, a pictorial, magnetic letters and a grocery list. This questionnaire was adapted from a sub-section of a questionnaire on parental literacy by Martini and Sénéchal (2012). In their work, these authors revealed that they also used the instrument to measure practices linked to formal literacy activities. In fact, any given material can naturally elicit a formal exchange instead of an informal one. For instance, the usage of magnetic letters can potentially lead to a formal exchange on the sound and form of letters. However, we decided to utilize the questionnaire as a measure of availability of different supports at home, rather than a measure of formal literacy activities because we thought it was important to maintain the distinction between concepts of diversity and formality. Where relevant, a collinearity of these two measures could be eventually considered by way of amendment to our statistical analysis. In any case, the diversity of supports was established by recording the number of supports marked on the questionnaire as being used by parents at home. The complete adaptation protocol is described in details in Charron et al. (2020).

Frequency of exchanges about literacy

For the purpose of measuring the frequency of parents-children exchanges about literacy, a list of six situations of typical exchange was presented to the parent (learn lullabies, discuss images in a book, name images in a book, ask the child to read simple words, point at words, and ask the child to trace or copy letters). This list was developed from scale 3 of a questionnaire developed by Martini and Sénéchal (2012). In its original English version, the scale was comprised of seven questions, including a question on the parent teaching letters, three questions on the parent teaching words, two questions related to exchanges on images in a book and one question on

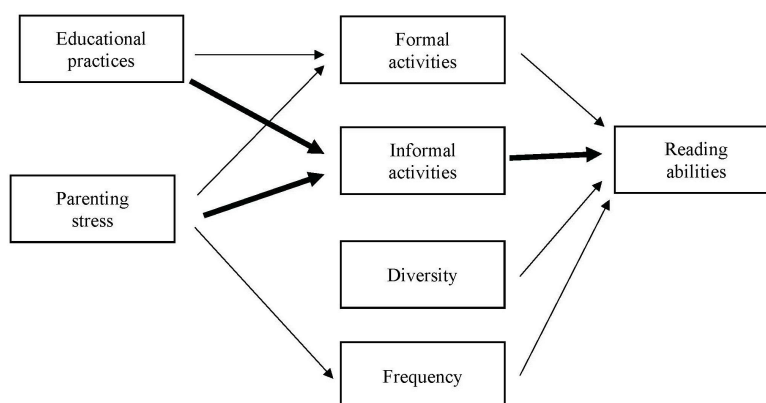


FIGURE 1

Hypotheses concerning the links between reading abilities, family literacy practices and socio-familial factors.

an activity unrelated to literacy, i.e., puzzles. This last question was eliminated because it did not directly pertain to literacy. As for the six questions left, the term “teach” was replaced by terms referring to exchanges, such as “speak with the child, discuss, etc.” Parent was to indicate, on a scale of 1–5, the frequency of activities performed with her child. Although the formal measure explicitly requested the parent to confirm the frequency at which she “shows” something to her child, separate items of exchanges frequency measurement were formulated in a less direct fashion. However, it is possible that the parent interpreted both questions in the same manner. During our statistical analysis, special attention was given to correlation between the measure of formal activities and measure of exchanges frequency.

Reading abilities

Children’s reading abilities were measured by means of the tests on reading words, decoding pseudo-words, and reading comprehension from the Wechsler Individual Achievement Test, 2nd edition (WIAT-II; Wechsler, 2005). In accordance with test protocols, an equivalency score was established for reading components. Regarding the use of this instrument, Wechsler (2005) reported excellent internal coherence coefficients ($\alpha > 0.90$) and very good temporal stability coefficients ($r > 0.80$) for measured abilities.

Educational practices

Parental educational practices were evaluated based on a questionnaire of 42 questions (Poulin et al., 2006). This questionnaire measures three aspects of practices that benefit the child – parental engagement with the child, use of positive educational practices, and sense of efficacy with discipline – as well as three aspects of unfavorable practices and negative educational practices: inconsistent parental practices, use of hostile educational practices, and affective rejection. The Likert-type response scale (1 = not at all to 5 = absolutely) generated

a mean score for favorable practices and a mean score for unfavorable practices. For this study, we used only a mean of scores on the three scales measuring favorable practices.

Parenting stress

To measure parenting stress, we used an abbreviated version of the *Index de stress parental* (Abidin, 1995; Bigras et al., 1996). The index is composed of 36 questions divided into two general categories of stressors: those linked to the child (distraction/hyperactivity, temper, acceptance by the parent, adaptation capacity, level of expectation) and those associated with the parent (sense of competency, attachment to the child, role restrictions, depression, spousal relationship, social isolation, physical health of the parent). A high score signifies that the stress experienced by the parent – the mother in this case – is high.

Data collection

Data were collected in two phases. First, during the month following the child’s seventh birthday, an evaluator visited the familial household to fill out questionnaires with the parents and control the handover sequence, which was identical for all children. The duration of home visits was approximately 120 min. In the second phase, between the months of April and June at the end of first grade, a research assistant assessed the child’s reading abilities with the WIAT-II (Wechsler, 2005) at the child’s residence.

Results

Prior to initiating any analysis, we ensured that our variables met the normality and uniformity criteria. Data were transformed for normality when necessary. Transforming

TABLE 1 Correlation between independent variables.

	Parenting stress	Educational practices	Informal activities	Formal activities	Diversity	Frequency
Parenting stress	1					
Educational practices	-0.503** (<i>n</i> = 146)	1				
Informal activities	-0.01 (<i>n</i> = 135)	0.115 (<i>n</i> = 132)	1			
Formal activities	-0.12 (<i>n</i> = 149)	0.245** (<i>n</i> = 146)	0.025 (<i>n</i> = 136)	1		
Diversity	-0.233** (<i>n</i> = 149)	0.293** (<i>n</i> = 146)	0.05 (<i>n</i> = 136)	0.493** (<i>n</i> = 151)	1	
Frequency	-0.301** (<i>n</i> = 149)	0.294** (<i>n</i> = 146)	0.115 (<i>n</i> = 136)	0.449** (<i>n</i> = 151)	0.553** (<i>n</i> = 151)	1

***p* < 0.01.

data for normality is a regular procedure to allow parametric analyses requiring normality assumption such as regression analysis. Pituch and Stevens (2016) reports that as long as the sample size exceeds 50, transforming data for normality has no impact on validity. In addition, a correlational analysis (Table 1) was done to verify the multicollinearity of independent variables. Not surprisingly, the parenting stress and parental educational practices variables were found to be correlated ($r = 0.503$, $p < 0.01$). Indeed, we can easily imagine that the stress associated with parental responsibilities can lead to less favorable educational practices and vice versa. However, although there is a correlation, the correlation value is considered moderate (Larsen-Hall, 2010). Therefore, although they are linked to each other, the two variables remain distinct constructs that can be reflected separately in modeling.

Regression analyses were done to determine whether parenting stress and educational practices influenced the nature, frequency and diversity of material used during FLAs. According to the regression analyses conducted (Figure 2), favorable educational practices predicted formal literacy activities ($R^2 = 0.06$, $F(2,143) = 5.56$, $p < 0.05$, $\beta = 0.26$, $p < 0.001$), the diversity of supports ($R^2 = 0.08$, $F(2,396) = 7.26$, $\beta = 0.25$, $p < 0.001$), and the frequency of exchanges ($R^2 = 0.09$, $F(2,143) = 8.45$, $p < 0.01$, $\beta = 0.21$, $p < 0.001$), whereas the parenting stress variable did not predict any aspects of at-home discussions of books. These results run counter to our hypotheses 1 and 2, which predicted that the diversity of supports available at home would not be linked to parenting stress or educational practices because they depended mainly on other socio-economic factors, such as family income. Instead, our findings show that it is the frequency of FLAs that is not linked to familial affective variables. Nevertheless, the diversity of supports available at home is linked to both a low level of parenting stress and favorable educational practices.

A second series of regression analyses allowed us to verify whether educational practices and parenting stress predicted reading abilities through family literacy activities (hypothesis 3) and whether their effect on reading was mostly driven by informal or formal literacy activities (hypothesis 4). To test hypothesis 3, we conducted a

regression analysis (Table 2) with a first set of independent variables composed of traditional measures of FLAs; namely formal activities, informal activities, diversity of supports, frequency of exchanges (Model 1). A second regression analysis was conducted (Model 2) with traditional measures of FLAs and familial affective variables (parenting stress and educational practices) as independent variables. Model 2 explains more of the reading variance ($R^2_{adj} = 0.21$, $F(1,123) = 4.7$, $p < 0.01$) than Model 1, ($R^2_{adj} = 0.185$, $F(4,124) = 7.0$, $p < 0.05$), thus confirming that reading abilities are partly predicted by educational practices and parenting stress.

As illustrated in Figure 3, the analyses conducted revealed different roles associated with parenting stress and educational practices. Parenting stress is directly linked to reading abilities whereas educational practices are indirectly linked to reading through their relations with the diversity of supports.

It is relevant to mention that Models 1 and 2 differ as to the significance of formal literacy activities as a predictor of reading abilities. In the first model, three variables are predictive of reading: informal literacy activities, formal literacy activities, and diversity of supports. In Model 2, formal literacy activities no longer have significant predictive value regarding reading abilities.

Discussion

This study investigated the effect of familial affective characteristics on reading abilities. We hypothesized that reading abilities would be indirectly influenced by educational practices and parenting stress, and that this effect would be mediated by informal literacy activities due to their prevalence in parent-child exchanges about books. Instead, our findings showed that educational practices are indirectly linked to reading abilities through the diversity of literacy supports at home, whereas parenting stress is directly correlated with reading abilities.

The results of our study establish relations between educational practices and certain aspects of family literacy (diversity of supports, frequency of exchanges, and formal

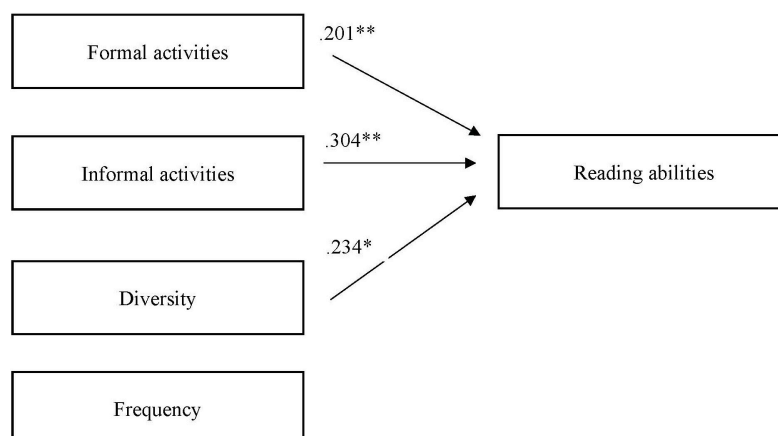


FIGURE 2

Explanatory model of reading abilities without familial affective characteristics. ** $p < 0.01$; * $p < 0.05$.

TABLE 2 Multiple regression analyses for reading abilities.

	<i>B</i>	<i>SE</i>	β
Model 1			
Constant	88.75	11.22	
Informal activities	28.58	7.69	0.304** ¹
Formal activities	0.840	0.406	0.201* ²
Diversity of materials	0.424	0.193	0.234*
Frequency	0.646	0.517	0.134
Model 2			
Constant	109.49	14.655	
Informal activities	28.93	7.58	0.307**
Formal activities	0.787	0.401	0.168
Diversity of materials	0.384	0.191	0.212
Frequency	0.420	0.520	0.088
Parenting stress	−0.221	0.103	−0.181*

¹** $p < 0.01$.

²* $p < 0.05$.

activities). These links have also been mentioned in a previous study of family characteristics and family literacy (Deniz-Can and Ginsburg-Block, 2016), in which participants originated in more challenging socio-economic environments with lower education levels than in our sample. In our study, the mothers' education level was relatively high and homogenous, as were their ethnocultural origins. Nevertheless, our findings suggest that parental engagement plays a significant role in various aspects related to at-home discussions about books.

Furthermore, our regression analysis highlighted evidence that parenting stress is an explanatory factor directly linked to the child's reading abilities. The largest explanatory variance of reading abilities stems from the model that integrates both FLAs and familial affective characteristics (Model 2). Our findings

echo other results regarding parenting stress, particularly its component linked to the parent–child relationship and its influence on various aspects of the child's development (Guajardo et al., 2009; Iruka et al., 2012). Our findings add reading abilities to the list of developmental aspects that may be affected by parenting stress. A remaining question, however, is whether this link can be explained mostly, or even exclusively, by the parent–child relationship component of parenting stress, as is the case for other developmental aspects.

The results of our study raise two important elements for further research and interventions in family literacy. On the one hand, they demonstrate the importance of integrating familial affective characteristics from the environment in which literacy activities take place when it comes to modeling predictors of reading abilities. In fact, a predictive model of reading abilities that does not include variables linked to familial affective factors may lead one to posit relations that no longer exist when family-related characteristics are considered. As shown by model 1 and like several other predictive models of reading abilities (LeFevre and Sénéchal, 2002), formal literacy activities are identified as a predictor of reading abilities. However, when model 2 integrates the familial affective characteristics in which literacy activities take place, formal activities' predictive value declines and only informal literacy activities continue to predict reading abilities. We are not claiming to present an all-inclusive explanatory model of reading abilities. The R^2 value indicates otherwise: close to 80% of the variance in reading abilities remains unexplained, even after the inclusion of parenting stress and educational practices as explanatory variables in the model. However, our findings confirm that the effect of FLAs on reading abilities is not independent of conditions of the family environment in which these activities occur.

Our findings also have implications for the implementation of family literacy intervention programs. In our study, informal literacy activities proved to be predictive of reading abilities,

regardless of the prevailing familial affective characteristics. At first glance, this result seems counterintuitive. We incorrectly believed that the informal activities are less restrictive concerning written work and thus are more susceptible to variations between families (Curenton and Justice, 2008; Shahaeian et al., 2018). However, this is not what we found. In fact, the beneficial effects of formal literacy activities on reading abilities are more affected by familial affective characteristics than those informal activities. It is possible that the responsiveness of formal literacy activities to ethnocultural (Hindman et al., 2014) and socio-affective factors resides in the educational nature of the exchange (e.g., teaching letters, explicitly presenting the sounds of letters, etc.). It has been demonstrated that an exchange with a high degree of educational connotation will be more susceptible to family conditions than an extratextual exchange with a more personal connotation, such as informal literacy activities (Cline and Edwards, 2013). We made the same observations.

Moreover, these findings lead us to question the value of organizing formal literacy activities, the cornerstone of some literacy programs in family settings (Brown et al., 2019). Like Hindman et al. (2014), we question the “one size fits all” recommendations for family literacy programs. In fact, a significant number of studies and intervention programs adhere to a formal, explicit instructional approach toward families presenting vulnerabilities (Justice and Ezell, 2000; Hindin and Paratore, 2007). Our findings call for caution when implementing such programs. In fact, our results suggest instead that informal literacy activities, namely activities involving conversation about the meaning of a text or the links between

the text and the child’s everyday reality, are potentially more beneficial in terms of effects on reading abilities than formal literacy activities where the parent adopts a teaching role or explicitly presents written codes. Furthermore, informal activities correspond to the “natural” role of a parent reading stories to a child (Bingham, 2007; Cline and Edwards, 2013) and represent the only activities, in our study, that had a beneficial effect on reading abilities while remaining permeable to familial affective characteristics.

Finally, another interesting observation, for which an in-depth analysis exceeds the scope of our study, concerns the ever-present influence of familial context on reading abilities at the end of first grade. In fact, after 2 years of schooling (kindergarten and first grade), during which considerable time is spent developing reading abilities and understanding written codes, school attendance does not seem to have diminished the influence of family context on reading abilities. Once again, this highlights the school’s enormous challenge of alleviating social inequalities, even in a largely favorable socio-economic context such as the one in which our study was conducted, where 73% of families came from privileged backgrounds.

Limitations

The correlations presented here are not causal and may equally be represented as two-directional. Moreover, our sample of 154 families included few vulnerable families and, in that sense, is not representative of the general population. As such,

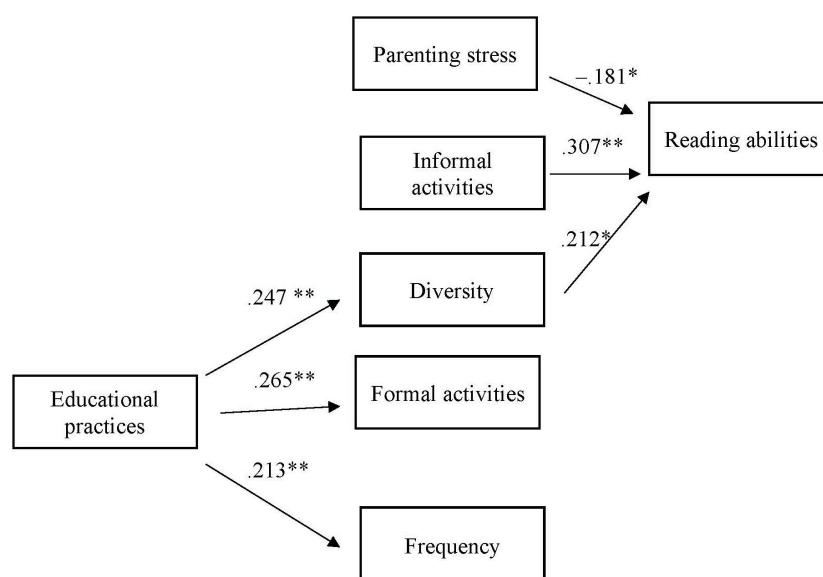


FIGURE 3

Explanatory model of reading abilities including familial affective characteristics. ** $p < 0.01$; * $p < 0.05$.

one of our conclusions concerned the varying nature of the predictive value of FLAs for reading abilities in relation to the family context. Consequently, it is important not to generalize this predictive model to the overall population.

Conclusion

The importance of our findings goes beyond the dissemination of a predictive model of reading abilities. Essentially, our study makes two contributions. First, it confirms the influence of affective-family factors on reading abilities. Second, it reveals that formal literacy activities are more responsive to family context than informal literacy activities. Two familial affective characteristics (educational practices and parenting stress) were integrated into the predictive model. Other factors – social, familial or individual – may also influence FLAs' impact on reading abilities in the same manner, substantially modifying their predictive value. More importantly, it would be surprising if the context in which literacy activities occur had variable effects on reading abilities in a family setting but not in a school environment. Based on these observations, we propose two areas for future research. First, it would be interesting to measure the effects of an intervention program in family literacy that includes family literacy components but also familial affective aspects, namely the parent–child relationship. It would also be worthwhile to measure whether the predictive value of literacy activities varies depending on affective characteristics in a school environment. Finally, we recommend that future studies examine certain aspects of socio-affective characteristics in the classroom, namely teaching practices and stress associated with the teacher–child relationship, as potential influences on the effects of literacy activities in schools.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

- Abidin, R. R. (1995). *Parenting stress index*. Charlottesville, VA: Psychological Assessment Resources.
- Anderson, K. L., Atkinson, T. S., Swaggerty, E. A., and O'Brien, K. (2018). Examining relationships between home-based shared book reading practices and children's language/literacy skills at kindergarten entry. *Early Child Dev. Care* 189, 2167–2182. doi: 10.1080/03004430.2018.1443921
- Armand, F., Lory, M. P., et Rousseau, C. (2013). Les histoires, ça montre les personnes dedans, les feelings. Pas possible si pas de théâtre. (Tahina)
- Ateliers d'expression théâtrale plurilingues en classe d'accueil. *Lidil. Rev. linguist. Didactique Lang.* 48, 37–55. doi: 10.4000/lidil.3311
- Arnold, J. (2006). Comment les facteurs affectifs influencent-ils l'apprentissage d'une langue étrangère? *Ela. Étud. Linguist. Appl.* 4, 407–425.
- Beaudoin, I. (2002). *La qualité des interventions du parent supervisant la lecture de son enfant: Relation avec le niveau d'habileté en lecture de l'enfant [The quality of interventions of parents supervising their children's reading: Correlation with the children's reading ability]*. PhD Thesis. Quebec City: Université Laval.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Université du Québec à Montréal. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This research was supported by a grant from the Canadian Social Sciences and Humanities Research Council (Grant number 435-2012-1133).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Bennett, K. K., Weigel, D. J., and Martin, S. S. (2002). Children's acquisition of early literacy skills: Examining family contributions. *Early Child. Res. Q.* 17, 295–317. doi: 10.1016/S0885-2006(02)00166-7
- Bigras, M., Lafrenière, P., and Abidin, R. R. (1996). *Indice de stress parental: Manuel francophone en complément à l'édition américaine* [Parenting stress index: French-language manual to complement the American edition]. Toronto, ON: Multi-Health Systems Inc.
- Bingham, G. E. (2007). Maternal literacy beliefs and the quality of mother-child book-reading interactions: Associations with children's early literacy development. *Early Educ. Dev.* 18, 23–49. doi: 10.1080/10409280701274428
- Boudreau, M., Beaudoin, I., and Mélançon, J. (2018). Implantation et évaluation d'un programme de formation sur la lecture interactive destiné à des parents d'enfants de maternelle 5 ans [Implementation and assessment of a training program on interactive reading for parents of children in 5-year-old kindergarten]. *Lang. Lit.* 20, 1–22. doi: 10.20360/langandlit29380
- Brown, C. L., Schell, R., Denton, R., and Knode, E. (2019). Family literacy coaching: Partnering with parents for reading success. *Sch. Comm. J.* 29, 63–86. doi: 10.1016/j.pedhc.2017.01.007
- Bus, A. G. (2001). "Joint caregiver-child storybook reading: A route to literacy development," in *Handbook of early literacy research*, eds S. B. Newman and D. K. Dickinson (New York, NY: Guilford Press), 79–191.
- Bus, A. G., and Van Ijzendoorn, M. H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Dev.* 59, 1262–1272. doi: 10.2307/1130489
- Bus, A. G., and Van Ijzendoorn, M. H. (1995). Mothers reading to their 3-year-olds: The role of mother-child attachment security in becoming literate. *Read. Res. Q.* 30, 998–1015. doi: 10.2307/748207
- Bus, A. G., Van Ijzendoorn, M. H., and Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Rev. Educ. Res.* 65, 1–21. doi: 10.3102/00346543065001001
- Charron, A., Gagné, A., Bigras, N., and Lemire, J. (2020). Comment évaluer les pratiques de littératie familiale? Développement d'un questionnaire mesurant l'exposition parentale aux livres [How should we evaluate family literacy practices? Development of a questionnaire measuring parental exposure to books]. *R. Int. Educ. Fam.* 47, 103–118. doi: 10.3917/rief.047.0103
- Christenson, S. L., Rounds, T., and Gorney, D. (1992). Family factors and student achievement: An avenue to increase students' success. *Sch. Psychol. Q.* 7, 178–206. doi: 10.1037/h0088259
- Cline, K. D., and Edwards, C. P. (2013). The instructional and emotional quality of parent-child book reading and early head start children's learning outcomes. *Early Educ. Dev.* 24, 1214–1231. doi: 10.1080/10409289.2012.697431
- Cline, K. D., and Edwards, C. P. (2017). Parent-child book-reading styles, emotional quality, and changes in early head start children's cognitive scores. *Early Educ. Dev.* 28, 41–58. doi: 10.1080/10409289.2016.1177392
- Crosnoe, R., Leventhal, T., Wirth, R. J., Pierce, K. M., Pianta, R. C., and Nichd Early Child Care Research Network (2010). Family socioeconomic status and consistent environmental stimulation in early childhood. *Child Dev.* 81, 972–987. doi: 10.1111/j.1467-8624.2010.01446.x
- Curenton, S. M., and Justice, L. M. (2008). Children's preliteracy skills: Influence of mother's education and beliefs about shared-reading interactions. *Early Educ. Dev.* 19, 261–283. doi: 10.1080/10409280801963939
- De Jong, P. F., and Leleman, P. P. (2001). Lasting effects of home literacy on reading achievement in school. *J. Sch. Psychol.* 39, 389–414. doi: 10.1016/S00224405(01)00080-2
- Deniz-Can, D., and Ginsburg-Block, M. (2016). Parenting stress and home-based literacy interactions in low-income preschool families. *J. Appl. Dev. Psychol.* 46, 51–62. doi: 10.1016/j.appdev.2016.07.002
- Dexter, C. A., and Stacks, A. M. (2014). A preliminary investigation of the relationship between parenting, parent-child shared reading practices, and child development in low-income families. *J. Res. Child. Educ.* 28, 394–410. doi: 10.1080/02568543.2014.913278
- Dornbusch, S. M., Ritter, P. L., Leiderman, P. H., Roberts, D. F., and Fraleigh, M. J. (1987). The relation of parenting style to adolescent school performance. *Child Dev.* 58, 1244–1237. doi: 10.2307/1130618
- Evans, M. A., and Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychol. Sci.* 16, 913–920. doi: 10.1111/j.1467-9280.2005.01636.x
- Evans, M. A., and Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Can. Psychol.* 49, 89–95. doi: 10.1037/0708-5591.49.2.89
- Ferguson, H. (1987). *Manual for multicultural education*. Yarmouth, MA: Intercultural Press.
- Fletcher, K. L., and Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Dev. Rev.* 25, 64–103. doi: 10.1016/j.dr.2004.08.009
- Fredrickson, B. L., and Losada, M. F. (2013). Positive affect and the complex dynamics of human flourishing": Correction to Fredrickson and Losada (2005). *Am. Psychol.* 68:822. doi: 10.1037/a0034435
- Frijters, J. C., Barron, R. W., and Brunello, M. (2000). Direct and mediated influences of home literacy and literacy interest on prereaders' oral vocabulary and early written language skill. *J. Educ. Psychol.* 92, 466–477. doi: 10.1037/0022-0663.92.3.466
- Grieshaber, S., Shield, P., Luke, A., and Macdonald, S. (2012). Family literacy practices and home literacy resources: An Australian pilot study. *J. Early Child. Lit.* 12, 113–138. doi: 10.1177/1468798411416888
- Guajardo, N. R., Snyder, G., and Petersen, R. (2009). Relationships among parenting practices, parental stress, child behaviour, and children's social-cognitive development. *Inf. Child Dev.* 18, 37–60. doi: 10.1002/icd.578
- Hamel, M. (2001). "Les relations parent-enfant et les pratiques éducatives des parents [Parent-child relations and parents' educational practices]" in *Les déterminants de la santé et du bien-être des enfants d'âge scolaire*, eds M. Hamel, L. Blanchet, and C. Martin (Sainte-Foy, QC: Les Publications du Québec), 9–49.
- Hindin, A., and Paratore, J. R. (2007). Supporting young children's literacy learning through home-school partnerships: The effectiveness of a home repeated-reading intervention. *J. Lit. Res.* 39, 307–333. doi: 10.1080/10862960701613102
- Hindman, A. H., Skibbe, L. E., and Foster, T. D. (2014). Exploring the variety of parental talk during shared book reading and its contributions to preschool language and literacy: Evidence from the early childhood longitudinal study-birth cohort. *Read. Writ.* 27, 287–313. doi: 10.1007/s11145-013-9445-4
- Hutton, J. S., Horowitz-Kraus, T., Mendelsohn, A. L., DeWitt, T., Holland, S. K., and C-Mind Authorship Consortium (2015). Home reading environment and brain activation in preschool children listening to stories. *Pediatrics* 136, 466–478. doi: 10.1542/peds.2015-0359
- Iruka, I. U., LaForett, D. R., and Odum, E. C. (2012). Examining the validity of the family investment and stress models and relationship to children's school readiness across five cultural groups. *J. Fam. Psychol.* 26, 359–370. doi: 10.1037/a0028290
- Jiménez-Pérez, E., Alarcón, R., and De Vicente, M. I. (2019). Intervención lectora: Correlación entre la inteligencia emocional y la competencia lectora en el alumnado en bachillerato. *Rev. Psicodid.* 24, 24–30. doi: 10.1016/j.psicod.2018.10.001
- Jiménez-Pérez, E., Barrientos-Báez, A., Caldevilla-Domínguez, D., and Gómez-Galán, J. (2020). Influence of mothers' habits on reading skills and emotional intelligence of university students: Relationships in the social and educational context. *Behav. Sci.* 10:187. doi: 10.3390/bs10120187
- Justice, L. M., and Ezell, H. K. (2000). Enhancing children's print and word awareness through home-based parent intervention. *Am. J. Speech Lang. Path.* 9, 257–269. doi: 10.1044/1058-0360.0903.257
- Justice, L. M., and Pullen, P. C. (2003). Promising interventions for promoting emergent literacy skills: Three evidence-based approaches. *Top. Early Child. Spec. Educ.* 23, 99–113. doi: 10.1177/02711214030230030101
- Justice, L. M., Bowles, R. P., and Skibbe, L. E. (2006). Measuring preschool attainment of print concept knowledge: A study of typical and at-risk 3- to 5-year-old children using item response theory. *Lang. Speech Hear. Serv. Sch.* 37, 224–235. doi: 10.1044/0161-1461(2006/024)
- Justice, L. M., Kaderavek, J., Bowles, R., and Grimm, K. (2005). Language impairment, parent-child shared reading, and phonological awareness: A feasibility study. *Top. Early Child. Spec. Ed.* 25, 143–156. doi: 10.1177/02711214050250030201
- Lacharité, C., Éthier, L., and Piché, C. (1992). Le stress parental chez les mères d'enfants d'âge préscolaire: Validation et normes québécoises pour l'Inventaire de stress parental [Parenting stress in mothers of preschool children: Validation and Quebec norms for the Parenting Stress Inventory]. *Santé Mentale QC* 17, 183–203. doi: 10.7202/502077ar
- Larsen-Hall, J. (2010). *A guide to doing statistics in second language research*. New York, NY: Routledge.
- LeFevre, J., and Sénéchal, M. (2002). Learning to read in a second language: Parent involvement still counts. *Paper presented at the biennial meeting of the international society for the study of behavioral development*, Ottawa, ON.
- Marcella, J., Howes, C., and Fuligni, A. S. (2013). Exploring cumulative risk and family literacy practices in low-income Latino families. *Early Educ. Dev.* 25, 36–55. doi: 10.1080/10409289.2013.780504

- Martini, F., and Sénéchal, M. (2012). Learning literacy skills at home: Parent teaching, expectations, and child interest. *Can. J. Behav. Sci.* 44, 210–221. doi: 10.1037/a0026758
- Mason, J., and Allen, J. B. (1986). A review of emergent literacy with implications for research and practice in reading. *Rev. Res. Educ.* 13, 3–47. doi: 10.3102/0091732X01300100
- Mol, S. E., and Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychol. Bull.* 137, 267–296. doi: 10.1037/a0021890
- Moody, S. M., Matthews, S. D., and Eslami, Z. R. (2020). “Bridging the gap: The use of translanguaging in shared readings,” in *Handbook of research on cultivating literacy in diverse and multilingual classrooms*, eds G. Neokleous, A. Krulatz, and R. Farrelly (Hershey, PA: IGI Global), 555–576.
- Nichols, S. (2000). Unsettling the bedtime story: Parents’ reports of home literacy practices. *Contemp. Issue Early Child.* 1, 315–328. doi: 10.2304/ciec.2000.1.3.7
- Niklas, F., Tayler, C., and Schneider, W. (2015). Home-based literacy activities and children’s cognitive outcomes: A comparison between Australia and Germany. *Int. J. Educ. Res.* 71, 75–85. doi: 10.1016/j.ijer.2015.04.001
- Pituch, K. A., and Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM’s SPSS*. New York, NY: Routledge.
- Poulin, F., Capuano, F., Vitaro, F., and Verlaan, P. (2006). “La prévention des problèmes de comportement chez les filles en début de scolarisation; stratégie de dépistage et profil des filles identifiées [The prevention of behavioural problems in girls at the start of schooling: Screening strategy and profile of girls identified],” in *Les conduites antisociales des filles: Comprendre pour mieux agir*, eds P. Verlaan and M. Déry (Québec, QC: Presses de l’Université du Québec), 235–256.
- Pramanik, A., and Dhir, S. (2020). “Student’s emotion: The Power of emotioneducation in school,” in *Emotion and information processing*, ed. S. N. Mohanty (Switzerland: Springer), 145–160.
- Price, L. H., Van Kleeck, A., and Huberty, C. J. (2009). Talk during book sharing between parents and preschool children: A comparison between storybook and expository book conditions. *Read. Res. Q.* 44, 171–194. doi: 10.1598/RRQ.44.2.4
- Provencher, J. (2014). Impliquer les parents dans le développement en lecture de leurs enfants, c’est gagnant [Getting parents involved in their children’s reading development is a winning move!] Qc. Fr. Available online at: <https://id.erudit.org/iderudit/72029ac> (accessed August 17, 2022).
- Rodríguez, J. I., Plax, T. G., and Kearney, P. (1996). Clarifying the relationship between teacher nonverbal immediacy and student cognitive learning: Affective learning as the central causal mediator. *Comm. Educ.* 45, 293–305. doi: 10.1080/03634529609379059
- Scarborough, H. S., and Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Dev. Rev.* 14, 245–302. doi: 10.1006/drev.1994.1010
- Sénéchal, M., and LeFevre, J.-A. (2002). Parental involvement in the development of children’s reading skill: A five-year longitudinal study. *Child Dev.* 73, 445–460. doi: 10.1111/1467-8624.00417
- Sénéchal, M., and LeFevre, J.-A. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Dev.* 85, 1552–1568. doi: 10.1111/cdev.12222
- Shahaeian, A., Wang, C., Tucker-Drob, C., Geiger, V., Bus, A. G., and Harrison, L. J. (2018). Early shared reading, socioeconomic status, and children’s cognitive and school competencies: Six years of longitudinal evidence. *Sci. Stud. Read.* 22, 485–502. doi: 10.1080/10888438.2018.1482901
- Snow, C. E., and Ninio, A. (1986). “The contracts of literacy: What children learn from learning to read books,” in *Emergent literacy: Writing and reading*, eds W. H. Teale and E. Schulzby (Westport, CT: Ablex Publishing), 116–138.
- Statistics Canada (2011). *Low income lines, 2011–2012*. Ottawa, ON: Statistics Canada
- Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Lang. Teach.* 46, 195–207. doi: 10.1017/S0261444811000486
- Tremblay, O. (2018). *Pour l’amour des mots : Enseigner le vocabulaire en partant du cœur ! [For the love of words: Teaching vocabulary from the heart!]*. Québec, QC: Blog post, Les Éditions passe-temps.
- Van Ijzendoorn, M. H., Van Der Veer, R., and Van Vliet-Visser, S. (1987). “Attachment three years later. Relationships between quality of mother-infant attachment and emotional/cognitive development in kindergarten,” in *Advances in psychology*, Vol. 44, eds L. W. C. Tavecchio and M. H. Van Ijzendoorn (Nirth-Holland: Elsevier), 185–224.
- Waters, T. E., Camia, C., Facompré, C. R., and Fivush, R. (2019). A meta-analytic examination of maternal reminiscing style: Elaboration, gender, and children’s cognitive development. *Psychol. Bull.* 145, 1082–1102. doi: 10.1037/bul0000211
- Wechsler, D. (2005). *Wechsler individual achievement test*, 2nd Edn. London: Psychological Corp.
- Zuniga, M., and Rueb, A. (2018). “Dans le flow : l’influence des tâches pédagogiques sur la qualité de l’expérience émotionnelle en classe de français langue seconde,” in *Emotissage. Les émotions dans l’apprentissage des langues [Emoticization. Emotions in language learning]*, ed. F. Berdal-Masuy (Belgium: Presses universitaires de Louvain), 263–278.



OPEN ACCESS

EDITED BY

Olivia López Martínez,
University of Murcia,
Spain

REVIEWED BY

Hui Helen Li,
Wuhan University of Technology, China
Charanjit Kaur Swaran Singh,
Sultan Idris University of Education,
Malaysia

*CORRESPONDENCE

Mevlüt Gülmez
mgulmez@akdeniz.edu.tr

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 09 September 2022

ACCEPTED 14 November 2022

PUBLISHED 05 December 2022

CITATION

Direkci B, Akbulut S, Şimşek B,
Gülmez M and Nalçacıgil Çopur E (2022)
Analysis of pre-service teachers'
argumentation-based academic writing
process.
Front. Psychol. 13:1040332.
doi: 10.3389/fpsyg.2022.1040332

COPYRIGHT

© 2022 Direkci, Akbulut, Şimşek, Gülmez
and Nalçacıgil Çopur. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Analysis of pre-service teachers' argumentation-based academic writing process

Bekir Direkci, Serdar Akbulut, Bilal Şimşek, Mevlüt Gülmez*
and Emel Nalçacıgil Çopur

Turkish Education Department, Faculty of Education, Akdeniz University, Antalya, Turkey

The purpose of this research is to examine the participating students' argumentation-based academic writing processes and the contributions of these processes to the students' academic writing skills. The participants of the study, which was conducted through mixed method research design, were a group of 53 pre-service Turkish Teachers who are in their first year of Turkish education program at a state university in Turkey during the 2020–2021 academic year. In this research, the data were obtained through student products, rubrics, reflective participant diaries, and a semi-structured interview form. SPSS 23 was used in the analysis of the quantitative data, and NVivo 12 programs were used in the analysis of qualitative data. When the results of the analyses are considered in general, it can be deduced that academic writing practices based on argumentation contributed to the development of students in the dimensions of "subject and content," "organization," "language use," "citation," and "writing process." In the data obtained from the reflective diaries and interviews, although some difficulties were pointed out, the statements of the students regarding their development came to the fore. In addition, it was pointed out that argumentation contributes not only to students' academic writing skills, but also their development of thinking, objectivity, research motivation, and critical thinking.

KEYWORDS

language education, writing education, academic writing, argumentation, mixed method research design

Introduction

Academic articles are texts in which the findings obtained from scientific research carried out for a specific purpose are explained in an academic style (Demirci, 2014). These texts are created for many purposes, such as reporting research, solving a problem, testing a hypothesis, discussing a topic, and giving information about it, and synthesizing the studies done by different researchers (Bailey, 2011; Dinçer, 2018). Academic writing, on the other hand, is the documentation of the research process and its results (Monippally and Pawar, 2010). According to Bahar (2014), academic writing is defined as the reporting process of a scientific research. The reporting process aforementioned is among the most

frequently encountered problems in the conduct of scientific research (Kozak, 2018).

Within the scope of reporting, problems related to developing a thesis statement, making connections between sentences, writing coherent paragraphs, organizing ideas, and the text, elaborating the argument, having problems with word choice and sentence structure, citing sources and reorganizing citations have been pointed out in existing publications (Pablo and Lasaten, 2018; Ivanova, 2020; Lin and Morrison, 2021; Aldabbus and Almansouri, 2022). Within this framework, teaching academic writing includes purposes such as creating the text based on scientific foundations and various evidences; being able to deal with the results obtained as a result of one's own work together with the results of other studies; gaining the ability to conduct research around established standards; being able to see from different perspectives; being able to present a text that is consistent in terms of arguments and content; and being linguistically proficient.

Academic writing should not be seen as just a writing skill. For a qualified study, first of all, an adequate literature review should be done, evidence should be sought for the relevant study and the data obtained should be written down. Thus, scientific studies provide progress in the light of the results of previous research (Kozak, 2018). For this reason, the data to be obtained within the framework of the research and data obtained from different studies should be assimilated and both sets of data should be synthesized. In this process, the concept of "citation" comes into prominence. Citation is one of the distinguishing indicators for academic papers (Coffin, 2009) in which information or finding is associated with a particular source (Swales, 2014; Kafes, 2017; Zhang, 2022). Citation, one of the main features of academic articles, is among effective strategies authors use to support their proposals (Hyland, 1999, p. 362). With this strategy, the author provides evidence for his argument and enriches his/her research with different perspectives. During the citation stage, there are points that the researcher should pay attention to formally and semantically. The author should use the citations in his/her own research by synthesizing it. Clear linking of ideas is important to help readers follow the text (Swales and Feak, 2012). This can be achieved with correct synthesis and the use of correct language. Consequently, in academic writing, the writer must support ideas well, order them logically, justify them rigorously, and weave these ideas together tightly (Fang, 2021). Academic language, on the other hand, is an objective and qualified language that has different functions from everyday speech, requires high-level thinking skills, and is guided by knowledge and technical terms (Sarıkaya, 2020). A scientific style is used in the academic writing process, and certain conclusions are reached by asserting reasons and evidence instead of definite judgments (Gillett et al., 2009). In academic writing, the author's opinion should be clear, s/he should support the arguments s/he put forward with the evidence, prove the accuracy of the statements s/he put forward in the text, and should prefer a convincing and causal style (Caine, 2008). In this respect, the argumentative narrative style has an important place

in academic writing. In the argumentative narrative, the author's opinion is revealed to be valid and valuable by presenting the reasons related to the subject, credible reasons and sufficient evidence (Beyreli and Konuk, 2018). The reasons put forward by the author are the elements that show how his claim can be successfully defended against attacks and how the counterclaim can be refuted (Van Eemeren et al., 2004). In addition, justifications are one of the basic elements that contribute to the development of the argument. The development of the argument presented in an academic text is considered an important feature of successful writing (Lea and Street, 1998). Furthermore, many students have been seen to be either unaware of developing an argument in their writing or have difficulty in doing so (Davies, 2008; Bacha, 2010). In this context, it is very important to develop students' argumentation skills and to use this in academic writing processes. It is necessary to use high-level reasoning and thinking skills in order to create a comprehensive scientific knowledge about a subject (Topdemir and Unat, 2014, p. 7). At the stage of creating scientific knowledge, researchers use scientific arguments to explain the experiment and observation processes (Bakırcı et al., 2017). In this process, researchers frequently use argumentation, which requires asking questions, making claims, and supporting their claims with evidence (Erduran and Jimenez-Aleixandre, 2007; Shi, 2019). The framework of argumentation, which is a structured scientific argumentation technique, was introduced by Toulmin (1958) and was handled by different researchers as a technique for justifying a claim with evidence and solving an existing problem (Jiménez-Aleixandre et al., 2000; Cho and Jonassen, 2002). The overlap between the nature of scientific research and the "data-claim-justification" process, which is at the core of argumentation, highlights the relationship between these two concepts (Uluçınar Sağır and Kılıç, 2013).

When the studies in the literature are reviewed, it is seen that argumentation has positive impacts on the areas such as academic achievement (Canoz et al., 2022), concept teaching (Boyraz et al., 2016), higher-level thinking skills (Erkek and Bostan, 2019; Viyanti et al., 2020), critical thinking and decision-making skills (Tonus, 2012; Giri and Paily, 2020), problem-solving skills (Kardaş, 2013), argumentation creation (Torun and Şahin, 2016), and teaching of socio-scientific issues (Deniz, 2014). However, it was determined that students had difficulties in producing arguments (Jonassen and Kim, 2010) and using elements such as data-justification-support and rebuttal to strengthen their claims (Jiménez-Aleixandre et al., 2000; Kuhn, 2010). As a matter of fact, the inability to verify the asserted claims with data is seen as one of the obstacles to argumentation (Zohar and ve Nemet, 2002). Teaching these factors, which are the basic elements of creating scientific texts, and enabling students to gain argumentation skills should be an important educational goal (Crowell and Kuhn, 2012; Bağ and Çalık, 2017). However, some studies in the literature emphasized that students' argumentation skills are low, and that sufficient attention is not paid to the development of these skills in schools (Weinstock et al., 2006; Crowell and Kuhn, 2012). In this

context, it is of great importance that students gain argumentation skills, which are closely related to scientific method and academic writing.

Although academic writing is seen as an area of interest for graduate students and academics, it is very important to gain this skill at previous levels. As a matter of fact, undergraduate students are expected to conduct research on different courses and subjects, and report these research studies (Siyez, 2016). In order to meet this expectation, it is seen that the content of the Turkish Language II course, which is included in the teacher training programs developed by YÖK (Council of Higher Education), is prepared for teaching academic writing. In addition, it is important to gain these skills not only for meeting the expectation of using academic writing skills during this course, but also for different purposes in other courses from different departments. As a matter of fact, the program aims to train Turkish teacher candidates competently in scientific texts as well as in other text types. Especially, the responsibility given to Turkish teachers in the teaching of informative texts in the *Milli Eğitim Bakanlığı* (2019) requires Turkish teachers to be competent in this field. In this context, it is planned to teach academic writing to Turkish teacher candidates in a way that will contribute to the content of the Turkish Language II course. Within the framework of the lesson plan, pre-service teachers are expected to present their argument as a writer and develop their writing around this argument. The content of the article should be created around the argument discussed and intellectual consistency should be ensured throughout the article. In addition, the author must organize the patterns of ideas about the argument s/he put forward, give them in a logical order, and correctly construct the sentence and paragraph connections during the writing phase. The author should know the correct citation, should not make citation mistakes, and should make the citation in correct form or style. In some studies, it was found that students had difficulty in supporting their claims with evidence (Shi, 2019) and could not evaluate the evidence by associating it with their claims (Sadler, 2004; Watson et al., 2004; Acar et al., 2010) and the evidence they presented contradicted their claims, but they continued to assert their original claims (Evagorou et al., 2012). As another element of academic writing, writers are expected to use a scientific style and objective expressions and write in a fluent and understandable language in accordance with the grammatical rules of the language. While creating a scientific text, the author examines the ideas and findings of different experts and includes them in his/her own article accordingly.

Based on this context, the aim of our research is to examine argumentation-based academic writing processes and the contributions of these processes to students' academic writing skills. The sub-problems of our research around this aim can be listed as follows:

1. What is the effect of argumentation-based academic writing practices on students' ability to create academic texts?
2. How are academic writing practices based on argumentation reflected in student diaries?
3. What are the students' views on academic writing practices based on argumentation?

Materials and methods

Research model

Mixed methods research design was used in the study (McCrudden et al., 2019; Kelle, 2022). In the quantitative part of the study, pre-service Turkish teachers' argumentation-based academic writings were analyzed and scored. Reflective diaries kept by prospective Turkish teachers during the argumentation-based academic writing training process, and the interviews conducted with the participants after the completion of the training process and academic writing phase constitute the qualitative data of this study.

Study group

In this study, purposive sampling technique was used to determine the study group, and pre-service teachers who took "Turkish Language II" course and volunteered to participate in the study were included in the study group in order to ensure the development of academic writing. The study group in the experimental process of the research consisted of 53 Turkish teacher candidates, of whom 28 were female and 25 were male, who were studying in the first grade in the Turkish language teaching undergraduate program of a state university in the spring term of the 2020/2021 academic year, and who had not received any training on academic writing before. The main rationale behind the selection of the study group from Turkish teacher candidates is that they, as a mother tongue teacher, should be competent in scientific texts as well as in different text types in terms of teaching four basic language skills and developing literacy skills. Thus, an important place is given to the teaching of informative texts in the *Milli Eğitim Bakanlığı* (2019) for Turkish teacher candidates in this field. In addition, the fact that Turkish Language II course, which has a course content aimed at gaining academic text writing proficiency in the undergraduate program, was for the first-year students, was also effective in determining the study group.

For the reflective diaries, which constitute one part of the qualitative dimension of the research, all participants who participated in the experimental process kept a weekly diary in the determined time period. In the semi-structured interviews, the other part of the qualitative dimension, 15 Turkish language teacher candidates, of whom seven were male and eight were female, and who participated in the experimental process and volunteered for the interview, constituted the study group.

TABLE 1 Subsections and levels of the rubric used in the research.

Academic writing assessment rubric

Criteria	Lev descriptors	Weight
Subject & Content	5	•highly focused & coherent • thorough & adequate development of thesis
	4	• mostly focused & coherent • good development of thesis
	3	• focused but sometimes incoherent • limited development of thesis
	2	• often unfocused & incoherent • weak development of thesis
	1	• mostly unfocused & incoherent • inadequate development of thesis
Organization	5	• very well organized throughout; clear logical sequencing • effective use of cohesive markers
	4	• well organized throughout; logical sequencing • good use of cohesive markers
	3	• adequately organized; logical but poor sequencing • limited range of cohesive markers
	2	• inadequate organization; lacks logic and poor sequencing • many inappropriate cohesive markers
	1	•serious disorganization; unclear sequence • cohesive markers almost inappropriate
Language Use	5	•writing flows smoothly; very few language errors • highly appropriate register
	4	•writing flows rather smoothly; some language errors • adequate & appropriate register
	3	• many language errors but writing comprehensible • some inappropriate register
	2	•many language errors; writing not comprehensible at times • many inappropriate register
	1	•dominated by language errors; writing mostly incomprehensible • little knowledge of register
Citation	5	• very strong ability to cite and quote accurately • accurate application of citation style
	4	• good ability to cite and quote accurately • apply citation style with occasional errors
	3	• reasonable ability to cite and quote accurately • apply citation style with some errors
	2	• weak ability to cite and quote accurately • apply citation style with many errors
	1	• very weak ability to cite and quote accurately • inaccurate application of citation style
Writing Process	5	• diligently reviewed & proofread • excellent incorporation of others' responses/ideas
	4	• good reviewing & proofreading • very good incorporation of others' responses/ideas
	3	• acceptable reviewing & proofreading • good incorporation of others' responses/ideas
	2	• weak reviewing & proofreading • poor incorporation of others' responses/ideas
	1	• poor reviewing & proofreading • insufficient proof of incorporation of others' responses/ideas

Data collection

In this research, “student products,” “academic writing assessment rubric,” “reflective participant diaries” and “semi-structured interview form” were used as data collection tools. The descriptions of these tools are given in the following parts.

Student products

Academic texts written by students. The students were made to write two academic texts one before and one after the argumentation-based academic writing training. In this framework, a total of 106 written products were obtained from 53 students.

Academic writing assessment rubric

Rubrics are explanatory/graded scoring schemes used to evaluate the learning process of individuals or the learning product that emerges at the end of the process (Brookhart, 2018). These charts, in which the expected/targeted things are defined at each stage, can be used both as an assessment tool and as a teaching tool. While teachers can follow the learning process according to the levels in the rubrics (Arter, 2002), students can also obtain information for the next stages (Moskal and Leydens, 2000). In this

study, an analytical rubric developed by Lallmamode et al. (2016) was preferred in order to evaluate the academic texts produced by the students. The rubric in question consists of academic writing and e-portfolio sections, and in this study, academic writing section consisting of 5 sub-dimensions, “subject and content, organization, language use, citation, and writing process,” was used. The rubric used in the evaluation of academic writing has a 5-stage grading system in each sub-dimension (see Table 1).

Reflective participant diaries

Reflective diaries are written documents including various data (such as analysis, figure, draft, quotation, comment, and impression) in which learners chronologically bring their feelings and thoughts together with their actions such as research, inquiry, experiment, observation, suggestion, etc. (Johnson, 2014). During the research, the students filled out a diary once a week and their opinions and thoughts about the process were determined with these diaries.

Semi-structured interview form

After 8 weeks of experimental process of academic writing implementations, a semi-structured interview form was used to conduct the interviews in order to identify the views of the

participants on the argumentation-based academic writing training.

Implementation process

The research was carried out within the scope of the 1st year spring term “Turkish Language II” course at the Faculty of Education, Turkish Language Teaching. The reason behind choosing this course for the implementation of this research is that the content of the “Turkish Language II” course, which is a common course under the dimension of “General Knowledge” for education faculties within the framework of “Teacher Training Undergraduate Programs” [Yüksek Öğretim Kurulu Başkanlığı (YÖK), 2018], is aimed at academic writing.

Turkish Language II course content:

Features of academic language and writing; using definitions, concepts and terms in academic writings; objective and subjective expression; structure and types of academic texts (articles, reports and scientific abstracts, etc.); making a claim, proposition (justifying, defending, or opposing an idea); formal features of scientific reports and articles; the steps of writing a report; explanation, discussion, establishing intertextual relations, citation (citing and footnotes, bibliography); writing titles, summarizing, writing keywords; ethical principles to be considered in scientific writings; academic text writing practices” [Yüksek Öğretim Kurulu Başkanlığı (YÖK), 2018].

An 8-week plan was constructed for the training and the planned activities for the experimental process to be implemented within the scope of the research. Accordingly, in the first phase, the participants were identified and informed about the study. In the first week, in order to determine the academic writing skills of the participants, the participants were asked to write an academic text on a subject of their own choice. They were informed that they could make use of the library, and the computer laboratory in the faculty during the process of text creation. The written products of the students were evaluated in detail according to the rubric employed in the research. The results were also shared with the students, and they were informed about the evaluation criteria before writing. In addition, the texts created by the students were checked through a plagiarism detection program, and students were delivered feedback.

Between weeks 2 and 7, 6 weeks of training activities were carried out to develop argumentation-based academic writing skills. During the time frame of 6 weeks, activities were carried out on the determined topics (see Table 2). Sample texts were shown, and practices were carried out regarding weekly topics. Furthermore, students were given feedback in line with the rubric criteria. Students were also asked to keep a reflective diary during the weeks of these activities.

In the eighth week, after the training activities, students were asked to create academic texts without any subject restriction.

They were informed that they could benefit from the library, and computer laboratories in the faculty. By collecting the academic texts prepared by the students after the process, both the experimental process was completed, and post-test data were obtained. The results of the evaluation of these texts were also shared with the students and their opinions about the process were obtained. The weekly workload and the subject distribution for the implementation process of argumentation-based academic writing training are shown in Table 2.

Analysis of data

During this research, the data obtained from the data collection tools were analyzed and interpreted in accordance with the structure of the mixed methods research design. The analysis of the data started while the implementation process of the research continued.

Quantitative data

At the beginning of the research process, the researchers examined the academic texts, which were the pre-test data written by the students at the beginning of the research process, and they were scored in line with the academic writing assessment rubric developed by a researcher and an academician with a PhD in Turkish education as an expert. At the end of the 8-week implementation process, the post-test academic texts written by the students were also scored by the same coders in line with the Academic Writing Assessment Rubric.

The reliability of the analysis of student products (academic texts) that were obtained during the research was demonstrated depending on the level of reliability between coders. Cohen's Kappa formula was used to calculate reliability among coders. The data obtained in the study were coded by an expert other than the researcher, and the inter-coder agreement was calculated with kappa. According to the results (see Table 3), the fit values of the pre-test score measurements of the academic writing rubric and its sub-dimensions were 0.86 for subject and content; 0.71 for organization; 0.77 for language usage; 0.78 for citation; 0.72 for writing process, and finally, 0.77 for the total. The fit values of post-test score measurements were 0.80 for subject and content; 0.82 for organization; 0.88 for language usage; 0.94 for citation; 0.80 for writing process; and for the post-test total score, the results were 0.85. These results demonstrate that the agreement between the coders was high in data coding. Any different encodings among the coders were re-evaluated by the coders and after the agreement was reached, the final scoring was calculated, and the data analysis for the experimental part of the research was carried out over the final scoring. In order to determine which statistical tests will be used in the analysis of these data, the results of the normality test of the distribution of the data were examined (see Table 4).

Table 4 demonstrates that according to the Kolmogorov–Smirnov test, the mean values of the academic writing

TABLE 2 Implementation process work-timeline.

Work-time schedule	Weeks							
	1	2	3	4	5	6	7	8
Identification of participants	X							
Preliminary information	X							
Pre-training text creation	X							
Academic text and academic writing		X						
Argumentation		X						
The place and importance of argumentation in academic writing		X						
Finding evidence for arguments			X					
Academic databases			X					
Literature review			X					
Citation rules				X				
Writing bibliography				X				
Creating a text for the argument					X			
Ensuring consistency throughout the text					X			
Organizing the text					X			
Argumentative narration						X		
Developing thesis and antithesis						X		
Language use in scientific writings						X		
Synthesizing							X	
Proofreading studies							X	
Creating post-training text								X
Evaluation								X

pretest-posttest total score and the mean values of pretest-posttest sub-dimension (namely subject and content, organization, language use, citation and writing process) scores do not meet the normality assumption ($Z=0.402; 0.281; 0.295; 0.231; 0.357; 0.162; 0.355; 0.333; 0.300; 0.211; 0.264; 0.161; p<0.05$). Based on these data, non-parametric tests were selected for all analyses. In this direction, the data obtained from the student products within the scope of the study were analyzed with the Wilcoxon Signed Ranks Test.

Qualitative data

Apart from the first and last week activities, students were asked to keep a reflective diary of the activities carried out within the framework of the research. The reason for not keeping a diary in the first and last week was that the participants were expected to see the positive and negative aspects and deficiencies in the texts they wrote in the first week in the process, while they were also expected to reflect their views in the last week in semi-structured interviews. After the 2nd, 3rd, 4th, 5th, 6th and 7th week activities, the students were asked to keep a reflective diary for the activities of that week. Reflective participant diaries were collected weekly during the research and analyzed by the researchers using the content analysis technique. After the analysis of the reflective diaries was completed, a meeting was held in order to eliminate the differences between the coders and the

differences in emerged contents and themes were resolved on the basis of unanimity and a majority of votes.

After the 8-week training implementation process and receiving academic texts from the participants as post-test data, the participants' views on the argumentation-based academic writing process were collected through semi-structured interviews. The interviews were audio recorded so as not to lose any data. The obtained data were transcribed and analyzed by two researchers with content analysis technique using NVIVO 12 program. A meeting was held to eliminate the differences between the researchers after the analysis, and the data analysis process was completed by eliminating the differences in emerged contents and themes on the basis of unanimity/majority vote.

Results

“What is the effect of argumentation-based academic writing practices on students' ability to create academic texts?” The results of the pre-test and post-test applications for this research question are shown in Table 5.

Table 5 demonstrates the total scores and for the scores of each sub-dimension, which are subject and content [\bar{x} (pretest) = 2.34; \bar{x} (posttest) = 4.02], organization [\bar{x} (pretest) = 2.28; \bar{x} (posttest) = 3.47], language use [\bar{x} (pretest) = 2.18; \bar{x}

TABLE 3 Fict indices results.

Rubric and sub-dimensions	Kappa
Subject and content pre-test	,86
Organization pre-test	,71
Language use pre-test	,77
Citation pre-test	,78
Writing process pre-test	,72
Academic writing pre-test total	,77
Subject and content post-test	,80
Organization post test	,82
Language use post test	,88
Citation final test	,94
Writing process post test	,80
Academic writing posttest total	,85

TABLE 4 Normality test results of scores from academic writing and its sub-dimensions.

Academic writing and its sub-dimensions	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Subject and content pre-test	,402	53	,000
Organization pre-test	,281	53	,000
Language use pre-test	,295	53	,000
Citation pre-test	,231	53	,000
Writing process pre-test	,357	53	,000
Academic writing pre-test total	,162	53	,001
Subject and content post-test	,355	53	,000
Organization post-test	,333	53	,000
Language use post-test	,300	53	,000
Citation final test	,211	53	,000
Writing process post-test	,264	53	,000
Academic writing post-test total	,161	53	,002

(posttest) = 3.72], citation [\bar{x} (pretest) = 2.21; \bar{x} (posttest) = 3.04], writing process sub-dimensions [\bar{x} (pretest) = 2.26; \bar{x} (posttest) = 3.53], and academic writing total [\bar{x} (pretest) = 11.28; \bar{x} (post-test) = 17.77]. When the pre-and post-test measurement averages are compared, it is seen that the post-test measurement scores are higher. Whether this difference was statistically significant or not was examined with the Wilcoxon Signed Ranks Test, and the results are presented in Table 6.

According to Table 6, the scores for sub-dimensions are participants' subject and content ($Z = -6.217$; $p < 0.01$), organization ($Z = -5.185$; $p < 0.01$), language use ($Z = -5.236$; $p < 0.01$), citation ($Z = -3.676$; $p < 0.01$), and writing process ($Z = -5.331$; $p < 0.01$), and academic writing total ($Z = -6.188$; $p < 0.01$). It is seen that there is a statistically significant difference between the pre-and post-test scores. Considering the mean ranks (in favor of positive ranks), it can be said that this difference is in favor of post-test scores.

Student's reflective diaries for argumentation-based academic writing practices

After the academic writing practices in the first week and the eighth week, the students were not asked to keep a reflective diary. In the articles written in the first week, it was aimed for the students to see the shortcomings they experienced step by step and to transfer their progress to their diaries. As a result of their academic text writing practices in the eighth week, data were collected from the students through a semi-structured interview form. The students were asked to keep a diary in the activities implemented except for these 2 weeks.

In the second week of the implementation, it was aimed to define the concepts of academic writing and argumentation. At this stage, firstly, the elements of academic writing were conveyed and then the concept of argumentation was taught to the students. The similarities in the nature of academic research have been brought to the fore with the argumentation being based on concepts such as data-claim-justification. In this context, it has been discussed in the classroom environment that academic research is actually intertwined with the concept of argumentation, and how argumentation will contribute to the academic writing process. When the student diaries of the second week were examined, it was seen that the students stated that their knowledge about academic writing and argumentation was insufficient. Some of the students also put forward some thoughts about the inadequacy of the articles they wrote in the first week. Moreover, some of the students emphasized that they were far from these concepts and this situation created anxiety for them. The thoughts of the students about the second week practice were reflected in their diaries as follows:

"Before class I realized that I had almost no knowledge of scientific writing. The things I knew weren't very worthwhile, either. I saw my own shortcomings in the course." (S17)

"I think I had some anxiety because I had never written an article before and because of the fear of not being able to do what I did not know." (S35)

"I had never written an academic text before. So I had no idea about the subject. This is the first time I've heard of the argument. It cannot be said that I wrote the concepts of data-claim-justification very carefully." (S25)

"I had not written a scientific article before and I did not know how to write a scientific article. This situation challenged me in my learning and writing process due to my inexperience." (S53)

TABLE 5 Descriptive statistics for academic writing pre-test and post-test measures.

	N	Minimum	Maximum	Mean	Std. deviation
Subject and content pre-test	53	1	4	2,34	,706
Organization pre-test	53	1	5	2,28	,794
Language use pre-test	53	1	5	2,18	,962
Citation pre-test	53	1	5	2,21	1,17
Writing process pre-test	53	1	4	2,26	,738
Academic writing pre-test total	53	7	18	11,28	2,87
Subject and content post-test	53	2	5	4,02	,604
Organization post-test	53	2	5	3,47	,723
Language use post-test	53	2	5	3,72	,794
Citation final test	53	2	4	3,04	,784
Writing process pos-test	53	2	5	3,53	,749
Academic writing post-test total	53	13	21	17,77	1,58

TABLE 6 Wilcoxon Signed Ranks Test results.

Subject and content post-test-pre-test	Negative Ranks	0	,00	,00	-6,217	,000
	Positive Ranks	49	25,00	1,225,00		
	Ties	4				
	Total	53				
Organization post-test-pre-test	Negative Ranks	5	17,10	85,50	-5,185	,000
	Positive Ranks	42	24,82	1,042,50		
	Ties	6				
	Total	53				
Language use post-test-pre-test	Negative Ranks	4	14,88	59,50	-5,236	,000
	Positive Ranks	41	23,79	975,50		
	Ties	8				
	Total	53				
Citation post-test-pre-test	Negative Ranks	11	14,77	162,50	-3,676	,000
	Positive Ranks	31	23,89	740,50		
	Ties	11				
	Total	53				
Writing process post-test-pre-test	Negative Ranks	3	18,50	55,50	-5,331	,000
	Positive Ranks	42	23,32	979,50		
	Ties	8				
	Total	53				
Academic writing total post-test-pre-test	Negative Ranks	3	3,50	10,50	-6,188	,000
	Positive Ranks	49	27,91	1,367,50		
	Ties	1r				
	Total	53				

“I want to learn how to write articles better. The point I need to learn is how to create a more persuasive writing style, how to write a complete error-free article, how to provide more support to rebuttal views. I think this course contributed to that.” (S10)

In the third week of the implementation, it is aimed to teach the use of literature review and some platforms that can be beneficial in this regard so that students can search for evidence

for their arguments. First of all, information about the purposes of the literature review, the use of keywords and how to reach the right data were given. In this context, the use of Elsevier, ERIC, ULAKBİM and Google Scholar platforms, which can be beneficial to students in the context of educational sciences and Turkish education, were demonstrated in the classroom. When the student diaries were examined, it was seen that the students had no knowledge of Elsevier, ERIC and ULAKBİM platforms, and some of them had used the Google Scholar platform before. While some

students stated that they used libraries as a source, some students emphasized that they benefited from various internet sites. The thoughts of the students about the third week practice were reflected in their diaries as follows:

"Since I didn't have much knowledge about this subject before the lesson, I realized that there were too many platforms to research." (S47)

"I learned where and how I can reach reliable sources during the academic text preparation process. I learned platforms such as Google Scholar, ULAKBİM, Elsevier and ERIC, I think I can use these platforms more often now." (S20)

"I did not know exactly how the literature review was done and what I should pay attention to. Thanks to this lesson, I learned how to do it. Since I did not know this before, I was worried about how to do it at first, but after getting informed, I saw that I could do it." (S27)

"The reason why I had difficulty in searching the sources was that I had never done such serious research. In order to strengthen my claim in what I wrote, I did not look at different studies. While I used to only find and write on the website, now I am trying to reach information by reading articles." (S42)

"I used to find resources by searching the literature on Google Scholar or going to the library. These were guiding my previous studies on the subject. I learned new ways in this lesson." (S13)

In the fourth week of the implementation, it was aimed to teach the citation rules and bibliography to the students. Citation is an element that contributes to the researcher in finding data and evidence to support his claims. Therefore, quoting correctly will contribute to both supporting the argument and improving academic writing skills. Since there are various standards in this regard, the teaching of the APA – 6 style, which is frequently used in the field of educational sciences, was preferred. When the students' views on the practice were examined, it was seen that some students did not have much experience in citation and preparing bibliography, did not attach much importance to citation, and had difficulties in writing a bibliography. The views in the student diaries for the fourth week's practice are as follows:

"Before the lesson, I had no knowledge of the subject. I had no idea that when using someone's word, we need to cite and write bibliography." (S5)

"When citing, I learned where and how to specify it. I learned how to correctly specify any representation that will ensure the integrity of the text, such as the layout of the bibliography, etc." (S29)

"During the course, I learned how important it is to prove a claim, opinion and defend it with a scientific study, by making references, and that the resources we have used on the subject will not be written randomly." (S32)

"Through this course, I first learned how to write a correct article. I did not know much about citing, citation, and plagiarism before the lesson. Thanks to this lesson, I learned how to use these correctly in the article." (S41)

"I knew about in-text citations before. So, it came easy to me. However, I had a hard time writing a bibliography because books, magazines, the internet, etc. in APA standards was different. I had some difficulty in learning because writing the bibliography of the articles was different. Of course, you don't use them all in the same text." (S21)

In the fifth week of the implementation, studies were carried out to create a text for the argument presented by the student and to ensure consistency throughout the text. While creating an academic text, it is of great importance for the student to organize the text correctly and give intellectual patterns in logical order. In this way, both the argument will be defended consistently, and the academic writing will be made into an organized report. In this regard, some students' difficulties were reflected in the statements in their diaries. Examples from the student diaries for the fifth week's practice are as follows:

"I had a hard time with this because I'm a beginner, I don't have a good command of the objective language, and I tend to take topic in a different direction." (S13)

"My arguments in the text I wrote today were qualified and convincing. the point I lacked was building a common bridge between the examples and having difficulty maintaining the objective language." (S6)

"My goals were to express and prove my point of view in the best possible way, (words of important people, etc.) and I believe I could prove, and express them beautifully. I think I have organized the sentences and paragraphs correctly." (S1)

"This course taught me how to defend and develop an idea when I come up with it because it is not enough to put forward the idea, it is necessary to convey it to the other party in a logical way." (S9)

"Thanks to the arguments, we were providing our evidence to the alleged issues and strengthening their credibility. But we were writing irrelevant ideas while doing this. When you read too much, I think one can get distracted from the subject." (S27)

In the sixth week of the implementation, the topic was presenting the argumentative expression in academic texts with a scientific and objective attitude; thus studies have been carried out to use a fluent and understandable style in accordance with the rules of the language. The activity, which was carried out using sample texts in the classroom, aimed at equipping students with the ability to use the elements such as argumentative expression, scientific value, and objectivity, which are in the nature of argumentation and academic writing. When the student diaries regarding this week's activity were examined, it was seen that the students understood the importance of argumentative expression, objective and scientific approach in an academic text and they want to improve themselves by doing exercises. Some of the statements taken from the student diaries written for the sixth week are as follows:

"When I saw the difference between what I first wrote about argumentation and the work I do with my friends now, I saw that I was getting closer to my goals. I realized that I would be more beneficial to my students in the future by using punctuation marks and spelling rules better." (S2)

"During the lesson, we learned how to write an article well, and how to use the language and expressions. Seeing different articles also contributed to this. Maybe we can learn to use a more objective language by taking other articles as an example." (S16)

"We learned that the academic article is more formal and the article we write needs to be proven. I learned that we should not approach the issues that we tell ourselves or that our friends tell us with prejudice because it should not be forgotten that every subject can be true in terms of provability." (S24)

"During the lesson, I learned the rules of a scientific writing, the steps of writing a scientific article, and how to defend our opinion while writing, and how to use argumentative speech to refute the opinion defended by the other party." (S53)

"In order to write a better article, I needed to learn how to do a better literature review and, as a result, better defend my opinion with scientific data and write it in a more scientific language. I think today's lesson contributed to that." (S18)

In the seventh week of the implementation, the ideas of different authors in the literature and providing evidence for the argument put forward by the student were combined in the academic text in line with the student's argument and a proofreading study was carried out. When the student diaries written about this activity were examined, it was seen that the students had difficulty in adding the statements that would support their arguments to their academic writing and experienced various concerns. In addition, it can be said that the proofreading study contributed to the writings of the students. Some statements taken from the student diaries written for the seventh week of the implementation are as follows:

"I find it difficult to make sentences with my own words without plagiarizing the subject because it is difficult to put together sentences without being influenced. I'm worried about plagiarism." (S22)

"I had difficulty in interpreting a person's thought in my own way and writing it down. I was worried and nervous. I was afraid of making mistakes." (S48)

"When I read the text I wrote from beginning to end, and I saw that there were places that needed correction. I've been doing this in other articles I've written before. I have benefited from this habit." (S45)

"I had never worked in an academic field before, or I used to call what we wrote as an article, but it turns out that it has nothing to do with it. Even before the writing phase, the teacher's explanation of the points that needed attention helped me to understand many of my shortcomings from the very beginning. With the activity we did in the last lesson, I can now write a better text." (S7)

"Sometimes I had a hard time combining the quotations in the text because it was a little difficult for me to rephrase. But I think I have achieved that, especially by rereading." (S19)

Student opinions on argumentation-based academic writing practice

In this section, the findings obtained through the semi-structured interviews are included. The opinions of the students regarding the last research question, "What are the student views on argumentation-based academic writing practices?" were gathered around three sub-themes: "the development of academic writing, the problems experienced, and the opinions on the concept of argumentation" (see Figure 1).

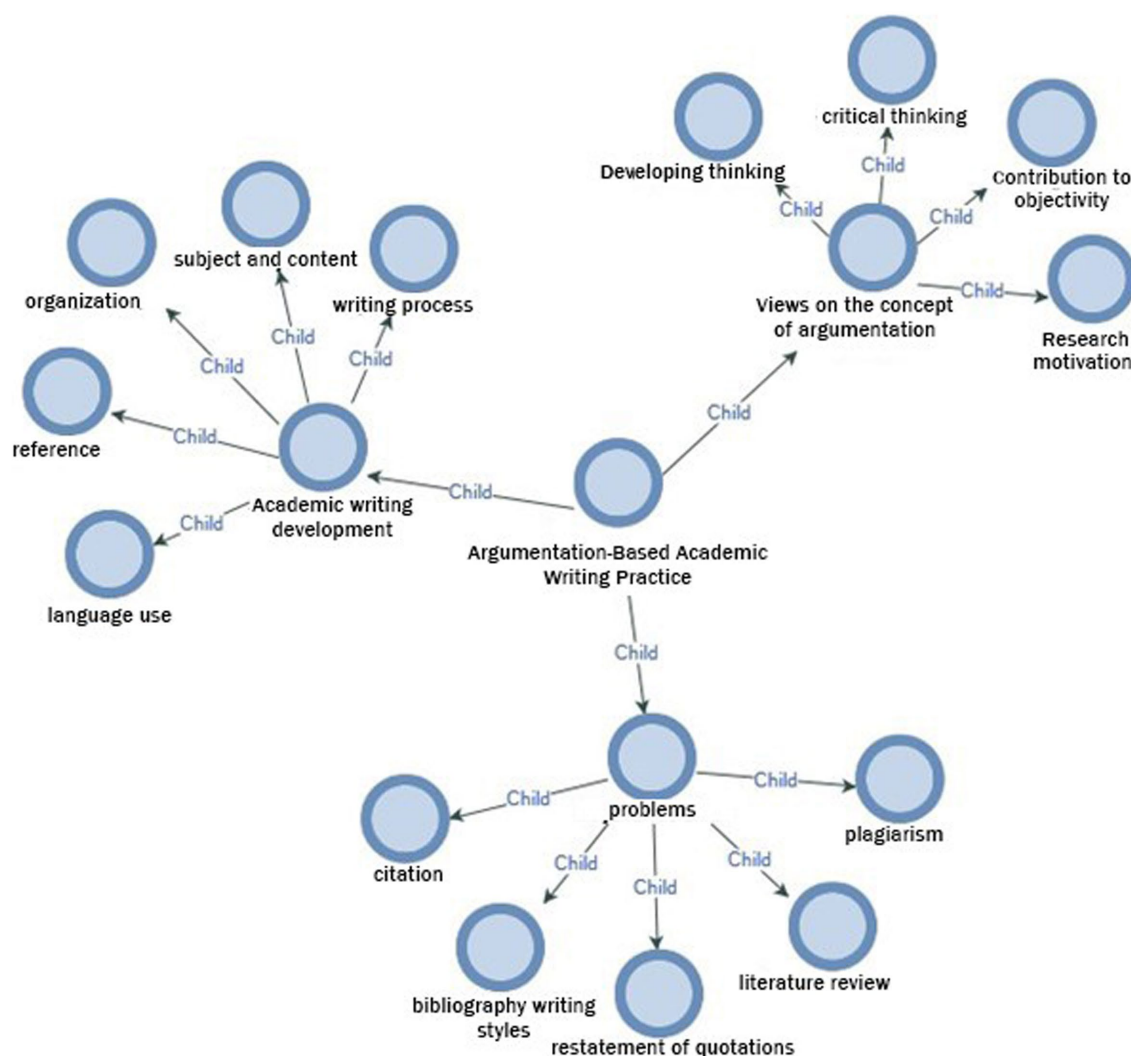


FIGURE 1
A model for student views on argumentation-based academic writing practice.

When the opinions related to the first sub-theme were examined, the statements of the students about the progress they made during the process of developing their argumentation-based academic writing skills were identified. The students' expressions about their development were discussed under the headings of the subject and content, organization, citation, language use and writing process in the rubric used. The students discussed their development in the context of the process and expressed their views on the effects of the practices in question. Some student opinions on the contribution of the practices to the development of academic writing are as follows:

"When we wrote a text around the subject we determined in our studies, I realized that I entered very different subjects in the first

text. The main thing was to support the argument we determined without deviating from the topic. We used this along with the lessons we learned in our last article." (S24)

"Considering the stages from the beginning to the end of the practice, I think that my academic writing proficiency has improved. I have made progress on the language and style to be used in writing an article. This is evident from the difference between the text when we first compose and the text at the end." (S12)

"We actually use concepts such as logical order and organization in different articles. But I think I should have paid a little more attention as I was trying to prove my argument because in order

to convince the other side, it is necessary to construct a logical article.” (S44)

“I think there has been improvement in my academic writing proficiency because I had a hard time citing and writing bibliography at first, but now I can do it without much difficulty.” (S37)

“I have never done proofreading in the texts I have written before. There may be places that need editing in the text, especially in an academic text, there should be no errors. In this respect, I see it as a positive aspect of the implementation.” (S34)

When the opinions of the second sub-theme are examined, it was seen that the students drew attention to some difficulties they experienced during the activities for the development of argumentation-based academic writing processes. Elements such as citation and bibliography writing styles, restatement of quotations, and plagiarism, which were seen for the first time for some of the students, brought along certain difficulties. Some of the students' views on the elements they had difficulties in the implementation process are as follows:

“I especially had problems in citing and plagiarism. because I had never defended a topic with such detail and based on evidence in any study before. What's more, I thought what if it was wrong, because their correct use is also important.” (S7)

“Like every new learner, I experienced difficulties. I had trouble citing because this step was a little more detailed. It was wrong to write the bibliography as I wrote the websites as they are, but I tried to edit them later and write them correctly.” (S19)

“I had problems with plagiarism at the beginning and the end of the implementation. You have trouble finding the appropriate words when conveying someone else's opinion with your own. I try to use synonyms, but I don't know if that's right or wrong.” (S24)

“I just had a hard time doing a literature review. I realized that I didn't know how to search for the topic. When I wrote the subject, I thought it would be right in front of me. I learnt that it was not like that, there were different platforms.” (S49)

“I had some difficulties in the citation and bibliography writing at the beginning because I was hesitant about whether I got it directly or from the citation in the in-text citations. In the

bibliography writing, there were various writing styles such as magazines, internet, and periodicals. So, I had some trouble.” (S12)

When the opinions of the third sub-theme are examined, the statements of the students about the concept of argumentation stand out. By emphasizing that argumentation contributes to academic writing, students highlighted its contribution to thinking and developing thinking skills. In addition, the students stated that the desire to defend an idea and refute the counter-idea provided more research motivation. Furthermore, it was revealed by some students that the effort to refute the counter-idea improved their critical thinking skills and that defending the idea by making use of different studies added objectivity to the academic article. Some student views on the concept of argumentation are as follows:

“Argumentation makes even obsessed people think on occasion. It enables people to look from different perspectives, develop their thoughts with different thoughts, and get rid of the stereotypes.” (S1)

“I think that since the argumentation-based teaching practice reminds us of the importance of research and teaches that not every source is the right source, the information you have obtained from the internet and here can actually be empty, so I think that the practice should be done by every educator.” (S31)

“I find it important to do research in the argumentation-based teaching process, to know different views, to research the subject from different sources, to reach real information and to verify it. I loved doing research from different sources and sites.” (S5)

“I find the activities in the argumentation-based teaching process useful because we need to base our claims on justifications in order to believe the subject we are defending, and we need to research it further. In this way, permanent learning is provided, and the subject is comprehended more. This gives us motivation to investigate further.” (S14)

“While trying to prove and document that the claims of the other party are false during the writing process, we think critically and approach the opinions of the other party in a questioning manner, and we can freely express their opinions.” (S29)

Conclusion and discussion

Considering the results of the research, it was determined that the arithmetic mean of the scores that the students got from the academic texts they wrote before the argumentation-based academic writing practices were carried out was 2,256. The score increased to 3,554 in the academic texts written after the implementation. When the pretest-posttest scores of the students are examined in the context of the dimensions in the rubric, it is seen that there is a significant difference in all of the dimensions, namely “subject and content,” “organization,” “language use,” “citation” and “writing process.” It can be said that the implementation carried out in this research had a positive effect on the academic text writing skills of the students. Indeed, the reflections of this development can be seen in the qualitative data collected during the research process. At the end of the activities, some students’ opinions came to the forefront that they could write the text they wrote within the framework of the arguments they put forward, organize the ideas in the text in a logical order, use a more scientific and objective language, apply the rules of citation and bibliography, integrate different ideas in their writing, and do proofreading. These views support the quantitative data. When the academic texts written before the implementation are examined, it is seen that the students had various deficiencies in writing academic texts. As a matter of fact, in other studies conducted with Turkish teacher candidates, the findings have also revealed that some students did not understand the language used in scientific texts (Yücelşen and ve Edizer, 2020), made mistakes in writing planning, had difficulty in finding the main idea and supporting idea (Arıcı, 2008), and could not provide logical integrity in the text (Göçer, 2010). Moreover, most of the Turkish teacher candidates who participated in the research conducted by Yücelşen and ve Edizer (2020) emphasized that they had difficulties in this regard because they had not practiced academic writing before. The aforementioned research results coincide with the pre-implementation statements of our study group. In our research, student diaries were used to obtain data on the implementation process. When the diaries were examined, it was seen that the students did not have enough knowledge about the concepts of academic writing and argumentation, but they gradually gained knowledge as the weeks progressed. In this process, the concepts of academic writing and argumentation were introduced, the information on literature review was given and the use of various platforms, citation and bibliography rules were taught. A diary was kept by the students during the implementation phase of the study, and weekly developments were observed based on the expressions in the diaries. These developments were also supported by quantitative data and student statements obtained from the semi-structured interviews.

Although there are significant differences in the development of students in the argumentation-based academic writing process, it was stated by the students that there are problems at some points. While some of these problems were reflected in the diaries during the implementation process, some of them continued as a

result of the implementation and were expressed in the interviews. The topics that the students expressed their difficulties experienced during the process only in the diaries they kept were identified as the consistency of subject and content, the organization of the text, the use of language and proofreading. When the studies in the literature are examined, it is seen that some students cannot use the information in the literature for a purpose (Groom, 2000, p. 67), they cannot relate the evidence they reach with their claims (Watson et al., 2004), they write statements irrelevant to their topic; they give irrelevant information, and they cannot express the ideas in a logical order (Wingate, 2012). They feel inadequate in their written expression skills (Bağcı, 2007), and they do not practice proofreading (Yücelşen and ve Edizer, 2020). In this context, it is acceptable for some students to experience difficulties in these topics during the education process. In addition, based on the interviews conducted at the end of the implementation, the continuing difficulties in the academic writing processes of the students were identified as the methods of citation and bibliography writing styles, restatement of citations, plagiarism, and use of recommended platforms for literature review. At the end of the implementation, learning the citation, and bibliography writing styles were expressed by the students as the most difficult subjects to master. The aforementioned student statements can also be found in the quantitative research results, and it is seen that the least developed topic is “citation.” Similarly, in another study, it was revealed that more than half of the Turkish language teaching students participating in the research conducted by Yücelşen and ve Edizer (2020) do not know that they should indicate the sources they use, and that there is no student who knows the correct way of citing sources. In addition, the fact that Google Scholar was preferred among the platforms used by the students in the literature review process and that ULAKBİM, Elsevier or ERIC platforms were not preferred can be seen as supporting data for similar study groups. Another essential element for writing an academic text is to exhibit good language and expression skills. Authors can synthesize and present texts from different research in their studies, and can write shorter versions by restating long texts. Thus, space for more data will be provided in the study and the problems that may rise regarding plagiarism will be minimized (Kozak, 2018). However, it is noteworthy that some students in our study group stated that they had difficulties in these matters though it was the end of the implementation period.

An academic paper in the social sciences requires the author to make a claim about a subject and support this claim with evidence to persuade the reader (Wood, 2001). In this context, it can be said that the concepts of argumentation and academic writing are intertwined. As a result of the argumentation-based academic writing practices based on this idea, it was revealed that the students put forward various opinions about the concept of argumentation. The students participating in the research emphasized that argumentation contributes to academic writing and highlighted its contribution to thinking and developing thinking skills. In this context, Doğan (2012) emphasized that in

educational discussion activities, basing claims on evidence is important in providing students with a different perspective. In different studies in the literature, it has been stated that argumentation contributes to the development of high-level thinking and reflective thinking, judgment and reasoning skills (Nussbaum, 2002; Erduran and Jimenez-Aleixandre, 2007; Kosko et al., 2014; Antonio, 2020). Besides, the students claimed that the desire to defend an idea and refute the counter-idea in academic writings provides motivation for doing more research. These thoughts coincide with the data in the literature that argumentation arouses a sense of curiosity, develops research-inquiry skills, directs the individual to research and motivates him/her (Apaydin and Kandemir, 2018; Özcan et al., 2018; Karaer et al., 2019). Within the scope of the present research, some students stated that argumentation contributed to the development of critical thinking skills. The approach of supporting a claim with evidence and refuting the counter-idea, which is inherent in argumentation, requires critical thinking. Within the framework of this necessity, the contribution of argumentation to critical thinking skills has been revealed in different studies (Namdar and Salih, 2017; Giri and Paily, 2020). In addition, it was stated by some students that defending the argument by making use of different studies added objectivity to an academic article. Argumentation encourages students to use scientific theories, data and evidence to defend their claims about a topic (Simon et al., 2006). The use of different scientific data in an academic article will contribute to more objective expressions in that article.

When the results of the research are considered in general, it was seen that academic writing practices based on argumentation contributed to the development of students in the sub-dimensions of “subject and content,” “organization,” “language use,” “citation” and “writing process.” In the data obtained from reflective diaries and interviews, although some difficulties were pointed out, the statements of the students regarding their development came to the fore. In addition, it was emphasized that argumentation contributed not only to academic writing skills, but also to development of thinking, objectivity, research motivation and critical thinking.

Based on the results of the study, the following suggestions can be made:

1. Considering the mistakes of the students in the texts they first wrote, it can be said that they did not have enough information about the elements of academic writing, so they did not have enough information in the previous stages. In this context, argumentation-based activities can be carried out to improve students' academic writing skills, especially in activities related to informative texts at pre-graduate education levels.
2. The research was carried out with undergraduate level participants who received Turkish language teaching education. The effect of argumentation in academic writing training in different teaching areas can be examined.
3. Considering the student scores before the implementation, it was seen that the academic writing skills of the students should be improved. In order to fulfil this requirement, requirement, besides argumentation-based education in academic writing education, different education programs can be prepared, and the effectiveness of these programs can be examined comparatively.
4. In the statements that emerged in the semi-structured interviews with the students, it was stated that argumentation is an improvement in the areas of development of thinking skills, objectivity, research motivation and critical thinking. In future studies, the relationship between these concepts can be examined.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Akdeniz University Rectorate Social and Human Sciences Scientific Research and Publication Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

BD, SA, and BŞ design of the study and wrote sections of the manuscript. MG and ENÇ organized the database and performed the quantitative and qualitative analysis. BD, SA, BŞ, MG, and ENÇ contributed to the manuscript revision. All authors read and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acar, O., Türkmen, L., and Roychoudhury, A. (2010). Student difficulties in socio-scientific argumentation and decision-making research findings: crossing the borders of two research lines. *Int. J. Sci. Educ.* 32, 1191–1206. doi: 10.1080/09500690902991805
- Aldabbus, S., and Almansouri, E. (2022). Academic writing difficulties encountered by university EFL learners. *Br. J. Engl. Linguist.* 10, 1–11. doi: 10.37745/bjel.2013/vol10n3111
- Antonio, R. P. (2020). Developing students' reflective thinking skills in a metacognitive and argument-driven learning environment. *Int. J. Res. Educ. Sci.* 6, 467–483. doi: 10.46328/ijres.v6i3.1096
- Apaydin, Z., and Kandemir, M. (2018). İlkokulda sınıf öğretmenlerinin fen bilimleri dersinde argümantasyon yöntemi kullanımına ilişkin görüşleri [Opinions of primary school teachers on the use of argumentation method in science lesson]. *J. Comput. Educ. Res.* 6, 106–122. doi: 10.18009/jcer.387033
- Arıcı, A. F. (2008). Üniversite öğrencilerinin yazılı anlatım hataları [written expression mistakes of university students]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi* 21, 209–220.
- Arter, J. (2002). "Rubrics, scoring guides, and performance criteria" in *Understanding scoring rubrics a guide for teachers*. ed. C. Boston (Washington: Office of Educational Research and Improvement)
- Bacha, N. (2010). Teaching the academic argument in a university EFL environment. *J. Engl. Acad. Purp.* 9, 229–241. doi: 10.1016/j.jeap.2010.05.001
- Bağ, H., and Çalık, M. (2017). İlköğretim düzeyinde yapılan argümantasyon çalışmalarına yönelik tematik içerik analizi [Thematic content analysis for argumentation studies at primary school level]. *Eğitim ve Bilim* 42, 281–303. doi: 10.15390/EB.2017.6845
- Bağcı, H. (2007). Türkçe öğretmeni adaylarının yazılı anlatıma ve yazılı anlatım derslerine yönelik tutumlarının değerlendirilmesi [Evaluation of Turkish teacher candidates' attitudes towards written expression and written expression lessons]. *Türklük Bilimi Araştırmaları* 21, 29–61.
- Bahar, M. A. (2014). Lisansüstü eğitimde akademik yazma ve önemi [Academic writing and its importance in graduate education]. *Int. J. Lang. Acad.* 2, 209–233.
- Bailey, S. (2011). *Academic writing a handbook for international students*. New York, NY: Routledge.
- Bakırcı, H., Çalık, M., and Çepni, S. (2017). The effect of the common knowledge construction model-oriented education on sixth grade students' views on the nature of science. *J. Balt. Sci. Educ.* 16, 43–55. doi: 10.33225/jbse/17.16.43
- Beyreli, L., and Konuk, S. (2018). Altıncı sınıf öğrencilerinin ikna edici yazma becerisinin geliştirilmesine yönelik bir araştırma [A research on improving the persuasive writing skills of sixth grade students]. *Eğitim ve Bilim* 43, 181–215. doi: 10.15390/EB.2018.7520
- Boyras, D. S., Hacıoğlu, Y., and Aygün, M. (2016). Argümantasyon ve kavram karmaşası: Erime ve çözünme [Argumentation and conceptual confusion: dissolution and dissolution]. *Gazi Eğitim Fakültesi Dergisi* 36, 233–267.
- Brookhart, S. M. (2018). Appropriate criteria: key to effective rubrics. *Front. Educ.* 3, 1–12. doi: 10.3389/feduc.2018.00022
- Caine, K. (2008). *Writing to persuade*. Portsmouth, NH: Heinemann.
- Canoz, G. M., Ucar, S., and Demircioglu, T. (2022). Investigate the effect of argumentation-promoted interactive simulation applications on students' argumentation levels, academic achievements, and entrepreneurship skills in science classes. *Think. Skills Creat.* 45:101106. doi: 10.1016/j.tsc.2022.101106
- Cho, K., and Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem-solving. *Educ. Technol. Res. Dev.* 50, 5–22. doi: 10.1007/BF02505022
- Coffin, C. (2009). Incorporating and evaluating voices in a film studies thesis. *Writ. Pedagogy* 1, 163–193. doi: 10.1558/wap.v1i2.163
- Crowell, A., and Kuhn, D. (2012). Developing dialogic argumentation skills: A three-year intervention study. *J. Cogn. Dev.* 15, 363–381. doi: 10.1080/15248372.2012.725187
- Davies, W. M. (2008). 'Not quite right': helping students to make better arguments. *Teach. High. Educ.* 13, 327–340. doi: 10.1080/13562510802045352
- Demirci, A. (2014). "Akademik araştırmalarda yazım aşaması [Writing phase in academic research]" in *Coğrafya araştırma yöntemleri [Geography research methods]*. eds. Y. Arive and İ. Kaya (Coğrafyacılar Derneği), 109–154.
- Deniz, T. (2014). Çevre eğitiminde toplumbilimsel argümantasyon yaklaşımının kullanımı [Use of sociological argumentation approach in environmental education], [unpublished master thesis]. Gazi Üniversitesi, Ankara.
- Dinçer, S. (2018). *Akademik yazım ve araştırmacılara öneriler [Academic writing and suggestions to researchers]*. Ankara: Pegem A Yayıncılık.
- Doğan, Y. (2012). "Ottomans in Tunisia through the eyes of ottomans and Spanish" in *Yenilikçi tarih öğretimi yaklaşımları [Innovative approaches to history teaching]*. ed. H. Köksal (Ankara: Harf Eğitim Yayıncılığı), 309–3014.
- Erduran, S., and Jimenez-Aleixandre, M. P. (2007). "Argumentation in science education: Recent developments and future directions" in *Science and Technology Education Library. Vol. 35*. Dordrecht: Springer.
- Erkek, Ö., and Bostan, M. I. (2019). Ortaokul matematik öğretmen adaylarının muhakeme süreçlerine farklı bir bakış: global argümantasyon yapısı [A different view on the reasoning processes of secondary school mathematics teacher candidates: global argumentation structure]. *Eğitim ve Bilim* 44, 75–101. doi: 10.15390/EB.2019.7867
- Evagorou, M., Jiménez-Aleixandre, M. P., and Osborne, J. (2012). Should we kill the grey squirrels? A study exploring students' justifications and decision-making. *Int. J. Sci. Educ.* 34, 401–428. doi: 10.1080/09500693.2011.619211
- Fang, Z. (2021). *Demystifying academic writing: Genres, moves, skills, and strategies*. New York, NY: Routledge.
- Gillett, A., Hammond, A., and Martala, M. (2009). *Successful academic writing*. Harlow: Pearson Longman.
- Giri, V., and Paily, M. U. (2020). Effect of scientific argumentation on the development of critical thinking. *Sci. & Educ.* 29, 673–690. doi: 10.1007/s11191-020-00120-y
- Göçer, A. (2010). Eğitim fakültesi öğrencilerinin yazılı anlatım becerilerinin süreç yaklaşımı ve metinsellik ölçütleri ekseninde değerlendirilmesi: Niğde üniversitesi örneği. [evaluation of written expression skills of education faculty students in the axis of process approach and textuality criteria: Niğde university example]. *Kastamonu Eğit. Derg.* 18, 271–290.
- Groom, N. (2000). "A workable balance: self and source in argumentative writing" in *Learning to argue in higher education*. eds. S. Mitchell and R. Andrews (Boynton/Cook: Heinemann), 65–145.
- Hyland, K. (1999). Academic attribution: citation and the construction of disciplinary knowledge. *Appl. Linguist.* 20, 341–367. doi: 10.1093/applin/20.3.341
- Ivanova, I. (2020). Making sense of research: university students' difficulties in writing academic summaries. *Stud. Linguist. Cult. FLT* 8, 16–34. doi: 10.46687/SILC.2020.v08i01.002
- Jiménez-Aleixandre, M. P., Rodríguez, A. B., and Duschl, R. A. (2000). "Doing the lesson" or "doing science": argument in high school genetics. *Sci. Educ.* 84, 757–792. doi: 10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F
- Johnson, A. P. (2014). Eylem araştırması el kitabı (Y. Uzunur and Y. Özten Anay, trans.). Anı Yayıncılık.
- Jonassen, D. H., and Kim, B. (2010). Arguing to learn and learning to argue: design justifications and guidelines. *Educ. Technol. Res. Dev.* 58, 439–457. doi: 10.1007/s11423-009-9143-8
- Kafes, H. (2017). Citation practices among novice and expert academic writers. *Eğitim ve Bilim* 42, 441–461. doi: 10.15390/EB.2017.6317
- Karaer, G., Karademir, E., and Tezel, Ö. (2019). Sınıf öğretmen adaylarının fen laboratuvarında argümantasyon tabanlı öğretime yönelik görüşlerinin incelenmesi [Examination of pre-service teachers' views on argumentation-based teaching in the science laboratory]. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi* 20, 217–241. doi: 10.17494/ogusbd.548346
- Kardaş, N. (2013). Fen eğitiminde argümantasyon odaklı öğretimin öğrencilerin karar verme ve problem çözme becerilerine etkisi [The effect of argumentation-oriented teaching on students' decision-making and problem-solving skills in science education] [unpublished master thesis]. Eskişehir Osmangazi Üniversitesi, Eskişehir.
- Kelle, U. (2022). "Mixed methods" in *Handbuch Methoden der empirischen Sozialforschung*. eds. N. Baur and J. Blasius (Wiesbaden: Springer VS)
- Kosko, K. W., Rougee, A., and Herbst, P. (2014). What actions do teachers envision when asked to facilitate mathematical argumentation in the classroom? *Math. Educ. Res. J.* 26, 459–476. doi: 10.1007/s13394-013-0116-1
- Kozak, M. (2018). *Akademik yazım: İlkeler-uygulamalar-örnekler [Academic writing: Principles-practices-examples]*. Ankara: Detay Yayıncılık.
- Kuhn, D. (2010). Teaching and learning science as argument. *Sci. Educ.* 94, 810–824. doi: 10.1002/sce.20395
- Lallmame, S. P., Daud, N. M., and Abu Kassim, N. L. (2016). Development and initial argument-based validation of a scoring rubric used in the assessment of L2 writing electronic portfolios. *Assess. Writ.* 30, 44–62. doi: 10.1016/j.asw.2016.06.001
- Lea, M. R., and Street, B. V. (1998). Student writing in higher education: an academic literacies approach. *Stud. High. Educ.* 23, 157–172. doi: 10.1080/03075079812331380364

- Lin, L. H., and Morrison, B. (2021). Challenges in academic writing: perspectives of engineering faculty and L2 postgraduate research students. *Engl. Specif. Purp.* 63, 59–70. doi: 10.1016/j.esp.2021.03.004
- McCrudden, M. T., Marchand, G., and Schutz, P. (2019). Mixed methods in educational psychology inquiry. *Contemp. Educ. Psychol.* 57, 1–8. doi: 10.1016/j.cedpsych.2019.01.008
- Milli Eğitim Bakanlığı (2019). Türkçe dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar) [Turkish lesson curriculum (primary and secondary school grades 1, 2, 3, 4, 5, 6, 7 and 8)]. Ankara: MEB (Milli Eğitim Bakanlığı).
- Monippally, M. M., and Pawar, B. S. (2010). *Academic writing – A guide for management students and researchers* Vivek Mehra for Response Books. Thousands Oaks, CA: Sage.
- Moskal, B. M., and Leydens, J. A. (2000). Scoring rubric development: validity and reliability. *Pract. Assess. Res. Eval.* 7, 1–6. doi: 10.7275/q7rm-gg74
- Namdar, B., and Salih, E. (2017). Fen bilgisi öğretmen adaylarının teknoloji destekli argümantasyona yönelik görüşleri [Views of pre-service science teachers on technology-assisted argumentation]. *Abant İzzet Baysal Üniv. Eğitim Fakültesi Dergi.* 17, 1384–1410. doi: 10.17240/aibuefd.2017.17.31178-338837
- Nussbaum, E. M. (2002). Scaffolding argumentation in the social studies classroom. *Soc. Stud.* 93, 79–83. doi: 10.1080/00377990209599887
- Özcan, R., Aktamış, H., and Hiğde, E. (2018). Fen bilimleri derslerinde kullanılan argümantasyon düzeyinin belirlenmesi [Determining the level of argumentation used in science lessons]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi* 43, 93–106.
- Pablo, J. C. I., and Lasaten, R. C. S. (2018). Writing difficulties and quality of academic essays of senior high school students. *Asia Pac. J. Multidiscip. Res.* 6, 46–57.
- Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *J. Res. Sci. Teach.* 41, 513–536. doi: 10.1002/tea.20009
- Sarıkaya, B. (2020). "Akademik dil ve yazının önemi [the importance of academic language and writing]" in *Türk dili 2 akademik yazma eğitimi [Turkish language 2 academic writing training]*. eds. M. N. Kardeş and R. Koç (Ankara: Pegem Akademi). 1–16.
- Shi, Y. (2019). Enhancing evidence-based argumentation in a mainland China middle school. *Contemp. Educ. Psychol.* 59:101809. doi: 10.1016/j.cedpsych.2019.101809
- Simon, S., Erduran, S., and Osborne, J. (2006). Learning to teach argumentation: research and development in the science classroom. *Int. J. Sci. Educ.* 28, 235–260. doi: 10.1080/09500690500336957
- Siyez, D. M. (2016). PDR'de kaynak tarama ve rapor yazma [source scanning and report writing in PDR]. Pegem A Akademi.
- Swales, J. M. (2014). Variation in citational practice in a corpus of student biology papers: from parenthetical plonking to intertextual storytelling. *Writ. Commun.* 31, 118–141. doi: 10.1177/0741088313515166
- Swales, J. M., and Feak, C. B. (2012). *Academic writing for graduate students: Essential skills and tasks*. Ann Arbor: University of Michigan Press.
- Tonus, F. (2012). Argümantasyona dayalı öğretimin ilköğretim öğrencilerinin eleştirel düşünme ve karar verme becerileri üzerine etkisi [the effect of argumentation-based instruction on critical thinking and decision-making skills of primary school students] [unpublished master thesis]. Hacettepe Üniversitesi, Ankara.
- Topdemir, H. G., and Unat, Y. (2014). *Bilim tarihi [history of science]*. Ankara: Pegem A Yayınları.
- Torun, F., and Şahin, S. (2016). Argümantasyon temelli sosyal bilgiler dersinde öğrencilerin argüman düzeylerinin belirlenmesi [determining the argument levels of the students in the argumentation-based social studies course]. *Eğitim ve Bilim* 41, 233–251. doi: 10.15390/EB.2016.6322
- Toulmin, S. (1958). *The uses of argumenti*. UK: Cambridge University Press.
- Uluçınar Sağır, Ş., and Kılıç, Z. (2013). İlköğretim öğrencilerinin bilimin doğasını anlama düzeylerine bilimsel tartışma odaklı öğretimin etkisi [the effect of scientific discussion-oriented teaching on primary school students' understanding of the nature of science]. *Hacet. Üniv. Eğit. Fak. Derg.* 44, 308–318.
- Van Eemeren, F., Grootendorst, R., and van Eemeren, F. H. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. UK: Cambridge University Press.
- Viyanti, V., Cari, C., Sunarno, W. S., and Prasetyo, Z. K. (2020). Reconstruction of higher order thinking skill through enriching student's argumentation skills. *Jurnal Pendidikan Progresif* 10, 327–335. doi: 10.23960/jpp.v10.i2.202016
- Watson, J. R., Swain, J. R. L., and McRobbie, C. (2004). Students' discussions in practical scientific inquiries. *Int. J. Sci. Educ.* 26, 25–45. doi: 10.1080/0950069032000072764
- Weinstock, M., Neuman, Y., and ve Glassner, A. (2006). Identification of informal reasoning fallacies as a function of epistemological level, grade level, and cognitive ability. *J. Educ. Psychol.* 98, 327–341. doi: 10.1037/0022-0663.98.2.327
- Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *J. Engl. Acad. Purp.* 11, 145–154. doi: 10.1016/j.jeap.2011.11.001
- Wood, N. V. (2001). *Perspectives on argument*. New Jersey: Prentice Hall.
- Yücelşen, N., and ve Edizer, Z. C. (2020). Türkçe öğretmeni adaylarının akademik yazma becerilerine ilişkin görüşleri [Turkish teacher candidates' opinions on academic writing skills]. *Ana Dili Eğitimi Dergisi* 8, 1165–1182. doi: 10.16916/aded.774465
- Yüksek Öğretim Kurulu Başkanlığı (YÖK) (2018). Öğretmen yetiştirme lisans programları [Teacher education undergraduate programs]. Ankara. Available at: https://www.yok.gov.tr/Documents/Kurumsal/egitim_ogretim_dairesi/Yeni-Ogretmen-Yetistirme-Lisans-Programlari/Turkce-Ogretmenligi-Lisans-Programi.pdf
- Zhang, G. (2022). The citational practice of social science research articles: an analysis by part-genres. *J. Engl. Acad. Purp.* 55:101076. doi: 10.1016/j.jeap.2021.101076
- Zohar, A., and ve Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *J. Res. Sci. Teach.* 39, 35–62. doi: 10.1002/tea.10008



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Maria Auxiliadora Robles-Bello,
University of Jaén, Spain
Paula López,
Universidad de León, Spain
Cristian R. Céspedes,
Andres Bello University, Chile

*CORRESPONDENCE

Wilter C. Morales-García
✉ wiltermorales@upeu.edu.pe

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

RECEIVED 15 June 2022

ACCEPTED 14 November 2022

PUBLISHED 06 January 2023

CITATION

León-Gutiérrez M,
Morales-García WC,
Morales-García M, Jaimes-Soncco J,
Saintila J and Mamani PGR (2023)
Psychometric properties and
invariance of the self-efficacy for
writing scale in Peruvian high school
students. *Front. Educ.* 7:969554.
doi: 10.3389/feduc.2022.969554

COPYRIGHT

© 2023 León-Gutiérrez,
Morales-García, Morales-García,
Jaimes-Soncco, Saintila and Mamani.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Psychometric properties and invariance of the self-efficacy for writing scale in Peruvian high school students

Matilde León-Gutiérrez¹, Wilter C. Morales-García^{2*},
Mardel Morales-García¹, Jania Jaimes-Soncco³,
Jacksaint Saintila⁴ and Percy G. Ruiz Mamani⁵

¹Unidad de Posgrado de Educación, Escuela de Posgrado, Universidad Peruana Unión, Lima, Peru, ²Unidad de Posgrado de Salud Pública, Escuela de Posgrado, Universidad Peruana Unión, Lima, Peru, ³Unidad de Posgrado de Psicología, Escuela de Posgrado, Universidad Peruana Unión, Lima, Peru, ⁴Escuela de Medicina Humana, Universidad Señor de Sipán, Chiclayo, Peru, ⁵Escuela Profesional de Enfermería, Facultad de Ciencias de la Salud, Universidad Privada San Juan Bautista, Lima, Peru

Introduction: Being able to write is a key competency in educational models; therefore, it is important to have a scale to assess writing self-efficacy.

Objective: The study evaluated the internal structure, providing evidence of reliability, validity, and factorial invariance of the self-efficacy for writing scale (SEWS) across sex and age in Peruvian schoolchildren.

Methods: An instrumental study was carried out on 1,107 Peruvian adolescents (early, middle, and late) of school age. A confirmatory factor analysis was carried out and the factorial invariance for sex and age was evaluated.

Results: A good fit was obtained for the total sample ($\chi^2 = 1229.08$, $df = 101$, $CFI = 0.997$, $TLI = 0.997$, $RMSEA = 0.057$ [90% CI: 0.0540.059], and $SRMR = 0.029$) and presented good α , ω , and H reliability (>0.89). Correlations between another measure of writing self-efficacy ($p = 0.76$; $p < .01$), study satisfaction ($p = 0.31$; $p < .01$), and burnout ($p = -0.12$; $p < .01$) scales supported convergent and concurrent validity. Multigroup factor analysis supported strict invariance for sex and age, for which the SEWS provides evidence of validity and reliability.

Discussion: This adaptation of the SEWS is a valid, reliable, and invariant measure that can be considered for assessing self-efficacy for writing in Peruvian high school students.

KEYWORDS

self-efficacy, high school, writing, Peru, invariance, validation

Introduction

It is important for adolescents at the school stage to develop a strong sense of confidence in an effective way to communicate through writing (Klassen, 2002). Confident students will have strong self-esteem regarding their writing (Pajares, 2010). Having the skills to perform a complex activity, however, does not ensure adequate performance; when faced with complex tasks, people with higher levels of functioning can have mediocre results if they do not trust their ability to face the challenge.

Self-efficacy affects the development of self-perception of competence by experiencing that the results were successful verbal messages and positive perceptions that encourage and empower (Bandura, 1977). Self-efficacy refers to the perception of personal abilities to perform a given action and makes possible the integration, organization, and application of cognitive, social, and behavioral skills to successfully perform a given task (Bandura, 1982; Pajares, 2002). In learning settings, self-efficacy provides students with: persistence in the face of difficulties, self-confidence, low levels of anxiety and the use of self-regulation and self-learning strategies, which allows greater commitment to tasks (Golparvar and Khafi, 2021). Thus, low self-efficacy for writing in students is related to a delay in the progressive growth of academic programs (Bai and Guo, 2018).

Being able to write is a key competency in educational models (Pajares, 2002). Self-efficacy for writing is the student's perception of his or her ability to write (Volkan and Seçkin, 2019). Previous studies indicate that self-efficacy plays an important role in the use of personal skills to cope with learning tasks. Students with high self-efficacy have greater motivation to learn (Honicke and Broadbent, 2016; Trautner and Schwinger, 2020) and provide solutions to academic problems (Ornelas et al., 2015). Furthermore, students who adopt a strategic study plan and are aware of the writing process (Demir, 2018) improve their willingness to write (Berk and Ünal, 2017) and are effective writers (Cassany, 1995), as opposed to students who have a simple or less sophisticated idea of the writing process (Villalón and Mateos, 2009).

Writing in the early years of adolescence becomes increasingly demanding and complex, requiring greater revision, planning, and self-regulation of the processes (Klassen, 2002). Academic motivation is sometimes considered to decline during adolescence due to hormonal changes, maladaptation among adolescents, social and academic competition, and the school environment (Klassen, 2010). Therefore, the usefulness of the perception of writing and motivation of adolescents is crucial to understand their development, since it is the area where students experience the greatest difficulties (Bruning and Horn, 2000). Thus, students with lower levels of self-efficacy have greater difficulty participating in writing tasks and tend not to persist in the task when they encounter difficulties or failures (Pajares, 2010). In addition, unlike primary school children where basic learning and organization strategies are used, sophisticated cognitive and metacognitive learning strategies are required in adolescence due to a higher expectation of learning (Klassen, 2010). Writing in adolescence facilitates access to information and is a key skill in the construction of knowledge. Therefore, the growing beliefs and motivations allow their development and are part of expressive writing with the social interaction of their emotions. New activities generally produce certain levels of anxiety and stress, which sometimes interfere with self-efficacy in writing (Bandura, 1986; Hooper et al., 2019; Azila-Gbette et al., 2021). Therefore, for Spanish,

an instrument that has the potential to make writing self-efficacy explicit for schoolchildren since this type of belief could be hidden, both for students and teachers. In fact, individuals with a lot of ability to write could be perceived as unprepared to face writing tasks, and this could generate a rejection of this type of task. Whereas when adolescents become aware of their own abilities and reflect on their self-efficacy, they are prepared to make the necessary corrective adjustments (Bandura, 2006; Zumbunn et al., 2020).

Self-efficacy in writing skills predicts academic performance, allows the evaluation of theoretical and applied domains, and provides guidelines and interventions in the results of instruction to improve writing skills in adolescents (Bruning et al., 2013). Therefore, the instrument should efficiently assess perceived writing ability while also allowing the student to make a judgment about what they can do when performing writing tasks. Currently, most of the instruments used in Latin America to measure writing self-efficacy were developed in English and for university students: *Description of Self-Assessment of Writing Measure* (McCarthy et al., 1985), *Writing self-efficacy* (Shell et al., 1989), *Writing Self-Efficacy Scale* (Pajares and Schunk, 2001), and *Writing Goals Items and Subscale* (MacArthur et al., 2016). Similarly, the Self-Efficacy for Writing Scale (SEWS) developed by Bruning et al. (2013) for a school context and its Spanish version applied to a university population (Ramos-Villagrasa et al., 2017) provide sufficient evidence regarding its reliability and validity, whereas other instruments do not report broad validation standards (González et al., 2019). The SEWS uses very readable language that incorporates clear and simple sentences in the first person (Pajares et al., 2001). Its use has made it possible to find the strengths and weaknesses of those being evaluated, making it valuable for guiding student interventions (Bruning et al., 2013; Ramos-Villagrasa et al., 2017). It focuses on the experience of mastery through positive self-perception of skills and neglects negative patterns, considering that the positive perception can be a stimulus and reinforcement, while the negative perception can weaken the beliefs regarding their capabilities (Volkan and Seçkin, 2019). Likewise, the number of SEWS items (16 items) makes it preferable for mass application in educational settings rather than larger scales (Sanders-Reio et al., 2014).

The SEWS version in Spanish (Ramos-Villagrasa et al., 2017), named in the same way as the English version by Bruning et al. (2013), consists of three dimensions that include ideation, conventions, and self-regulation of writing (Ramos-Villagrasa et al., 2017). On the contrary, the initial version (Bruning et al., 2013) was evaluated in high school students in the US through a confirmatory factor analysis (CFA). In comparison with other measures of self-efficacy assessed in adolescents that explore the dimensions of success and failure attributed to various aspects such as luck, effort, and the ability to perform a task (Álvarez-Fernández and García-Sánchez, 2014), the writing self-efficacy model described by Bruning et al.

(2013) maintains harmony with other writing process models that centralize working memory (Hayes, 2006, 2012), writing representations, and writing development (Bruning and Horn, 2000; Schunk and Zimmerman, 2007). Thus, the dimensions of the SEWS are as follows: (a) ideation, which measures cognitive processes, idea generation linked to semantics, and schematic knowledge (Schraw and Egory, 2015); (b) writing conventions, which refers to the articulation of ideas in writing forms and aligned to translation (Hayes, 2006, 2012); (c) self-regulation, which extends from writing activities to their management, monitoring, and evaluation (Zimmerman and Kitsantas, 2006). The proposed dimensions classify the perception of the student's ability to perform tasks in relation to writing. To date, the SEWS has not been validated in the population of high school students for whom it was initially built, and given the lack of validation studies in Spanish high school students, it is imperative to analyze it in this context.

In addition, in the initial study of the SEWS carried out on basic education students (Bruning et al., 2013) and the Spanish version carried out on university students (Ramos-Villagrasa et al., 2017), they show that male students have higher writing self-efficacy than female students. Other studies in school-age populations indicated that female students have higher self-efficacy for writing than male students (Pajares et al., 2007; Andrade et al., 2010). Indeed, female students tend to maintain goals (or mastery), while male students tend to have goals focused on performance (or ego) (Pajares et al., 2000). Women also tend to be more concerned with mastering a writing task than men, who, on average, tend to be more concerned with showing people what they are capable of. That is, women may derive more satisfaction and confidence from self-generated evidence of progress on a writing task, while men seek confirmation of their progress from people around them, including teachers and peers (Andrade et al., 2010). Likewise, age is also important for self-efficacy in writing, since early adolescence (10–13 years old), they experience changes in attitudes and motivational beliefs; however, these changes can be observed in middle adolescence (14 to 16 years old) and late adolescence (17–21 years old), as they achieve greater autonomy in their activities (Klassen, 2002; UNICEF, 2020).

In this case, no study has examined the invariance of SEWS measurement between sex and age. Measurement invariance is necessary to make meaningful comparisons between groups since it tests the equivalence of the meaning of the elements between both groups (Byrne and Stewart, 2009; Inglés et al., 2017). In addition, because psychometric properties may vary, measurement invariance has been considered a prerequisite for making comparisons between groups and examining whether the properties remain invariant, thus making it possible to control for and distinguish differences between groups and avoid concluding erroneous or unfounded data (Hopwood and Donnellan, 2010; Steinmetz, 2019). If the

instrument presented a lack of invariance, the comparisons would be partial and not significant, so the validity of the empirical conclusions would not be guaranteed (Byrne and Stewart, 2009).

Therefore, the main objective of the study was to determine the suitability of the SEWS in its Spanish version for its application in Peruvian school adolescents at the secondary level from first to fifth grade. The following objectives were established: (a) to evaluate the proposed initial structure of the SEWS, through confirmatory factor analysis (CFA), and internal consistency; (b) to examine the convergent validity based on the relationship of the SEWS with another measure of Self-Efficacy for Writing (SEW) and the concurrent validity based on correlations with other measures of Brief Scale of Study Satisfaction (BSSS) and academic burnout; and (c) to evaluate the factorial invariance of SEWS according to schoolchildren, male students, and female students and according to the stages of adolescence (early, middle, and late).

Method

Study design and participants

A methodological study was carried out. To determine the sample size, we analyzed the effect size which considers the number of observed and latent variables in the model, the anticipated effect size ($\lambda = 0.10$), the desired statistical significance ($\alpha = 0.05$), and the statistical power level ($1 - \beta = 0.95$) which considers a recommended minimum sample of 184 participants. The inclusion criteria were as follows: students enrolled in educational institutions who agreed to participate in the research after signing informed consent. A non-probabilistic sample was used for data collection and the exclusion criteria were as follows: students who transferred to other institutions or were discharged, students who do not wish to participate in the research or have unresolved questionnaires, and students with some special educational needs. Furthermore, they were selected using a convenience sampling method. The final sample consisted of 1,107 Peruvian adolescent high school students enrolled in the 2021 period. The proportions according to sex, 47.6% ($n = 527$) were male students and 54.4% ($n = 580$) were female students, with a mean age of 14.43 ($SD = 1.75$) ranging from 11 to 19 years. On the contrary, most of the students were adolescents in the middle phase (36%), in the first grade of secondary school (34.8%), from public schools (89.3%), and from a geographical unit or coastal region of Peru (86.8%). Finally, the students were asked how often they practiced reading; 49.3% of students stated that they read sometimes, 32.2% read almost always, 14.3% always read, and only a smaller proportion (5%) read rarely or never (Table 1).

TABLE 1 Characteristics of the participants.

Characteristic	n	%	
Sex	Male	527	47.6
	Female	580	52.4
Phases			
Adolescence	Early (11–13)	398	36.0
	Mid (14–16)	561	50.7
	Late (17–19)	148	13.4
Study grade	First	385	34.8
	Second	198	17.9
	Third	159	14.4
	Fourth	166	15.0
	Fifth	199	18.0
College	Public	989	89.3
	Private	118	10.7
Region	Coast	961	86.8
	Sierra	70	6.3
	Jungle	76	6.9
How often do you practice reading?	Always	158	14.3
	Almost always	356	32.2
	Sometimes	546	49.3
	Rarely	42	3.8
	Never	5	0.5

Measurements

Self-efficacy for writing scale

The Spanish version of the SEWS was used for the Spanish university population (Ramos-Villagrasa et al., 2017), which consists of three dimensions (see Annex): ideation (items 1–5), linguistic conventions (items 6–10), and self-regulation (items 11–16). Internal consistency using Cronbach's alpha for the Spanish version ranged between 89 and 90 for all dimensions. For this study, a 7-point Likert-type ordinal scale was considered (0 = nothing and 7 = total certainty). The version of this study differs from the original version in that a scale from 0 to 100 is used, justified by the need to standardize and simplify the evaluation measures (Krosnick and Stanley, 2009).

Self-efficacy for writing

This version was adapted to Spanish in its version for high school students (Pérez et al., 2015). The one-dimensional scale allows the evaluation of the perception of the student's effectiveness with respect to writing conventions, based on their beliefs and skills of composition, grammar, use, and mechanics according to their academic level. It contains 10 items on a Likert-type scale from 1 (I'm sure I can't do this activity correctly) to 10 (I'm totally sure I can do this activity correctly), and its internal consistency is $\alpha = 0.83$.

Burnout unique item (IUB)

It is a measure that integrates only one item with five responses to assess academic burnout in students at different levels (Merino-Soto and Fernández-Arata, 2017). It allows you to perceive both mental and physical exhaustion, understood as "burned out" by spending a lot of time studying. The content analysis has been satisfactory both in the clarity of the content and in the ordering of the responses. The intensity rating analysis of the responses was from 1 (minimum perceived intensity) to 5 (maximum perceived intensity).

BSSS

This is a unidimensional measure composed of three items that evaluate the student's satisfaction with his or her way of studying, performance, and overall experience with his or her studies (Merino-Soto et al., 2017). The level of internal consistency of the scale was adequate ($\alpha = 0.78$). The scale presents five response options regarding agreement or disagreement with each of the statements (from strongly disagree to strongly agree).

Procedure and ethical aspects

The study protocol was reviewed and approved by the Research Ethics Committee of the Universidad Peruana Unión (Reference: CE-EPG-000012). Four directors of public and private educational institutions from the three regions of Peru (coast, highlands, and jungle) were contacted. An application was sent, indicating the objective of the study, and authorization was subsequently requested to apply the instrument. Likewise, informed consent was sent through Google forms and social networks (WhatsApp groups and Facebook Messenger accounts) to parents and in the same way to schoolchildren whose parents agreed to participate. The survey was administered to students during their virtual classes. The teaching collaborators informed the purpose of the study and then administered the surveys and answered the questions that arose. Participants were also informed that they could withdraw from the study at any time if they wished. Finally, the study was carried out following the ethical guidelines established in the Declaration of Helsinki, which implies the guarantee of protection of participant's privacy and the confidentiality of personal information, as well as the minimization of the possible effects of the study on the participant's physical, mental, and social apparatus (Puri et al., 2009; AMM, 2013).

Data analysis

Statistical analysis was performed using the free software R 4.1.1 (R Foundation for Statistical Computing, Vienna,

Austria; <http://www.R-project.org>). Descriptive statistics for each SEWS item were performed by calculating the mean, standard deviation, skewness, kurtosis, and corrected inter-test correlation analysis. For skewness and kurtosis, values between -1 and $+1$ were considered adequate (Ferrando and Lorenzo-Seva, 2014). Item-test correlation analysis corrected for item recall in case of $r(i\text{-}tc) \leq 0.2$ or multicollinearity ($i\text{-}tc) \leq 0.2$ was used, and internal consistency was estimated using the ordinal α coefficient (Kline, 2016).

For the confirmatory factor analysis (CFA), it was estimated using the *lavaan* library of the RStudio interface, and the weighted least-square method (WLSM) was used due to the ordinal nature of the items (Brown, 2015), and the mean-variance extracted from the SEWS was calculated. The following indicators were considered for the evaluation of the fit models: the chi-square test (χ^2), confirmatory fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residuals (SRMRs), where CFI and TLI values >0.90 indicate an acceptable fit and >0.95 indicates a good fit (Schumacker and Lomax, 2016). For the RMSEA and SRMR, values below 0.05 indicate a good fit and below 0.8 are considered acceptable (Kline, 2016). The convergent internal validity was estimated by calculating the average variance extracted (AVE), the values indicate that more than 50% of the variance is due to its indicators (Fornell and Larcker, 1981), and the values above 0.50 in the factor loadings (λ) are considered adequate (Dominguez-Lara, 2018). The reliability of the latent model for the total sample and the subsamples of men and women was calculated with the McDonald's ω (McDonald, 1999), and H (Hancock and Mueller, 2001) magnitudes >0.80 are considered adequate (Dominguez-Lara, 2016).

Statistical analysis was performed using the *semTools* package to calculate invariance. Factorial invariance was evaluated according to sex (males vs. females) and stages of the early, middle, and late adolescence of the participants through a sequence of hierarchical models, evaluating the most restricted CFA (Chen, 2007). The first level, configural invariance (M1), evaluates the reference model; the second level, metric invariance (M2), evaluates the equality of factor loading; the third level, scalar invariance (M3), evaluates the equality of factor loading and intersection; and the fourth level, scalar invariance (M4), evaluates the equality of factor loadings, intersections, and residuals. Because χ^2 ($\Delta\chi^2$) is sensitive to sample size, the absolute difference CFI (ΔCFI) value was used, where values <0.01 indicate that the factor structure is invariant between the groups compared (Cheung and Rensvold, 2009). For evidence of validity based on the relationship with other SEWS variables at the latent level, a structural equation model was performed with the factors of self-efficacy for writing, satisfaction with life, satisfaction with studies, self-efficacy for studying, and academic burnout.

TABLE 2 Descriptive statistics and reliability of the SEWS.

Items	M	SD	g^1	g^2	r.cor	α ordinal
ID 1	5.03	1.31	−0.44	−0.13	0.79	0.95
ID 2	5.01	1.35	−0.44	−0.32	0.77	0.95
ID 3	5.04	1.39	−0.48	−0.33	0.75	0.95
ID 4	4.95	1.43	−0.44	−0.42	0.73	0.95
ID 5	4.86	1.39	−0.30	−0.51	0.80	0.95
CO 6	4.55	1.41	−0.16	−0.63	0.65	0.95
CO 7	5.11	1.36	−0.48	−0.33	0.78	0.95
CO 8	4.89	1.43	−0.38	−0.52	0.73	0.95
CO 9	4.60	1.36	−0.17	−0.60	0.82	0.95
CO 10	4.97	1.37	−0.39	−0.35	0.76	0.95
SR 11	4.98	1.50	−0.43	−0.56	0.75	0.95
SR 12	4.59	1.57	−0.25	−0.70	0.70	0.95
SR 13	4.88	1.48	−0.40	−0.47	0.71	0.95
SR 14	4.77	1.56	−0.39	−0.56	0.67	0.95
SR 15	4.98	1.39	−0.45	−0.22	0.79	0.95
SR 16	4.91	1.46	−0.46	−0.34	0.72	0.95

ID, ideation; CO, conventions; SR, self-regulation; M, mean; SD, standard deviation; g^1 , asymmetry; g^2 , Kurtosis.

Results

Descriptive statistics of the SEWS items

Descriptive statistics were analyzed in the total sample (Table 2), where the highest mean was found in item 7 ($M = 5.11$; $SD = 1.36$), which belongs to the language convention dimension, and the lowest mean was observed in item 12 ($M = 4.59$; $SD = 1.57$) corresponding to the self-regulation dimension. Skewness (g^1) and kurtosis (g^2) values were $< \pm 1$ for all items. All corrected item-total correlations were greater than the acceptable limit of 0.3, indicating that each item is related to the overall scale.

Evidence of validity and reliability

The confirmatory factor analysis was used to evaluate the three-dimensional structure derived from the hypothetical structure of the original version (Bruning et al., 2013; Ramos-Villagrasa et al., 2017). The goodness-of-fit indices for the total sample were $\chi^2 = 1,229.08$, $df = 101$; CFI = 0.997, TLI = 0.997, RMSEA = 0.057 (90% CI: 0.054–0.059), and SRMR = 0.029, indicating that the Peruvian version model fits the observed data adequately. In addition, all λ were >0.71 and the AVE values are adequate ($AVE > 0.50$), indicating that the latent factors are adequately explained by their observed variables, and α , ω , and H coefficients are shown in Table 3.

TABLE 3 Factor loadings, goodness-of-fit index, and reliability of the total sample, men, and women.

	F ₁ (λ)	F ₂ (λ)	F ₃ (λ)
ID 1	0.83		
ID 2	0.82		
ID 3	0.80		
ID 4	0.77		
ID 5	0.84		
CO 6		0.69	
CO 7		0.82	
CO 8		0.79	
CO 9		0.87	
CO 10		0.81	
SR 11			0.79
SR 12			0.74
SR 13			0.75
SR 14			0.71
SR 15			0.83
SR 16			0.76
AVE	0.66	0.64	0.58
α	0.90	0.89	0.89
ω	0.91	0.90	0.89
H	0.91	0.91	0.90
χ^2	1,229.08		
df	101		
<i>p</i>	<0.001		
CFI	0.997		
TLI	0.997		
RMSEA	0.057		
90% CI	0.054–0.059		
SRMR	0.029		

ID, ideation; CO, conventions; SR, self-regulation; F1= factor ideation; F2, conventions; F3, self-regulation; λ , factor loadings. AVE, average variance extracted; α , ordinal alpha coefficient; ω , coefficient omega; H, coefficient H.
 χ^2 , chi-square; df, degrees of freedom; RMSEA, root mean square error of approximation; CFI, comparative fit index; SRMR, standardized root mean square residual.

Factorial invariance

Comparisons of the factorial invariance models by sex and stages of adolescence are shown in Table 4. Again, it is shown that the three-factor structure for the groups of men and women was adequate. Furthermore, the three-factor model was suitable for all three phases: (1) early (11–13), (2) middle (14–16), and (3) late (17–19). The configurational invariance model (M1) presented a good fit for both groups. This model serves as a baseline for the evaluation of others. Subsequently, the metric invariance (M2) in the groups was evaluated and presented as a good fit for the data. Once the metric invariance was established, the scalar invariance (M3) was evaluated. Finally, strict invariance (M4) was evaluated in the groups whose

values were significant, the Δ CFI coefficients were <0.01, and Δ RMSEA coefficients were <0.01. Thus, the models M1, M2, M3, and M4 are within the expected range, which confirms the factorial invariance of SEWS and the different measures can be compared in the sex and age groups.

Convergent and concurrent validity

This type of validity was examined, and a structural equation model was evaluated where four latent variables are modeled: self-efficacy for writing, satisfaction with life, satisfaction with studies, and academic burnout. The model had a good fit: $\chi^2 = 2,008.292$, $df = 400$; CFI = 0.993, TLI = 0.992, RMSEA = 0.035 (90% CI: 0.034–0.037), and SRMR = 0.041. As expected, the SEWS significantly correlated positively with the other measure of writing self-efficacy (SEW) ($p = 0.76$; $p < 0.01$) and BSSS ($p = 0.31$; $p < 0.01$) and negatively with burnout ($p = -0.12$; $p < 0.01$), evidencing acceptable convergent and concurrent validity (Figure 1).

Discussion

The objective of the present research was to analyze the factorial structure, as well as the reliability and adequate psychometric properties. The study suggests a promising instrument for assessing writing self-efficacy in Peruvian schoolchildren. The CFA supports the three-factor structure proposed in the original version. Likewise, the results supported reliability, validity (convergent, concurrent, and discriminant), and the invariance of the measurement at the strict level between sex and age.

The CFA was performed, which supported the three-factor structure of the original questionnaire, so the Peruvian version maintains the same factors as the original questionnaire in English (Bruning et al., 2013) and the Spanish version (Ramos-Villagrasa et al., 2017). The CFA provided a good fit to the model data, thus showing that the 16-item version of the SEWS classified on the dimensions of ideation, language conventions, and meaningful theory-based self-regulation adequately assessed writing-related tasks. In the first dimension of the Flower and Hayes model of writing, students generate ideas from thematic and world knowledge. The second dimension is linked to writing, having the ability to successfully express ideas in linguistic forms. Finally, self-regulation is part of the value judgments you have about yourself. Therefore, the instrument seems suitable for future applications with other multidimensional constructions associated with the wellbeing of schoolchildren (Bruning et al., 2013).

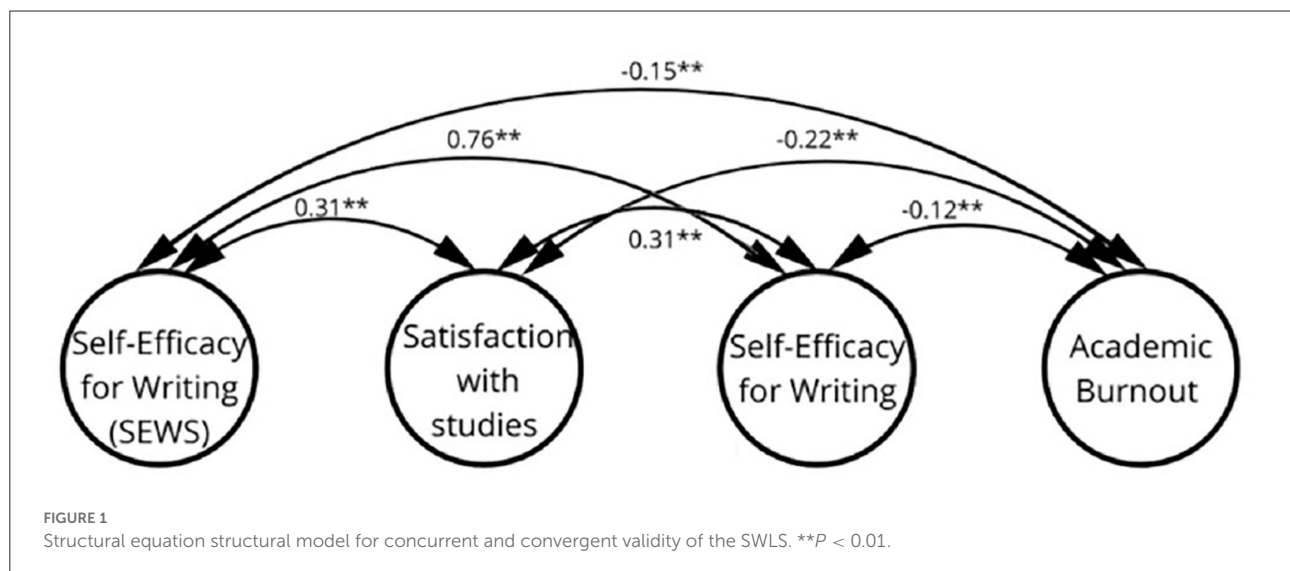
The results also showed good internal consistency. Reliability, based on ordinal alpha, was acceptable for the total scale and the three factors, with a range of 0.89–0.90. Reliability

TABLE 4 Measurement invariance between sex and age groups.

Groups	χ^2	df	RMSEA	[IC 90%]	p	SRMR	TLI	CFI	Δ CFI	Δ RMSEA
Sex										
M1	197.888	124	0.017	0.012–0.021	0.000	0.025	0.999	0.999	-	-
M2	215.011	134	0.020	0.015–0.024	0.000	0.031	0.998	0.998	0.001	-0.003
M3	226.244	144	0.019	0.014–0.024	0.000	0.031	0.998	0.998	0.000	0.001
M4	245.636	157	0.020	0.015–0.024	0.000	0.033	0.998	0.998	0.000	-0.001
Phases										
M1	253.371	186	0.015	0.000–0.020	0.000	0.027	0.999	0.999	-	-
M2	255.110	206	0.016	0.000–0.022	0.000	0.033	0.999	0.999	0.000	-0.001
M3	313.831	226	0.020	0.014–0.025	0.000	0.035	0.998	0.998	0.001	-0.004
M4	339.486	252	0.019	0.014–0.024	0.000	0.038	0.998	0.998	0.000	-0.001

M1, configural; M2, metric; M3, scalar; M4, strict.

χ^2 , chi-square; df, degrees of freedom; RMSEA, root mean square error of approximation; CFI, comparative fit index; Δ CFI, comparative fit index difference; Δ RMSEA, root mean square error of approximation difference.



was also consistent with the original version, indicating a good ability to similarly assess SEWS (Ramos-Villagrasa et al., 2017). Unlike other studies, we calculated the coefficient ω reflecting the proportion of variance in the scale scores associated with an overall factor (Zinbarg et al., 2005) and the H coefficient that evaluates the reliability of the construct, reflecting the influence of the construct in the overall model and subsamples. The greater its magnitude, the better it is represented (Dominguez-Lara, 2016). These coefficients are considered better estimators than alpha, which tends to underestimate reliability. The values of the corrected item-total correlations were good, indicating adequate homogeneity.

Convergent internal validity and validity in relation to other measures and concurrent validity were examined. Regarding the internal convergent validity of the SEWS, it showed adequate factor loadings ($\lambda > 0.50$) and an acceptable AVE (AVE > 0.50)

in the three models (Fornell and Larcker, 1981). In addition, the related convergent validity of the SEWS was assessed with a homologous one-dimensional scale of Self-Efficacy for Writing (SEW) by Pajares et al. (2001) that focuses on the linguistic conventions of self-efficacy, and the results indicated an adequate correlation between the scores of each test, despite the fact that SEW by Pajares et al. (2001) only evaluates one dimension of the SEWS. Likewise, for concurrent validity, the latent variable of the SEWS and BSSS were correlated, indicating statistically significant and positive correlations.

Self-efficacy is an important element in how students feel about themselves in general (Sabouripour et al., 2021). Self-efficacy is also a universal psychological need that controls an individual's cognitive aspect, emotions, and decisions related to psychological wellbeing (Bartimote-Aufflick et al., 2016). Likewise, student self-efficacy is a fundamental variable in

the satisfaction of schoolchildren with their studies and, consequently, in the subjective results of education (Wach et al., 2016). Since it increases positive emotional states, it contributes to wellbeing and improved academic performance (Bresó et al., 2011). In the current study, a positive relationship was found between self-efficacy and satisfaction with studies. In another study, it is stated that high self-efficacy positively impacts people's wellbeing (Pajares and Schunk, 2001). Furthermore, it is likely that self-efficacy influences the amount of stress and anxiety that students experience when participating in a course (Doménech-Betoret et al., 2017). There is a large body of empirical literature that evidences the positive effects of self-efficacy on student wellbeing and study satisfaction (DeWitz and Walsh, 2002). A model has even been proposed in which beliefs of high efficacy in studies constitute a determinant of satisfaction with studies since they make possible a better adjustment between the level of demand of the task and the abilities that the student perceives he/she has to face it (Bebermeier et al., 2022). Bebermeier et al. (2022), this has been documented in some studies detailing the relationship between self-efficacy and satisfaction with studies (DeWitz and Walsh, 2002; Shehadeh et al., 2020). However, the mechanisms mediating the relationship between self-efficacy and student satisfaction in studies need to be further studied. Understanding these motivational mechanisms is crucial for implementing promotion programs that increase student satisfaction.

On the contrary, concurrent validity between SEWS and academic burnout indicated a negative relationship; previous studies have also identified a negative relationship between stress and self-efficacy (Zajacova et al., 2005; Rayan, 2018), taking into account that when a student perceives that he/she has fewer personal resources to face a task, he/she will seek to avoid and put less effort into solving that task, which will end up worsening the academic situation and generating greater stress (Krumrei-Mancuso et al., 2013).

The main advantage of the SEWS is that it allows the assessment of the three different areas of writing by finding the weaknesses and strengths of each student and allows for interventions at the individual level. Furthermore, in relation to the test scores of the same construct, they showed that all SEWS dimensions were positively correlated. Previous studies have shown high associations between ideation and self-regulation dimensions (Bruning et al., 2013). As indicated, individuals with high levels of ideation linked to cognitive processes in idea generation, systematic domains, and schematic knowledge may exhibit high levels of self-regulation that extend well beyond writing activities by having ideas to write about and mastery in the conventions of writing (Schraw, 2006; Zimmerman and Kitsantas, 2007; Bruning et al., 2013).

Discriminant validity results showed that all SEWS dimensions were negatively correlated with burnout, which is consistent with previous studies in other populations

(Hall et al., 2019). Therefore, it is evident that the two scales are sufficiently independent and also measure separate constructs (Eignor, 2013). Previous studies suggest that the lower the feeling of self-efficacy, the more burnout one experiences (Cappe et al., 2021). Given the results of writing self-efficacy correlating negatively with a burnout in the current study, adolescent assessments could be conducted to assess the purpose, emotional reactions, effort, coping, and endurance. In this way, it is necessary to examine the effectiveness of interventions in coping with difficult tasks and activities or understanding problems that lead to the development of stress, depression, and weak problem-solving that allow continuous improvement in students (Rahmati, 2015).

In addition, the factorial invariance of the SEWS measure was reported for the first time in a sample composed of school adolescents, male students, and female students, and in the early, middle, and late stages. The configurational, metric, scalar, and strict invariance of SEWS were acceptable in the present study, indicating that it can be evaluated with the same accuracy in groups of female and male adolescents. This analysis is important, considering the psychological characteristics of male and female populations that can differentially affect their behavior (Hyde, 2005). Thus, findings from initial studies conducted with school students found that female students reported lower self-efficacy for male writing (Bruning et al., 2013). This is consistent with the Spanish study on university students which suggests that male students report higher self-efficacy than female students (Ramos-Villagrasa et al., 2017). However, other studies in school populations show the opposite, where boys are the ones who reach lower levels of self-efficacy, mainly due to stereotyped beliefs in the socialization process (Pajares and Valiente, 2001), while other findings do not show any difference (Pajares et al., 2001). Likewise, written self-efficacy may differ partially in the context according to the stages of school adolescence, given that in the early stage self-efficacy beliefs are less established (Klassen, 2002), while in middle and late adolescence it allows greater external and internal regulation (Cattelino et al., 2019). Moreover, future studies should explore the strict invariance of SEWS in other populations, in addition to exploring other groups, such as socioeconomic status, cultures, and clinical groups.

Limitations

The results showed appropriate psychometric properties for the Peruvian version of the SEWS. However, some limitations are considered. On the contrary, the study was cross-sectional and did not consider a longitudinal design, which prevented us from evaluating causal relationships between the variable's satisfaction with studies, self-efficacy for writing, and academic burnout. Likewise, the self-report techniques used in the study may be influenced by social desirability, introspection, and

memory, among other biases. Test–retest reliability was also not examined; therefore, its incorporation in future studies is recommended.

Conclusion

The validity of the internal structure of the SEWS was satisfactory, and a three-factor structure like the original one was determined, with adequate and stable psychometric properties. Strict factorial invariance was demonstrated for sex and age, which is an important contribution to the measurement of adolescence. Therefore, the SEWS is a valid and reliable measure of writing self-efficacy in the Peruvian school context.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The study was approved by the Ethics Committee of the Universidad Peruana Unión (CE-EPG-000012). Written informed consent to participate in this study was provided by the participants' or their legal guardian/next of kin.

References

- Álvarez-Fernández, M. L., and García-Sánchez, J. N. (2014). Development of the writing process from primary to secondary education. *Revista de Psicodidáctica*. 19, 5–26. doi: 10.1387/RevPsicodidact.6518
- AMM (2013). *Declaración de Helsinki de la AMM - Principios éticos para investigación médica en seres humanos*. Asamblea Médica Mundial. Available online at: <https://www.wma.net/es/politicas-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/> (accessed April 2, 2022).
- Andrade, H., Wang, X., Du, Y., and Akawi, R. (2010). Rubric-referenced self-assessment and self-efficacy for writing. *J. Educ. Res.* 102, 287–302. doi: 10.3200/JOER.102.4.287-302
- Azila-Gbetor, E. M., Mensah, C., Abiemo, M. K., and Bokor, M. (2021). Predicting student engagement from self-efficacy and autonomous motivation: a cross-sectional study. *Cogent Education*. 8, 1. doi: 10.1080/2331186X.2021.1942638
- Bai, B., and Guo, W. (2018). Influences of self-regulated learning strategy use on self-efficacy in primary school students' english writing in Hong Kong. *Read Writ Q.* 34, 523–536. doi: 10.1080/10573569.2018.1499058
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioural change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *Am. Psychol.* 37, 122–147. doi: 10.1037/0003-066X.37.2.122
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. New Jersey, NJ: Prentice Hall.
- Bandura, A. (2006). "Adolescent development from an agentic perspective," in *Self-Efficacy Beliefs of Adolescents*, eds F. Pajares and T. Urdan (Greenwich, CT: Information Age), 169.
- Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., and Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Stud. Higher Edu.* 41, 1918–1942. doi: 10.1080/03075079.2014.999319
- Bebermeier, S., Austers Schmidt, K. L., and Nussbeck, F. W. (2022). Determinants of psychology students' study satisfaction. *Psychol. Learn. Teach.* 21, 19–36. doi: 10.1177/1475725720985223
- Berk, R. R., and Ünal, E. (2017). Comparison of writing anxiety and writing dispositions of sixth, seventh and eighth grade students. *Int. J. Instr.* 10, 237–254. doi: 10.12973/iji.2017.10115a
- Bresó, E., Schaufeli, W. B., and Salanova, M. (2011). Can a self-efficacy-based intervention decrease burnout, increase engagement, and enhance performance? A quasi-experimental study. *Higher Edu.* 61, 339–355. doi: 10.1007/s10734-010-9334-6
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*, 2nd Edn. New York, NY: Guilford Press.
- Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., and Zumbunn, S. (2013). Examining dimensions of self-efficacy for writing. *J. Educ. Psychol.* 105, 25–38. doi: 10.1037/a0029692
- Bruning, R., and Horn, C. (2000). Developing motivation to write. *Educ. Psychol.* 35, 25–37. doi: 10.1207/S15326985EP3501_4
- Byrne, B. M., and Stewart, S. M. (2009). Teacher's corner: the MACS approach to testing for multigroup invariance of a second-order structure: a walk through the process. *Struct. Equ. Modeling*. 13, 287–321. doi: 10.1207/s15328007sem1302_7
- Cappe, E., Poirier, N., Engelberg, A., and Boujut, E. (2021). Comparison of teachers in France and in Quebec working with autistic students: self-efficacy, stress, social support, coping, and burnout. *Teach. Teach. Educ.* 98, 103244. doi: 10.1016/j.tate.2020.103244
- Cassany, D. (1995). *La cocina de la escritura*. Anagrama.
- Cattellino, E., Morelli, M., Baiocco, R., and Chirumbolo, A. (2019). From external regulation to school achievement: the mediation of self-efficacy at school. *J. Appl. Dev. Psychol.* 60, 127–133. doi: 10.1016/j.appdev.2018.09.007
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Modeling*. 14, 464–504. doi: 10.1080/10705510701301834

Author contributions

WM-G and ML-G participated in the conceptualization. PM, JJ-S, MM-G, and JS were in charge of the methodology and software. WM-G, MM-G, and PM validated the study, involved in formal analysis, and performed the research. WM-G and ML-G were involved in the curation of data and collected resources. WM-G, JS, and MM-G were involved in first draft writing—review and editing, visualization, and supervision. All authors have read and approved the final version of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Cheung, G. W., and Rensvold, R. B. (2009). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Modeling*, 9, 233–255. doi: 10.1207/S15328007SEM0902_5
- Demir, S. (2018). The relation between the awareness level of writing strategies and writing. *Self-Efficacy*, 11, 59–73. doi: 10.5539/ies.v11n5p59
- DeWitz, S. J., and Walsh, W. B. (2002). Self-efficacy and college student satisfaction. *J. Career Assess.*, 10, 315–326. doi: 10.1177/10672702010003003
- Doménech-Betoret, F., Abellán-Rosell, L., and Gómez-Artiga, A. (2017). Self-efficacy, satisfaction, and academic achievement: the mediator role of students' expectancy-value beliefs. *Front. Psychol.*, 8, 1193. doi: 10.3389/fpsyg.2017.01193
- Dominguez-Lara, S. (2018). Proposal for cut-offs for factor loadings: a construct reliability perspective. *Enfermería Clínica (English Edition)*, 28, 401–402. doi: 10.1016/j.enfcl.2018.06.006
- Dominguez-Lara, S. A. (2016). Evaluación de la confiabilidad del constructo mediante el Coeficiente H: breve revisión conceptual y aplicaciones. *Psychologia*, 10, 87–94. doi: 10.21500/19002386.2134
- Eignor, D. R. (2013). "The standards for educational and psychological testing," in *APA Handbook of Testing and Assessment in Psychology, Vol. 1: Test Theory and Testing and Assessment in Industrial and Organizational Psychology*, eds K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, and M. C. Rodriguez (American Psychological Association). doi: 10.1037/14047-013
- Ferrando, P. J., and Lorenzo-Seva, U. (2014). El análisis factorial exploratorio de los ítems: algunas consideraciones adicionales. *Anales de Psicología*, 30, 1170–1175. doi: 10.6018/analesps.30.3.199991
- Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.*, 18, 382–388. doi: 10.1177/002224378101800313
- Golparvar, S. E., and Khafi, A. (2021). The role of L2 writing self-efficacy in integrated writing strategy use and performance. *Assess. Writing*, 47, 100504. doi: 10.1016/j.asw.2020.100504
- González, M., Meza, P., and Castellón, M. (2019). Medición de la Autoeficacia para la Escritura Académica. *Una Revisión Teórico-Bibliográfica. Formación Universitaria*, 12, 191–204. doi: 10.4067/S0718-50062019000600191
- Hall, N. C., Lee, S. Y., and Rahimi, S. (2019). Self-efficacy, procrastination, and burnout in post-secondary faculty: an international longitudinal analysis. *PLoS ONE*, 14, e0226716. doi: 10.1371/journal.pone.0226716
- Hancock, G. R., and Mueller, R. O. (2001). "Rethinking construct reliability within latent variable systems," in *Structure Equation Modeling: Present and Future Festschrift in Honor of Karl Joreskog*, eds R. Cudeck, S. D. Toit, and D. Soerboom (Lincolnwood, IL: Scientific Software International), 195216.
- Hayes, J. R. (2006). "New directions in writing theory," in *Handbook of Writing Research* (New York, NY: Guilford Press).
- Hayes, J. R. (2012). "Evidence from language bursts, revision, and transcription for translation and its relation to other writing processes," in *Translation of Thought to Written Text While Composing: Advancing Theory, Knowledge, Research Methods, Tools, and Applications* (New York, NY: Psychology Press), 1525.
- Honick, T., and Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: a systematic review. *Edu. Res. Rev.*, 17, 63–84. doi: 10.1016/j.edurev.2015.11.002
- Hooper, S. R., Montgomery, J. W., Brown, T. T., Swartz, C. W., Reed, M. S., Wasileski, T. J., et al. (2019). Prevalence of writing problems across three middle school samples. *School Psych. Rev.*, 22, 610–622. doi: 10.1080/02796015.1993.12085677
- Hopwood, C. J., and Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Pers. Soc. Psychol. Rev.*, 14, 332–346. doi: 10.1177/1088868310361240
- Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.*, 60, 581–592. doi: 10.1037/0003-066X.60.6.581
- Inglés, C. J., Marzo, J. C., Hidalgo, M. D., Zhou, X., and García-Fernández, J. M. (2017). Factorial invariance of the questionnaire about interpersonal difficulties for adolescents across Spanish and Chinese adolescent samples. *Meas. Eval. Couns. Dev.*, 41, 89–103. doi: 10.1080/07481756.2008.11909824
- Klassen, R. (2002). Writing in early adolescence: a review of the role of self-efficacy beliefs. *Educ. Psychol. Rev.*, 14, 173–203. doi: 10.1023/A:1014626805572
- Klassen, R. (2010). Confidence to manage learning: the self-efficacy for self-regulated learning of early adolescents with learning disabilities. *Learn. Disabil. Q.*, 33, 19–30. doi: 10.1177/073194871003300102
- Kline, R. B. (2016). *Principles and Practice of Structure Equation Modeling, 4th Edn.* Guilford Press.
- Krosnick, J. A., and Stanley, P. (2009). "Question and questionnaire design," in *Handbook of Survey Research, 2nd Edn.*, eds J. A. Krosnick, J. D. Wright, and P. V. Marsden (San Diego, CA: Elsevier).
- Krumrei-Mancuso, E. J., Newton, F. B., Kim, E., and Wilcox, D. (2013). Psychosocial factors predicting first-year college student success. *J. Coll. Stud. Dev.*, 54, 247–266. doi: 10.1353/csd.2013.0034
- MacArthur, C. A., Philippakos, Z. A., and Graham, S. (2016). A multicomponent measure of writing motivation with basic college writers. *Learn. Disabil. Q.*, 39, 31–43. doi: 10.1177/0731948715583115
- McCarthy, P., Meier, S., and Rinderer, R. (1985). Self-efficacy and writing: a different view of self-evaluation. *Coll. Compos. Commun.*, 36, 465. doi: 10.2307/357865
- McDonald, R. P. (1999). *Test Theory: A United Treatment*. New York, NY: Lawrence Erlbaum.
- Merino-Soto, C., Dominguez-Lara, S., and Fernández-Arata, M. (2017). Validación inicial de una Escala Breve de Satisfacción con los Estudios en estudiantes universitarios de Lima. *Educación Médica*, 18, 74–77. doi: 10.1016/j.edumed.2016.06.016
- Merino-Soto, C., and Fernández-Arata, M. (2017). Single-item burnout in higher education students: Content validity study. *Edu. Médica*, 18, 195–198. doi: 10.1016/j.edumed.2016.06.019
- Ornelas, M., Blanco, H., Viciano, J., and Rodríguez, J. M. (2015). Percepción de autoeficacia en la solución de problemas y comunicación científica en universitarios de ingeniería y ciencias sociales. *Formación Universitaria*, 8, 93–100. doi: 10.4067/S0718-50062015000400011
- Pajares, F. (2002). *Overview of Social Cognitive Theory and of Self-Efficacy*. Available online at: <http://www.emory.edu/EDUCATION/mfp/eff.html> (accessed April 2, 2022).
- Pajares, F. (2010). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Read Writ. Q.*, 19, 139–158. doi: 10.1080/10573560308222
- Pajares, F., Britner, S. L., and Valiante, G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemp. Educ. Psychol.*, 25, 406–422. doi: 10.1006/ceps.1999.1027
- Pajares, F., Hartley, J., and Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Meas. Eval. Couns. Dev.*, 33, 214–221. doi: 10.1080/07481756.2001.12069012
- Pajares, F., Johnson, M. J., and Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Res. Teach. English*, 42, 104–120.
- Pajares, F., and Schunk, D. (2001). "Self-beliefs and school success: self-efficacy, self-concept, and school achievement," in *International Perspectives on Individual Differences: Self-Perception* Riding, R. J., and Rayner, S. G. (Eds.). p. 239–265.
- Pajares, F., and Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: a function of gender orientation? *Contemp. Educ. Psychol.*, 26, 366–381. doi: 10.1006/ceps.2000.1069
- Pérez, E., Medrano, L., Mattus, J., and Ayllón, S. (2015). *Adaptación de escalas de autoeficacia para escritura y Lengua*. Perspectivas en Psicología y Ciencias Afines. 5, 86–92.
- Puri, K. S., Suresh, K. R., Gogtay, N. J., and Thatte, U. M. (2009). Declaration of Helsinki, 2008: implications for stakeholders in research. *J. Postgrad. Med.*, 55, 131–134. doi: 10.4103/0022-3859.52846
- Rahmati, Z. (2015). The study of academic burnout in students with high and low level of self-efficacy. *Procedia Soc. Behav. Sci.*, 171, 49–55. doi: 10.1016/j.sbspro.2015.01.087
- Ramos-Villagrasa, P. J., Sánchez-Iglesias, I., Grande-de-Prado, M., Oliván-Blázquez, B., Martín-Peña, J., and Cancer-Lizaga, P. (2017). Spanish version of "Self-Efficacy for Writing Scale" (SEWS). *Anales de Psicología*, 34, 1. doi: 10.6018/analesps.34.1.264931
- Rayan, A. (2018). Mindfulness, self-efficacy, and stress among final-year nursing students. *J. Psychosoc. Nurs. Ment. Health Serv.*, 57. doi: 10.3928/02793695-20181031-01
- Sabouripour, F., Roslan, S., Ghiami, Z., and Memon, M. A. (2021). Mediating role of self-efficacy in the relationship between optimism, psychological well-being, and resilience among Iranian students. *Front. Psychol.*, 12, 675645. doi: 10.3389/fpsyg.2021.675645
- Sanders-Reio, J., Alexander, P. A., Reio, T. G., and Newman, I. (2014). Do students' beliefs about writing relate to their writing self-efficacy, apprehension, and performance? *Learn. Instruct.*, 33, 1–11. doi: 10.1016/j.learninstruct.2014.02.001

- Schraw, G. (2006). "Knowledge: structures and processes," in *Handbook of Educational Psychology*, eds P. A. Alexander and P. H. Winne (New York, NY: Routledge).
- Schraw, G., and Egory. (2015). "Knowledge: Structures and Processes," in *Handbook of Educational Psychology*.
- Schumacker, R. E., and Lomax, R. G. (2016). *A Beginners Guide to Structure Equation Modeling, 4th Edn.* New York, NY: Routledge. doi: 10.4324/9781315749105
- Shunk, D. H., and Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Read Writ. Q.* 23, 7–25. doi: 10.1080/10573560600837578
- Shehadeh, J., Hamdan- Mansour, A., Halasa, S., Hani, M., Nabolsi, M., Thultheen, I., et al. (2020). Academic stress and self-efficacy as predictors of academic satisfaction among nursing students. *Open Nurs. J.* 14, 92–99. doi: 10.2174/1874434602014010092
- Shell, D. F., Murphy, C. C., and Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *J. Educ. Psychol.* 81, 91–100. doi: 10.1037/0022-0663.81.1.91
- Steinmetz, H. (2019). "Estimation and comparison of latent means across cultures," in *Cross-Cultural Analysis*. Oxford: Routledge. p. 95–126. doi: 10.4324/9781315537078-4
- Trautner, M., and Schwinger, M. (2020). Integrating the concepts self-efficacy and motivation regulation: how do self-efficacy beliefs for motivation regulation influence self-regulatory success? *Learn. Individ. Differ.* 80, 101890. doi: 10.1016/j.lindif.2020.101890
- UNICEF (2020). 'Qué es la adolescencia?' Available online at: <https://www.unicef.org/uruguay/que-es-la-adolescencia> (accessed April 2, 2022).
- Villalón, R., and Mateos, M. (2009). Concepciones del alumnado de secundaria y universidad sobre la escritura académica. *Infancia y Aprendizaje*. 32, 219–232. doi: 10.1174/021037009788001761
- Volkan, M. D., and Seçkin, I. A. (2019). Examining high school students' writing self-efficacy perceptions. *Int. J. Educ. Res.* 7, 106. doi: 10.7575/aiac.ijels.v.7n.4p.106
- Wach, F. S., Karbach, J., Ruffing, S., Brünken, R., and Spinath, F. M. (2016). University students' satisfaction with their academic studies: personality and motivation matter. *Front. Psychol.* 7, 55. doi: 10.3389/fpsyg.2016.00055
- Zajacova, A., Lynch, S., and Espenshade, T. (2005). Self-efficacy, stress, and academic success in college. *Res. High. Educ.* 46, 677–706. doi: 10.1007/s11162-004-4139-z
- Zimmerman, B. J., and Kitsantas, A. (2006). "A writer's discipline: The development of self-regulatory skill," in *Writing and Motivation*. Leiden, Netherlands: Brill. p. 51–69. doi: 10.1163/9781849508216_005
- Zimmerman, B. J., and Kitsantas, A. (2007). "A writer's discipline: The development of self-regulatory skill," in Hidi, S., and Boscolo, P. (Eds.). *Motivation and writing: Research and school practice*. Alphen aan den Rijn, Netherlands: Kluwer. p. 51–69.
- Zinbarg, R. E., Revelle, W., Yovel, I., and Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω H : their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*. 70, 123–133. doi: 10.1007/s11336-003-0974-7
- Zumbrunn, S., Broda, M., Varier, D., and Conklin, S. (2020). Examining the multidimensional role of self-efficacy for writing on student writing self-regulation and grades in elementary and high school. *Br. J. Educ. Psychol.* 90, 580–603. doi: 10.1111/bjep.12315



OPEN ACCESS

EDITED BY

Olivia López Martínez,
University of Murcia,
Spain

REVIEWED BY

Nesrin Menemenci Bahceleri,
Near East University,
Cyprus
Fong Peng Chew,
University of Malaya,
Malaysia

*CORRESPONDENCE

Isolda Margarita Castillo-Martínez
✉ isoldamcm@hotmail.com

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 28 September 2022

ACCEPTED 21 December 2022

PUBLISHED 16 January 2023

CITATION

Castillo-Martínez IM, Cerros Regalado CP,
Glasserman-Morales LD and
Ramírez-Montoya MS (2023) Academic
literacy among the university students in
Mexico and Spain: A holistic perspective.
Front. Psychol. 13:1055954.
doi: 10.3389/fpsyg.2022.1055954

COPYRIGHT

© 2023 Castillo-Martínez, Cerros Regalado,
Glasserman-Morales and Ramírez-
Montoya. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Academic literacy among the university students in Mexico and Spain: A holistic perspective

Isolda Margarita Castillo-Martínez^{1*},
Cynthia Patricia Cerros Regalado²,
Leonardo David Glasserman-Morales¹ and
María Soledad Ramírez-Montoya³

¹School of Humanities and Education, Tecnológico de Monterrey, Nuevo León, Mexico, ²Business School, Tecnológico de Monterrey, Ciudad Juárez, Chihuahua, Mexico, ³Institute for the Future of Education, Tecnológico de Monterrey, Nuevo León, Mexico

Introduction: Currently, young people have access to a large amount of information, so they must have the ability to critically analyze the texts they are exposed to in order to choose those that are useful for their training or research process, as well as to have the necessary skills to interact efficiently with the texts, especially with those specialized documents corresponding to their area of study. In this regard, this article aims to identify how cognitive, emotional, attitudinal, digital and personality aspects influence the development of academic literacy skills in university students.

Method: Research was conducted with the mixed method, in which quantitative instruments were applied and analyzed: a Likert scale questionnaire to measure the perceived level of mastery of academic literacy skills, which was applied to a sample of 595 students from higher education institutions in Mexico and Spain. In addition, a test was applied to identify personality type. For the qualitative part, the case study was chosen and the qualitative instruments applied were interviews with a teacher and a student and a focus group with five students.

Results: The findings identified were as follows: (a) the cognitive aspect of academic literacy is the one with the lowest perceived mastery by students, (b) having a positive attitude favors the development of academic literacy, (c) by knowing the aspects of their personality that can favor the development of academic literacy, students can seek strategies to improve that competency, (d) the emotional part has repercussions in the process of developing the competence of academic literacy, (e) students prefer to interact with texts in digital spaces and therefore must learn to interact critically in virtual environments, (f) Mexican students perceive themselves to have a higher level of mastery of the academic literacy competency than Spanish students.

Conclusion: The literature review and the mixed methods study allowed identifying the relevance of approaching academic literacy in university environments in a holistic manner through the analysis of the influence of cognitive, attitudinal, emotional, digital and personality aspects.

KEYWORDS

academic literacy, higher education, university students, learning strategies, academic writing, educational innovation

1. Introduction

The complex scenario currently faced by university students makes it necessary to reflect on the competencies that students should have as part of their graduate profile. One of the transversal competencies that is considered fundamental is academic literacy, since it can facilitate the learning process to be able to develop competencies such as efficient information search, comprehension of texts in the area of study to which the students belong, and writing texts with academic rigor. The mixed methods approach was chosen for the research to have a broader view of the object of the study. This study aimed to answer the question: How do cognitive, emotional, attitudinal, digital, and personality aspects influence the development of academic literacy competencies in university students?

A literature review was carried out to identify previous studies related to the topic addressed in this research. Two scoping reviews were identified, one with the objective of exploring and describing methods and strategies to promote the development of academic literacy (Klarare et al., 2022) and the other with the objective of completing a characterization of complex thinking in academic literacy and higher education (Suarez-Brito et al., 2022). Five articles were found that addressed interventions, two of them focused on writing (Olivier, 2019; Zashikhina, 2021), one aimed at academic writing in English (Roux et al., 2018), one about a Center for Reading, Writing and Orality (Soares Sito et al., 2019), and one research focusing on a program that was designed in such a way as to associate learning with the production and analysis of texts specific to the students' field of study (Urzúa-Martínez et al., 2021), and another article dealt with a cross-cutting program of academic literacy (Pardo-Espejo and Villanueva-Roa, 2019). In contrast, an investigation was also identified, the objective of which focuses on recognizing poor black youth in rural communities as knowledge givers and knowledge takers or "epistemic contributors" (Soares Sito, 2018). Technology was also an element present in some of the studies (e.g., Orlando et al., 2018). To provide information on the condition of academic literacy in Mexico and Spain, a review of studies related to academic literacy in those countries was carried out.

For the specific search of studies in Mexico and Spain, the Scopus database was chosen. A total of 12 studies were found in Mexico using the TITLE-ABS-KEY ("academic literacy" AND "Mexico") search string. Only those studies that involved a higher education level were selected. A study was identified about the subaltern practices of academic literacy, specifically regarding the use of quotation marks, making a comparison between indigenous and non-indigenous students at the higher education level; in this study, it was concluded that the meanings and practices of literacy that constitute unauthorized variation at the micro and macro levels should be a matter of reflection and educational praxis (Perales-Escudero et al., 2022). Another quantitative research identified aims to determine the relationship between teachers' prosocial personality traits and their research production and achievements. The authors conclude with a list of training areas

that can help alleviate the needs identified from the perspective of information, and scientific and academic literacy (Agredo-Machin et al., 2022). A study was identified that aims to analyze how research-academic literacies are constructed in university programs in the area of social sciences, in three public universities in northern, central, and southern Mexico (Boldo et al., 2018). The article by Mein (2012) analyzes the implications of gender-based research and pedagogy for multilingual students who cross linguistic, cultural, and disciplinary boundaries in their academic studies. Morgan and Carey's (2009) study considered that one of the main challenges to international students' right of access to higher education is academic literacy in English and proposes that the adoption of open course models in traditional universities, through blended or online learning, can offer benefits to institutions and non-English-speaking students. Using virtual ethnographic interviews and two online questionnaires, students' perceptions of the impact of blogs on their academic literacy were analyzed in the study developed by Reyes Angona et al. (2013).

Few studies of academic literacy in Spain could be found. Using the search string TITLE-ABS-KEY ("academic literacy" AND Spain) in the Scopus database, only seven articles were found. A study corresponding to high school students was eliminated. An integrative systematic review was identified to analyze the contextual and pedagogical variables associated with the development and evaluation of the Final Degree Project (TFG). This review made it possible to locate Europe as the continent where most research on TFG has been published (51.85%), with Spain in the first place (64.29%). The study concludes that the TFG involves academic literacy and is an indicator of professional development achieved (Trujillo et al., 2022). Another study related to doctoral theses in Spain focused on analyzing why doctoral students omitted the discussion of results in their theses (Hewitt and Lago, 2010). In another research, the data obtained affirmed that there are differences in students' literacy associated with the writing and reading practices they perform at the university in response to the demands of their professors (Guzmán-Simón and García-Jiménez, 2017). Other studies identified are specifically situated to English language teaching. The results of Mancho-Barés and Arnó-Macià's (2017) study shed light on practices and expectations related to the pedagogy of the discipline's disciplinary genres, an area that is at the crossroads between English-medium instruction (EMI) and English for specific purposes (ESP). In Lillis and Curry's (2006) research, the results indicate that a significant number of mediators, "literacy brokers," involved in text production influence texts in different and important ways. In contrast, Taillefer's (2005) study addresses how students abroad, a product of their own academic literacy culture, face the challenge of integrating quickly into a foreign academic literacy community.

As can be seen, the studies conducted on academic literacy in Mexico and Spain focused mainly on specific aspects, but it is important to conduct studies that analyze academic literacy in a broader sense. It was possible to identify that the studies found focus mainly on the cognitive aspect, only one article was located

in Mexico related to the personality aspect, and only two articles were found that addressed the digital aspect, one in each country. Therefore, the present study whose objective is to analyze the influence that cognitive, attitudinal, emotional, digital, and personality aspects may have on the development of academic literacy in Mexican and Spanish university students is relevant.

1.1. Academic literacy

It is important to start with the definition of academic literacy that governs this research. Academic literacy is the teaching process that may (or may not) be put in place to promote students' access to the different written cultures of disciplines. Academic literacy is currently closely linked to competencies to be able to evaluate digital content to subsequently produce knowledge (Carlino, 2013; Solimine et al., 2020).

The components that integrate academic literacy competence are as follows. (a) Attitudes toward academic literacy: the development of self-regulation skills, self-evaluation, metacognition, self-efficacy, and self-esteem as a writer and the attitudes of persistence and frustration tolerance are integrated (Errázuriz, 2017); (b) academic literacy knowledge: the production of texts incorporates knowledge of academic discursive genres and their characteristics, the management of the register to be used in each one, and reflection on the writing process, in addition to the application of cohesion mechanisms (Neira and Ferreira, 2011); (c) academic literacy practice: academic literacy practice involves the use of appropriate strategies for reading and producing (writing and speaking) texts that are considered linguistically and technically appropriate within various academic contexts (Papashane and Hlalele, 2014); and (d) problem-solving skills: they allow for building the problem scenario and similarity of situations that are connected to the learning objectives. They enhance critical thinking (Choon et al., 2021).

In addition to knowing the components that make up academic literacy, among which the cognitive aspect is included, it is important to identify the aspects that influence the development of academic literacy; therefore, it is important to address theoretically the emotional, attitudinal, digital, and personality aspects. In this research, the aforementioned aspects are theoretically supported as follows: the digital element is supported by the Digital Competencies Framework for Citizenship (DigComp 2.1), specifically in Competency Area 1: Information and Digital Literacy, which seeks the development of the following competencies: 1.1 navigate, search, and filter information and digital content, 1.2 evaluate information and digital content, and 1.3 manage information and digital content (Carretero et al., 2017). The attitudinal part is based on the theory of planned behavior (TPB). According to TPB, people's important behaviors are intentional and, although external and personal constraints make it difficult to act, the immediate determinant of the behavior is the person's intention to person to perform that behavior (Nuttavuthisit and Thøgersen, 2017). Therefore, TPB proposes the

following three variables that influence behavioral intentions: attitude, subjective norms, and perceived control. Attitudes reflect the degree of favorability or unfavorability toward a behavior based on the individual's acquired beliefs. The subjective norm refers to perceptions of social pressure to engage in a certain behavior by the individual. These arise from the importance people place on the beliefs of others relevant to them, thus regulating their behavior and motivation. Finally, perceived behavioral control considers the perceptions of the relative ease or difficulty of performing the behavior (Galleguillos-Cortés et al., 2022). TPB variables can be identified in academic literacy based on the facts that students can show favorable or unfavorable attitudes toward academic writing, subjective norms are reflected in how students perceive that their teachers, mentors, or facilitators consider that they should perform in the process of writing academic texts pertaining to their area of study, and perceived control is related to how students perceive the ease or difficulty in performing in the process of writing the texts they must produce throughout their student career. The emotional aspect is supported by Lazarus (1968), which holds that, in the face of any threatening event, first, a cognitive appraisal occurs; then, an emotional response is triggered, and, finally, a behavior is adopted to cope with the situation. In this case, when students are faced with the challenge of writing documents specific to their disciplinary area and are not familiar with the process, they may perceive the situation as a threat that leads them to experience negative emotions when faced with the challenge, or they may feel motivated by the challenge they are facing. The theory behind the personality aspect is explained in detail in the following section.

1.2. Personality in learning

Each person is different with respect to their abilities, strengths, and the way they perceive their environment, which allows their personality type to influence the way they perform in the different areas of their lives. In the educational field, it is not surprising that the personality and learning style of each student will result in success or failure in their academic life, as well as the development of learning skills (Rosas Prado et al., 2019). Personality is detected through people's behavior, constitutes a general way of behaving in the circumstances of life, and is mainly generated because of genetic inheritance and interaction with the environment and society (Muela et al., 2010). Thus, personality refers to those traits that unite and differentiate us from the rest of the people to achieve a balance. An important theory of personality is the theory of personal constructs. Kelly (1955) emphasized the creative capacity of living things to represent their environment, as opposed to simply reacting to it. These representations are known as constructs, the patterns that we create in our mind and attempt to fit over the realities of the world. Since our constructs do not always fit with reality, we are constantly modifying them as well as trying to increase our repertoire of constructs. Over time, we tested our constructs for

the ability to predict what will happen in our lives. With sufficient time and experience, and if we are willing to learn from our mistakes, we can evaluate all of our interpretations of the world in which we live. In terms of academic literacy, the constructs that people can form are derived not only from the interaction with reality and the perception that arises from it but also from the exposure to content in physical media and digital environments that provide information already digested and that reflect the perception of those who design such content, hence the importance of promoting critical thinking in students when choosing and analyzing the information they have access to through multiple platforms and spaces.

1.3. The attitude and emotional part of learning

Attitude is made up of three essential components, which are related to (a) the cognitive component, which is based on the ideas and perceptions about the object of the attitude, (b) the affective component, characterized by the feelings that the person has and their intensity (acceptance-rejection), and (c) the behavioral component, given by the response that the subject has, in reaction to the object of the attitude (Carfora et al., 2022). There are experiential referents, which creates predispositions or attitudes that affect the incorporation of the student into the learning process and affect their school achievement or failure (Gómez Barbosa et al., 2021). Attitudes are learned and relatively stable, so they could be more persistent than habits (Andrade-Valles et al., 2018). Another important element is the emotional aspect. Emotions are closely related to the educational process and the inability to regulate emotional processes can significantly impair the student's academic performance (Instituto Superior de Estudios Psicológicos, 2021). Attitudinal and emotional aspects are important influencing factors in learning and, therefore, can positively or negatively affect the development of academic literacy.

Several studies were identified related to the influence of attitude with learning and specifically with the performance of complex writing tasks as a part of academic literacy. In education, learning theory, and research, there are a number of constructs that relate closely to attitude and attitudinal learning, including learner beliefs, motivation, self-regulation, and self-efficacy (Watson et al., 2018). Self-efficacy is a type of belief that refers to an individual's judgment of his or her ability to perform the actions necessary to achieve specific goals (Bandura, 1977). Students with high self-efficacy may have higher confidence and, therefore, be better able to manage the cognitive demands of learning by managing their learning environments through strategies such as self-regulated learning strategies (Komarraju and Nadler, 2013). Self-efficacy is particularly important in completing complex writing tasks. In a study exploring self-efficacy in writing, Pajares et al. (2007) found that how students interpret the results of their own past writing performance, such as how successful they believe they were at completing a writing

task, can make a key contribution to their sense of self-efficacy. Graham et al. (2018) found that students' beliefs (i.e., sense of self-efficacy) contributed to 10% of the variance in predicting students' writing outcomes and the percentage is even higher (16.3%) for students with disabilities. As can be seen, one aspect that is important to consider for the adequate development of academic literacy is self-efficacy.

The emotional aspect and its link to academic literacy development have also been addressed in previous research. To carry out the coordinated mental activities involved in constructing meaning from text, especially when reading complex texts, a reader must be motivated and put forth effort (Cho et al., 2021). A lack of motivation can be detrimental to reading development because motivation determines how students respond to the challenges of accessing complex readings (Dweck, 1999). Some students may find it motivating to tackle complex readings, but others may feel demotivated and their academic literacy development may suffer.

1.4. Academic literacy in the digital scenario

Academic literacy is currently developing in a more complex scenario: the digital scenario. The evolution of technology has always had an enormous impact on all fields of human activity and, in this sense, the digital environment in terms of information search and interaction with texts is no exception. The complexity of the digital scenario not only presents difficulties but also has advantages (Selfa Sastre and Falguera Garcia, 2022). Digital literature presents the reader with a set of texts in which the word is not only the goal of reading but also broadens its horizons toward other analog, visual, and sound codes, thus allowing experimentation linked to the technique, software, and digital possibilities surrounding the literary message (Torres and Côrtes, 2021) and the same can occur with scientific texts in digital environments. The interest in the study of academic literacy is related to the recognition that the initial literacy acquired during primary and secondary education is not sufficient, as it does not allow students to successfully face the demands of a given field of knowledge. Upon entering the university context, students encounter specialized knowledge and, therefore, need to develop specific strategies to engage in the different activities of text analysis and production required for learning in the context of higher education (Carlino, 2013). Academic literacy in higher education becomes an essential element to promote, not only from reading but also from other much more complex competencies, all the human potential (Suarez-Brito et al., 2022). Academic literacy is positioned as a competency worth investigating in the context of higher education.

In México, research has been conducted on academic literacy and digital scenarios. With virtual ethnographic interviews and two online questionnaires, an analysis was made of students' perceptions of the blogs' influence on their academic literacy, the

links and tensions between the two types of writing, and the possible lessons for educational innovation in the area (Reyes Angona et al., 2013). Another research conducted at the Benemérita Universidad Autónoma de Puebla (BUAP) was identified, which focuses on analyzing the possibilities that a virtual environment through a commercial video game presents for academic writing and publication within the framework of university education (Ponce Carrillo and Alarcón Pérez, 2020). The study of Fernández-Cárdenas and Piña-Gómez (2014) presents the development of a resource portal for academic writing, which illustrates a set of actions and processes that constitute the academic writer's craft as a social practice. This modeling of academic writing is achieved using a sociocultural and a visual semiotic paradigm. These are some of the studies that show how digital resources and platforms can favor the development of academic literacy. It is important to integrate the development of academic literacy into the graduation profile of university students.

As can be seen earlier, the theoretical framework of this research is based on Carlino's definition (2013), which conceives academic literacy as the way in which students manage to interact with texts from their field of study and from different disciplines (cognitive aspect) and is enriched with the theories that, in previous paragraphs, explain the attitudinal (TPB), emotional (Lazarus' Cognitive Appraisal Theory), digital (Digital Competencies Framework for Citizenship), and personality aspects (Kelly's theory of personal constructs). This theoretical approach allows for a holistic view of academic literacy development.

2. Method

For the present research, the Mixed Methods approach was used since the information was collected and analyzed by means of quantitative and qualitative instruments to know the perception on the levels of mastery regarding academic literacy competencies. An explanatory sequential design was chosen since the data obtained through quantitative instruments are enriched with the data collected through qualitative instruments. In the quantitative part, the instrument called eResearch&Literacy (Castillo-Martínez and Ramírez-Montoya, 2020) was applied to 595 higher education students from different institutions in Mexico (383) and Spain (212). **The participating institutions were Tecnológico de Monterrey (various campuses), Universidad del Valle de México (various campuses), and Universidad de Sevilla and Universidad de Cantabria.** As the study was part of the NOVUS OpenResearchLab project of the Tecnológico de Monterrey, the research professors who were part of the project team were in charge of providing support for the application of the eResearch&Literacy questionnaire in their institutions. The application of the questionnaire was carried out online using a Google form.

In the qualitative part, the Case Study approach was chosen. Case studies are analyses in detail of single phenomena that are studied holistically by one or more methods. These single phenomena under study—the cases—may be persons, events,

decisions, periods, projects, policies, institutions, or other phenomena or systems (Thomas, 2020). For the present research, an invitation was made to undergraduate professors from universities in Mexico and Spain, obtaining a response from a professor at a university in northern Mexico, so that the case study was constituted with a group of 18 first-semester undergraduate students of the course "Business Strategy and Talent." The course was chosen to carry out the case study because its professor stated in an interview that his students had difficulties in understanding texts in their area of study, in academic writing, in efficiently searching for information, and in citing according to APA standards. An interview was conducted with a student and a teacher of that course, and a focus group was also carried out with five students of the same course. In the quantitative part, it was also applied to the group of the course analyzed by means of a case study the 16Personalities test (Neris Analytics, 2022), to know their personality type, and through the focus group (qualitative part), it was possible to identify which aspects the students consider that can be useful for the development of the academic literacy competency and which ones can be an obstacle. They also shared the strategies that they find suitable for the development of academic literacy competencies. The analysis of the data collected with the Likert-type eResearch&Literacy questionnaire was carried out with Microsoft Excel, and for the analysis of the qualitative data, the Atlas Ti program was used to identify the codes that later gave rise to the series of categories shown in the "Results" section.

The eResearch&Literacy questionnaire was subjected to a validity and reliability process. Criterion validity was determined, which refers to the strength of the relationship between the measures intended to predict the final criterion of interest and the criterion measure itself (Borneman, 2012). The experts showed a high level of agreement on the ratings given to the criteria of clarity, coherence, relevance, and sufficiency shown in the instrument. A high validity was established between the items of the instrument and the content to be evaluated since an average of 90% was obtained according to the scale used with respect to the degree of agreement. The degree of consistency of this instrument is high since, when applying the split-half reliability statistical method, a result of 0.94 was obtained. A detailed description of the validity and reliability process can be found in the article "Experts' validation of an instrument for self-perception of research skills to develop academic literacy" (Castillo-Martínez and Ramírez-Montoya, 2020). The questionnaire that was answered as a diagnostic tool to identify students' perception of their level of mastery of research and academic literacy competencies is shown in Table 1. A four-point Likert-type scale was used for the questionnaire, namely, 1 = strongly disagree, 2 = disagree, 3 = agree, and 4 = strongly agree.

2.1. Test to identify personality types

The teacher of the "Business Strategy and Talent" course decided to apply the 16Personalities test so that students could learn about their personality types. This resource was used for the

TABLE 1 eResearch&Literacy questionnaire.

Indicators	Items
Attitudes towards research	1. I fulfill my responsibilities in research projects with the guidance of my advisor and/or professor.
	2. I attempt to seek feedback from my advisor or research-related subject teacher regarding my progress on research deliverables at least every 2 weeks.
	3. When I fail to fulfill some academic responsibility, I usually evaluate myself to detect my own mistakes.
	4. I ask for directions if I have any doubts regarding my research work.
	5. I take advantage of opportunities to learn something new in the field of research.
Attitudes towards academic literacy	6. I can accept criticism of my writing and improve it as often as necessary.
	7. I am in the habit of reviewing my papers thoroughly when I complete them before submitting them to my advisor or teacher for revision.
Research knowledge	8. I am familiar with quantitative research methods.
	9. I am familiar with qualitative research methods.
	10. I have knowledge of how to formulate a research question.
	11. I know how to conduct database searches to carry out my research.
	12. I have the theoretical and methodological elements for the elaboration of hypotheses.
Academic literacy knowledge	13. I know the elements of the structure of a research report according to the guidelines of my career.
	14. Regarding a research paper, I know the elements of the structure that are handled in my career.
Research practices	15. I can design research instruments consistently in relation to the research method used.
	16. I can select and/or extract from databases information relevant to my research.
	17. I have the ability to organize and systematize information effectively.
	18. I have the skills for data analysis and interpretation.
	19. I know how to use computer programs for data analysis (Example: SPSS, Minitab, Atlas.ti, QSR Nvivo).
Academic literacy practice	20. I correctly apply spelling and grammatical rules in my writings to report the results of a research process.
	21. I prepare research reports following the structure established for my career.
	22. I elaborate research articles following the structure requested by the scientific journals to which I am addressing.
Research problem solving	23. I contribute to the solution of problems in my area of study through participation in research projects.
	24. I can understand what is happening and what is needed to solve the research problems that are presented to me.
	25. In addition to proposing ideas, I execute actions to solve research problems.
	26. I apply creativity and look for innovative solutions when solving research problems.
	27. I tend to critically evaluate the solutions derived from a given research problem.
Problem solving skills	28. If I do not have enough knowledge about the elaboration of a type of text, I am capable of doing research to elaborate it with the appropriate structure.
	29. If I do not understand a text, I look for strategies that make it easier for me to understand it.

present research to identify the relationship between personality type and the development of academic literacy competencies. The test, as its name suggests, yields 16 personality types, which are classified into four groups, namely, Analysts, Diplomats, Sentinels, and Explorers. The personality types are shown in Table 2.

The personality types we focused on are those that correspond to the five focus group participants. A **Campaigner** is someone with extraverted, intuitive, feeling, and prospecting personality traits. These people tend to embrace big ideas and actions that reflect their sense of hope and goodwill toward others. Their vibrant energy can flow in many directions. A **Debater** is a person with extraverted, intuitive, thinking, and prospecting personality

traits. They tend to be bold and creative, deconstructing and rebuilding ideas with great mental agility. They pursue their goals vigorously despite any resistance they might encounter. A **Protagonist** is a person with extraverted, intuitive, feeling, and judging personality traits. These warm, forthright types love helping others, and they tend to have strong ideas and values. They back their perspective with the creative energy to achieve their goals. An **Entrepreneur** is someone with extraverted, observant, thinking, and prospecting personality traits. They tend to be energetic and action-oriented, deftly navigating whatever is in front of them. They love uncovering life's opportunities, whether socializing with others or in more personal pursuits.

TABLE 2 Test 16 personalities.

Personality types			
Analysts			
Intuitive and Thinking personality types, known for their rationality, impartiality and intellectual excellence.			
Architect	Logician	Commander	Debater
Diplomats			
Intuitive and Feeling personality types, known for their empathy, diplomatic skills, and passionate idealism.			
Advocate	Mediator	Protagonist	Campaigner
Sentinels			
Observant and Judging personality types, known for their practicality and focus on order, security and stability.			
Logistician	Defender	Executive	Consul
Explorers			
Observant and Prospecting personality types, known for their spontaneity, ingenuity and flexibility.			
Virtuoso	Adventurer	Entrepreneur	Entertainer

3. Results

As already mentioned in the “Method” section, the participants in this study were 595 university students from Mexico (383) and Spain (212) for the application of the eResearch&Literacy instrument but, for the qualitative part, were carried out through a case study, the group analyzed consisted of 18 first-semester undergraduate students from a university in northern Mexico, of whom 5 participated in the discussion group, one participated in an interview, and a professor was also interviewed. In addition, all 18 students were administered the 16Personalities test. The research results emerged from the application and analysis of quantitative and qualitative instruments, as well as their triangulation.

3.1. Quantitative instruments

After applying the eResearch&Literacy instrument to 595 higher education students, we obtained the averages for each dimension of the instrument as well as the overall average regarding the perception of research and academic literacy competencies. For the present research, we focused on the part corresponding to academic literacy, as shown in Table 3.

As can be seen in Table 3, the overall average shows a perception regarding their mastery of academic literacy competencies that is at a high level. In the Learning by Knowing and Learning by Doing dimensions, the average is below 75%. In contrast, the Learning by Solving and Learning by Doing dimensions are above 80%.

An analysis was also performed to identify the averages by item according to academic literacy, as shown in Table 4.

It can be seen in Table 4 that participants perceive a higher level of mastery with respect to the correct application of spelling and grammar rules, which is reflected in their research reports. In contrast, they perceive a lower level of

TABLE 3 Analysis by dimensions of the questionnaire eResearch&Literacy.

	Mean	Standard deviation
Learning by being	3.41	0.64
Learning by knowing	2.78	0.74
Learning by doing	2.94	0.84
Learning by solving	3.24	0.60
Global	3.09	0.76

mastery with respect to preparing research articles following the structure requested by the journals to which they are addressed. The attitudinal part of the instrument obtained a perception of the high level of mastery, which is reflected in the averages obtained in items 28 and 29, related to the attitude they show when they do not have enough knowledge about the elaboration of a type of text or do not understand a text.

A comparative analysis of the perception of university students in Spain and Mexico is shown in Table 5.

Table 5 shows that there is a perception of a higher level of mastery of the academic literacy competency in Mexican students, and when comparing by dimensions, it can be seen that the perception of the level of mastery in all dimensions is higher in Mexican students than in Spanish students. The variation in the responses is similar in the students of both countries.

3.2. Qualitative instruments

An interview was conducted with a teacher and other with a student, in addition to a focus group with five students, and important data were obtained regarding their perception of the personality strengths assigned to them according to the 16Personalities test they answered as well as the points that could be an obstacle to the development of these competencies.

TABLE 4 Means by items corresponding to academic literacy.

Number	Item	Mean	Standard deviation
13	I know the elements of the structure of a research report according to my career guidelines.	2.79	0.75
14	Regarding a research paper, I know the elements of the structure that are handled in my career.	2.78	0.73
20	I correctly apply spelling and grammatical rules in my writings to report the results of a research process.	3.40	0.62
21	I prepare research reports following the structure established for my career.	2.91	0.75
22	I prepare research articles following the structure requested by the scientific journals I write for.	2.52	0.87
28	If I do not have enough knowledge about the elaboration of a text type I am able to research to elaborate it with the appropriate structure.	3.21	0.61
29	If I do not understand a text, I look for strategies that will make it easier for me to understand it.	3.27	0.59

TABLE 5 Comparative analysis between Mexico and Spain.

		Learning by being	Learning by knowing	Learning by doing	Learning by solving	Overall mean
Mexico	Mean	3.37	2.97	2.96	3.14	3.11
	Standard deviation	0.04	0.05	0.11	0.07	
Spain	Mean	3.19	2.59	2.63	2.93	2.84
	Standard deviation	0.07	0.06	0.11	0.06	

3.2.1. Personality type can promote or hinder the development of academic literacy competencies

In the focus group, although the students were chosen randomly, it turned out that the participants possessed different personality types. They shared the strengths of their personality to facilitate the development of academic literacy competencies and the points that could be an obstacle to achieving their development, as shown in Table 6.

3.2.2. Research competencies can support the development of academic literacy competencies

Although the eResearch&Literacy instrument was analyzed only with respect to the academic literacy competencies portion of the instrument, when conducting the interviews with the teacher and student and the focus group, it was reinforced that there is a close relationship between both types of competencies and that it was important to address it also in this study, which aims to provide a holistic view. The teacher's (PR) view on this point is reflected as follows:

I believe that everything is governed by the curiosity to investigate. I believe that we can start from there so that a person is able to review, investigate, inquire, formulate, etc. I believe that it is a whole, we could not say that this competence is better for literacy, because it is a whole and somehow in research you will have to go through all the phases, even if it is multidisciplinary.

The student interviewed (S6) responded as follows regarding research skills that can foster the development of academic literacy skills:

Look for reliable information and if you have the support of a teacher, ask him/her as well. Generate that criterion of saying this information is useful, this information is not useful, what am I looking for. If I am going to give proposals to McDonalds, I am not going to give them information that they already have, I am going to generate new information, so how to have my own criteria to say what I am looking for to see what I write.

3.2.3. Personality aspects favor the development of research and academic literacy skills

The participants in the focus group were asked about what strengths they considered their personality type to favor the development of their research and academic literacy competencies. In this regard, S4 comments:

I think one strength of the animator personality is that he is very communicative, so it is easy for him to communicate and get sources of information, also another one could be that he is a person with a lot of imagination and so on. So, I think that could help a lot.

Also, S3 shares:

The strong point in the personality type that came out for me is that we really like to work in a team. I think that if you do

TABLE 6 Helpful and hindering points in the development of academic literacy skills.

	Student 1	Student 2	Student 3	Student 4	Student 5
Personality type	Campaigner	Debater	Protagonist	Protagonist	Entrepreneur
Useful points	<i>I am very good at generating new skills.</i>	<i>I got to be an innovator. I think it's important to be one step ahead of others, because you are researching, you want to know more and more.</i>	<i>The strong point in my personality type is that we like to work as a team. I think that if you do research in a team it is a good way to learn more and have more knowledge.</i>	<i>I think a strong point of the animator personality is that he is very communicative, so it is easy for him to communicate and get sources of information, also another one could be that he is a person with a lot of imagination and so on. So, I think that could help a lot.</i>	<i>I think I would have to work on that because I tend to get distracted very easily.</i>
Obstacles	<i>The obstacles are that if I do not find what I want or expected I get disappointed very easily.</i>	<i>One obstacle of my personality I think is my patience, because if I do not see results soon, I need more, more, more, more.</i>	<i>Also my personality is kind of easily distracted.</i>	<i>As they are so changeable, you do not focus on a single point and you want to cover a lot, so it gets complicated in the end.</i>	<i>I think it would be a disadvantage, because I was saying that we cannot concentrate easily, follow a routine, when I'm doing research I get distracted by something else, by the ads that are on the page or so.</i>
Strategies that facilitate academic literacy development	<i>When you make a mistake they put in there that you made a grammatical mistake in this and that could help me.</i>	<i>Reading helps me more, I understand more when I read. It is always easier for me to understand the text</i>	<i>Reading, well, the more I read, the better I get at writing and I think that, yes, so does reading.</i>	<i>I feel that the feedback they give you to improve your spelling is more useful to me. If they tell you that you made a mistake, I'll check my work more.</i>	<i>I am more auditory, I have to listen to learn. When I have to learn readings regularly I record myself and then I start listening. If I just read it I do not learn.</i>

research in a team it is a good way to learn more and have more knowledge.

For his part, S2 explains, “I was an innovative person. I think it is important to be one step ahead of others, because you are researching, you want to know more and more” and S1 expresses, “The useful ones would be that I am, well, I said that I am very good at generating new skills.”

3.2.4. Personality aspects hinder the development of research and academic literacy skills

In the focus group, students were led to reflect on what they felt were the aspects of their personality type that hindered the development of research and academic literacy skills. S1 expressed that, for her, the obstacle was, “If I do not find what I want or expected I get disappointed very easily.” S2 shares that, “One obstacle of my personality I think is my patience, because if I do not see results soon is to see what is happening, I need more, more, more, more.” S3 and S5 agree that the main obstacle of their personality type is that they are easily

distracted. In contrast, S4 comments, “Being so changeable, you do not focus on a single point and you want to cover a lot, so it gets complicated in the end.”

3.2.5. Difficulties of university students with respect to academic writing

For students, academic writing is not a simple process. Sitting down to write a report, a research report, or an essay, among others, is a complex process. S6 states “I consider myself very good in spelling and grammar and in research in general, but what I struggle with the most is to land my ideas, that is, to draw conclusions, that is what I struggle with the most.”

Also, teachers detect the difficulty students have regarding academic writing. PR shares:

There are serious spelling problems that we have to help them clean up, secondly, they kind of go off into their own thoughts and write themselves, trying to use writing as a kind of outlet, which is fine, I mean, it's part of what we are working on, but sometimes it's so much the daily type that they put in there, that suddenly they forget about the academic part.

3.2.6. Influence of the emotional side on academic literacy

A question that has been less explored is how the emotional side can influence proper academic writing and thus efficient academic literacy. PR reflects on how this plays out in students and concludes:

I do have my reflection, I am clear about what I am thinking, but then how do I ground that to the concepts we are seeing. That part of aligning mine with the theoretical part because they go more to the emotional or to what I bring and feel, although maybe in some way they do thread it, when I review it in writing it is not seen and that is what they need to integrate.

The emotional side is also reflected in research work that can serve as a basis for academic literacy, as exemplified in S1's comment: I started reading everything, but I did not know what I was going to use, almost nothing helped, that's as far as I got. When the student was asked if he had become unmotivated, he responded: Yes, they were moving very fast and I had to study a lot.

3.2.7. Influence of the attitudinal part on academic literacy

Another aspect that can influence how academic literacy is developed is attitude. An important issue that can positively influence academic writing is the attitude to research, is the attitude to be curious, willing, and also open to constructive criticism regarding research reports to improve academic writing skills and is the attitude research has on a better development of academic literacy competencies. PR expresses that attitude even goes beyond the realm of academic literacy, which impacts everything, as she explains below:

I believe that for everything, organizations can check this part, the attitude, the quality of service, assertiveness, all these personal skills that a person already has as personal skills help you to develop the rest. From having to apply a test to a certain group of the population, to be able to reach a certain group of the population you need to know how to negotiate and communicate effectively and be very assertive, be friendly, if you are not like that, they will hardly want to answer you something, even if it is to know the quality of a product or for marketing purposes, etc.

Some students showed a positive attitude toward research competencies that can favor the development of academic literacy, as in the case of S2: "I like to read other people's research and then be able to implement them, at the moment if I am liking research." With respect to academic writing, a positive attitude was identified, as shown in S4's response: "That is why it is good to get a second opinion, or read it after 2 days to see if it makes sense what you are expressing, maybe there are things you understand but others do not."

3.2.8. Students prefer to interact with texts on digital platforms

Currently, students can access a large number of texts at the click of a button; they are no longer limited to going to a library to consult a limited collection compared to the vast amount of information provided by the web, the main site where they search for their schoolwork or research. PR reinforces this statement when it shares:

I think that more than anything the digital part, the printed part is not much to their liking, with the digital part they respond better, but if you put a video where you integrate the message you want to provide them with, it is better for them. They prefer listening, watching, not so much the reading part. I mean, there are those who, yes, have that ability more developed, it is a taste they have for reading, but in general, let's say, 80% are more visual than anything else.

Regarding the search for information on Internet sites, they prefer the web since it is more complex for them to access databases of the digital library of their institution. S6 comments, "Well, the most common is to search in Google, in the open one, Scholar I always forget to use it." PR also agrees, which is reflected in the following comment:

Google (laughs), the whole web, Wikipedia is almost not used, because they also have already demonized it. I feel that for a while there was a very negative message regarding Wikipedia, so they have already identified it as a space where they can find information, but teachers are not going to accept it. More than anything else on the web, there are some very smart people who use Google Scholar, but it is rare, it is very rare who really uses digital library sources.

4. Discussion

Adequate academic literacy in higher education requires solid knowledge of academic discursive genres and the structure of scientific texts. Table 3 shows that university students perceive that they have a lower level of mastery precisely in the cognitive aspect, included in the Learning by Knowing dimension of the eResearch&Literacy instrument with a 2.78. This is specifically observed in the items corresponding to this dimension: "I know the elements of the structure of a research report according to my career guidelines" (item 13 with 2.79) and "Regarding a research paper, I know the elements of the structure that are handled in my career" (item 14 with 2.78). As can be seen in the "Research competencies can support the development of academic literacy competencies" section, S5 reflects on the cognitive aspect that it is very useful to be able to turn to a teacher to improve the knowledge one already has to carry out an efficient search for

information. According to [Neira and Ferreira \(2011\)](#), the production of texts should incorporate, in addition to knowledge of academic discursive genres, the management of the register in each one, reflection on the writing process, and knowing how to use cohesion mechanisms. It can be observed that the cognitive aspect of academic literacy is the one with the lowest perceived mastery by students. The cognitive aspect, in addition to being based on knowledge of discursive genres, text structure, and writing process, is also related to an efficient search for information, for which the teacher's or facilitator's guidance can be very valuable.

The attitudinal part has an impact on learning in general and, therefore, on the development of academic literacy. [Table 4](#) shows that the items corresponding to attitudes are located in the highest averages with respect to the students' perception of their level of mastery, and these items are related to having an attitude of openness to search for information if one does not have sufficient knowledge about how to write certain types of texts (item 28) and to the willingness to seek strategies that can facilitate the understanding of texts that are difficult to understand (item 29). In the "Influence of the attitudinal part on academic literacy" section, it is identified that attitude even has repercussions in any field, not only school or academic. Attitude can affect the learning process ([Gómez Barbosa et al., 2021](#)). Perceived behavioral control considers the perceptions of the relative ease or difficulty of performing the behavior ([Galleguillos-Cortés et al., 2022](#)). Students having a positive attitude about seeking strategies to help them improve the aspects of academic writing with which they are having difficulty is already a critical step, as it will enable them to commit to making progress in developing their academic literacy competence.

Knowing their personality type can have the benefit of helping them identify those areas of opportunity that they can transform to achieve better learning and, in this case, better academic literacy development. [Table 6](#) shows that students with diverse personality types participated in the focus group, and all were able to identify useful points and points that could hinder the development of their academic literacy competencies. Regarding the favorable aspects that the students identified, the creative part stands out, as shown in the responses of S4: to have imagination, S5: to be innovative, and S1: to generate new skills, as well as S4: communication skills and S3: teamwork. It was possible to identify that an important obstacle is impatience, which is shown in the responses of students S1 and S2 regarding feeling disappointed when not finding what they are looking for and wanting to see results quickly, respectively. Another unfavorable point is that they are easily distracted, as in the case of S3 and S5. In the educational setting, it is not surprising that each student's personality and learning style will result in success or failure in their academic life, as well as the development of learning skills ([Rosas Prado et al., 2019](#)). The constructs that students have formed throughout their lives also influence learning ([Kelly, 1955](#)) and, in this case, the development of academic literacy. By knowing the aspects of their personality that can favor or disfavor the development of their

competencies, in this case, academic literacy, students can seek strategies that lead them to modify the constructs that are preventing them from having a positive disposition for learning and thus, they will achieve potentiating their strengths and improve their areas of opportunity in favor of the development of academic literacy.

The emotional part can also influence adequate academic writing. In the "Influence of the emotional side on academic literacy" section, it is reflected that sometimes students get carried away by the emotional part of their writings and do not take care to integrate what they are writing with the theoretical part of the course for which they elaborate their writings. Emotions are closely related to the educational process and the inability to regulate emotional processes that can significantly impair the student's academic performance ([Instituto Superior de Estudios Psicológicos, 2021](#)). The emotional aspect supported by [Lazarus \(1968\)](#) holds that, before adopting a coping behavior, the threatening event is first cognitively appraised and then an emotional response is triggered. In the case of the group of participants in the study, it is not identified that they perceive the act of writing as a challenge that leads them to experience negative emotions; on the contrary, it seems that the writing process can be for them a catharsis, which leads them to an emotional outlet, which is very positive, but can be an important distractor to achieve the objectives of academic writing and the subject for which they are writing the text. The emotional part has repercussions in any situation and context and the educational environment is no exception, the teaching-learning process may be favored if the student has adequate emotional regulation and whether the facilitator can help students find a balance between expressing their emotions and meeting the academic writing goal.

Learning is now largely framed in digital environments, which also impacts how university students interact with texts. The content of the "Students prefer to interact with texts on digital platforms" section shows how students prefer to read in digital environments; unfortunately, they do not have an adequate level of mastery regarding their digital competencies. Upon entering the university context, students encounter specialized knowledge, and therefore, they need to develop specific strategies to engage in the different text analysis and production activities required for learning in the higher education context ([Carlino, 2013](#)). As if that were not enough, they must also be prepared to be able to perform an efficient search for information on digital platforms ([Carretero et al., 2017](#)) that put a large amount of information within their reach. In this way, academic literacy implies that the student has the knowledge and skills to interact with texts not only physically but also in digital environments. The students need to improve their ability to understand specialized texts at the university level, especially those in his/her field of study.

An important aspect was to establish an analysis between Spanish and Mexican students to know the difference in their perception of their level of mastery of academic literacy. [Table 5](#) shows that the perceived level of mastery is higher for Mexican students than for Spanish students, taking into consideration the

overall mean and mean for each dimension. More studies on academic literacy have been carried out in Mexico than in Spain, according to the literature review carried out in the Scopus database and whose explanation is detailed in the “Introduction” section. The search allowed us to identify that the studies in both countries address specific topics that are fundamentally related to the cognitive aspect, which shows the relevance of studying academic literacy considering a vision that also integrates personality, emotional, attitudinal, and digital aspects.

The literature review allowed us to have a notion of the path taken by researchers regarding the development of academic literacy in university students. In the “Introduction” section, two reviews were found, which were a good contribution to learn about research that shared strategies for the development of academic literacy. It was also valuable to identify several articles that showed evidence of interventions, but it could be seen that the focus was on text analysis and production. The literature review confirmed that the present study generates a contribution to the field of the study of academic literacy in the university context since it contemplates aspects that have not received much attention but which, in the theoretical framework and in the “Results” section of this article, show their importance on aspects, including attitudinal, emotional, and personality, that have been analyzed in various studies involving cognitive and digital, giving rise to a holistic view of the development process of academic literacy competence.

5. Conclusion

The present study found that, in relation to the cognitive aspect, it is necessary to reinforce the knowledge of the different types of academic discursive genres and the structure of texts from the time students enter the undergraduate program and to facilitate the comprehension of texts in their area of study and even enable them to understand texts from other disciplinary areas, to favor their participation in interdisciplinary projects. It was also identified that the attitude that students show toward learning may depend on the learning environments that teachers provide so that innovative strategies can be motivating to promote academic literacy. The emotional part should also be considered; the school environment can provide resources that promote emotional regulation in students so that they can have an adequate school performance, which can influence the development of competencies, such as academic literacy in this case.

With respect to personality, it was found that it can be of great benefit to generate spaces and facilitate resources so that students can identify their personality type and their way of learning since they will become aware of their strengths and areas for improvement to design strategies that facilitate their teaching-learning process, allow them to improve their skills and competencies, and favor the development of an efficient academic literacy. Another finding was that students largely interact with

texts through digital environments, so universities must seek to promote critical and creative thinking so that they can distance themselves from the contents to which they are exposed and can choose those that are truly in line with their research topics, and it is important to adapt the training to the use of the digital library, taking special care that the information is clear and simple. To ensure that students are trained, it may be necessary to establish formal courses as the teacher interviewed suggests, since having students enter training voluntarily has not yielded results.

It was possible to identify that Mexican students have a higher perception regarding their level of mastery of academic literacy competence than Spanish students. Also, through a literature review, it was found that more studies have been conducted on academic literacy in Mexico than in Spain and that, in both countries, the research focuses on specific aspects, largely in relation to the cognitive and digital aspects. Therefore, a study that addresses academic literacy in university students with a holistic approach is relevant.

The design of strategies that contemplate the elements of the cognitive, attitudinal, emotional, digital, and personality aspects will make it possible to provide students with the possibility of favoring their learning process, generating greater motivation, and making it more meaningful. This will have an impact on the development of competencies necessary for today's complex environment, among them, the academic literacy competency, which is so important in the graduate profile of university students.

For future research, we recommend the design and validation of an instrument to evaluate the cognitive, attitudinal, emotional, digital, and personality aspects to have more information to provide greater support for the design of strategies that favor the development of academic literacy in university students. Strategies could even be designed to be the curricular integrated into the last grade of high school to generate a valuable bridge that allows students to enter with greater knowledge and skills for academic reading and writing, facilitating their progress in their level of mastery in their academic literacy competence.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the (patients/participants or patients/participants legal guardian/next of kin) was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

IC-M was primarily responsible for designing the study, collecting and analyzing data, and drafting the manuscript. CC participated in the methodology and the writing of the manuscript. LG-M and MR-M contributed to the revision and improvement of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This publication is a product of the projects “OpenResearchLab: innovation with artificial intelligence and robotics to scale domain levels of reasoning for complexity” (ID Novus N21-207), funded by the Institute for the Future of Education (IFE) and “Challenge-Based Research Funding Program 2022”. Project ID # I005 - IFE001 - C2-T3 – T. Also,

academic support from Writing Lab, IFE Tecnológico de Monterrey, México.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Agredo-Machin, D. M., Romo-González, J. R., Machin-Mastromatteo, J. D., and González-Quinones, F. (2022). “Personality traits as drivers of the scientific production: Information, scientific and academic literacies implications,” in *Information Literacy in a Post-Truth Era*, ed. S. Kurbanoglu 290–301.
- Andrade-Valles, I., Facio-Arciniega, S., Quiroz-Guerra, A., Alemán-de la Torre, L., Flores-Ramírez, M., and Rosales-González, M. (2018). Actitud, hábitos de estudio y rendimiento académico: Abordaje desde la teoría de la acción razonada. *Enferm. Univ.* 15, 342–351. doi: 10.22201/eneo.23958421e.2018.4.533
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavior change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191
- Boldo, C. S., Villavicencio, M. G. T., and Lozano, E. G. S. (2018). Pendiente. [¿Y esto para qué me va a servir? Jóvenes universitarios y literacidades investigativas*]. *Revista De La Educacion Superior* 47, 139–156. doi: 10.36857/resu.2018.188.511
- Borneman, M. (2012). “Criterion validity” in *Encyclopedia of Research Design*, ed. N. Salkind (Thousand Oaks, CA: SAGE Publications), 292–296.
- Carfora, V., Biella, M., and Catellani, P. (2022). Affective components in promoting physical activity: a randomized controlled trial of message framing. *Front. Psychol.* 13, 1–16. doi: 10.3389/fpsyg.2022.968109
- Carlino, D. P. (2013). *Leer y escribir para aprender en las diversas carreras y asignaturas de los IFD que forman a profesores de enseñanza media: concepciones y prácticas declaradas de los formadores de docentes*. Buenos Aires: Ministerio de Educación de la Nación.
- Carretero, S., Vuorikari, R., and Punie, Y. (2017). DigComp 2.1: the digital competence framework for Citizens with eight proficiency levels and examples of use [Marco de Competencias Digitales Para la Ciudadanía]. Available at: <https://epale.ec.europa.eu/es/content/marco-europeo-de-competencias-digitales-digcomp>
- Castillo-Martínez, I. M., and Ramírez-Montoya, M. S. (2020). “Experts’ validation of an instrument for self-perception of research skills to develop academic literacy,” in *Proceedings of the 8th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2020)*. Spain: University of Salamanca.
- Cho, E., Kim, E. H., Ju, U., and Lee, G. A. (2021). Motivational predictors of reading comprehension in middle school: role of self-efficacy and growth mindsets. *Read. Writ.* 34, 2337–2355. doi: 10.1007/s11145-021-10146-5
- Choon, L. K., Mohamad, S. S., and Lee, C. S. (2021). Going remote during COVID-19 pandemic: effects of problem-based learning towards improving students’ critical thinking and problem-solving skills. *Turk. J. Comput. Math. Educ.* 12, 2342–2356. doi: 10.17762/turcomat.v12i3.1216
- Dweck, C. S. (1999). *Self-Theories: Their role in motivation, personality, and development*. New York: Psychology Press.
- Errázuriz, M. C. (2017). Teorías implícitas sobre la escritura académica en estudiantes de programas de formación inicial docente: ¿inciden en el desempeño escrito? *Signo y Pensamiento* 36, 34–50. doi: 10.11144/javeriana.syp36-71.tiea
- Fernández-Cárdenas, J. M., and Piña-Gómez, L. (2014). El oficio del escritor académico. *Rev. Mex. Investig. Educ.* 51, 187–212. doi: 10.53556/rec.v51i0.25
- Galleguillos-Cortés, C., Silva-Munar, J. L., and Hurtado-Cailly, R. (2022). Influence of creative thinking beliefs on sustainable development behavior in higher education students. An approach from the theory of planned behavior. *Estudios Pedagogicos* 48, 71–87. doi: 10.4067/S0718-07052022000100071
- Gómez Barbosa, D., Prada Nuñez, R., and Hernández Suárez, C. (2021). Influencia de las actitudes en los ambientes de aprendizaje de las prácticas pedagógicas del docente de matemáticas. *Revista Boletín Redipe* 10, 238–255. doi: 10.36260/rbr.v10i8.1402
- Graham, S., Daley, S., Aitken, A., Harris, K., and Robinson, K. (2018). Do writing motivational beliefs predict middle school students’ writing performance? *J. Res. Read.* 41, 642–656. doi: 10.1111/1467-9817.12245
- Guzmán-Simón, F., and García-Jiménez, E. (2017). Academic literacy of future teachers: A comparative study on several Spanish universities. *Revista de Investigacion Educativa* 35, 317–335. doi: 10.6018/rie.35.2.246011
- Hewitt, E., and Lago, A. F. (2010). Academic style and format of doctoral theses: the case of the disappearing discussion chapter. [Estilo académico y formato de tesis doctorales: El caso de la desaparición del capítulo de discusión (sección de debate académico)]. *Iberica* 19, 119–140.
- Instituto Superior de Estudios Psicológicos (2021). *La influencia de las emociones en el proceso de aprendizaje de los niños*. Madrid: ISEP.
- Kelly, G. (1955). *The Psychology of Personal Constructs*. New York: Norton-Company.
- Klarare, A., Rydeman, I. B., Kneck, Å., Sparén, E. B., Winnberg, E., and Bisholt, B. (2022). Methods and strategies to promote academic literacies in health professions: a scoping review. *BMC Med. Educ.* 22:418. doi: 10.1186/s12909-022-03288-9
- Komaraju, M., and Nadler, D. (2013). Self-efficacy and academic achievement: why do implicit beliefs, goals, and effort regulation matter? *Learn. Individ. Differ.* 25, 67–72. doi: 10.1016/j.lindif.2013.01.005
- Lazarus, R. (1968). “Emotions and adaptation: conceptual and empirical relations” in *Nebraska Symposium on Motivation*, ed. I. W. J. Arnold (Nebraska: University of Nebraska Press)
- Lillis, T., and Curry, M. J. (2006). Professional academic writing by multilingual scholars: interactions with literacy brokers in the production of english-medium texts. *Writ. Commun.* 23, 3–35. doi: 10.1177/0741088305283754
- Mancho-Barés, G., and Arnó-Macià, E. (2017). Emi lecturer training programmes and academic literacies: a critical insight from ESP. *ESP Today* 5, 266–290. doi: 10.18485/esptoday.2017.5.2.7
- Mein, E. (2012). Bilingual literacy in context: the use of L1/L2 genre knowledge in graduate studies. *Int. J. Biling. Educ. Biling.* 15, 653–667. doi: 10.1080/13670050.2012.699946

- Morgan, T., and Carey, S. (2009). From open content to open course models: increasing access and enabling global participation in higher education. *Int. Rev. Res. Open Distan. Learn.* 10, 1–16. doi: 10.19173/irrodl.v10i5.632
- Muela, J., García-León, A., Augusto, J., and López-Zafra, E. (2010). The teaching of psychology through learning activities from a multidisciplinary approach: increasing motivation and performance. *Eur. J. Educ.* 3, 33–44.
- Neira, A., and Ferreira, A. (2011). Escritura académica: Un modelo metodológico efectivo basado en tareas y enfoque cooperativo. *Literatura y Lingüística* 24, 143–159. doi: 10.4067/s0716-58112011000200008
- Neris Analytics (2022). 16Personalities. Available at: <https://www.16personalities.com/personality-types>
- Nuttavuthisit, K. Y., and Thøgersen, J. (2017). The importance of consumer trust for the emergence of a market for green products: the case of organic food. *J. Bus. Ethics* 140, 323–337. doi: 10.1007/s10551-015-2690-5
- Olivier, L. (2019). The effect of a compulsory academic literacy module on the academic writing skills of Afrikaans mother-tongue first-year nursing students. *Ikala, Revista de Lenguaje y Cultura*. 1, 67–86.
- Orlando, J., Hanham, J., and Ullman, J. (2018). Exploring intentional use of a technological proxy, Turnitin, to enhance student academic literacy practices. *Australas. J. Educ. Technol.* 34, 44–56. doi: 10.14742/ajet.3575
- Pajares, F., Johnson, M. J., and Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Res. Teach. Engl.* 42, 104–120.
- Papashane, M., and Hlalele, D. (2014). Academic literacy: a critical cognitive catalyst towards the creation of sustainable learning ecologies in higher education. *Mediterr. J. Soc. Sci.* 5, 661–671. doi: 10.5901/mjss.2014.v5n10p661
- Pardo-Espejo, N. E., and Villanueva-Roa, J. D. D. (2019). Diseño, Implementación y Evaluación del Programa Transversal de Alfabetización Académica Lector-es. *Información Tecnológica* 30, 301–314. doi: 10.4067/s0718-07642019000600301
- Perales-Escudero, M. D., Díaz, M. R., and Sima Lozano, E. G. (2022). Prácticas subalternas sobre el sistema ortográfico entre estudiantes universitarios indígenas y no indígenas. Un estudio de literacidades académicas. *Perfiles Educativos* 44, 39–57. doi: 10.22201/iisue.24486167e.2022.177.60078
- Ponce Carrillo, R., and Alarcón Pérez, L. M. (2020). Entornos virtuales para la escritura académica. Un modelo en Minecraft. *Alteridad* 15, 78–87. doi: 10.17163/alt.v15n1.2020.06
- Reyes Angona, S., Fernández-Cárdenas, J. M., and Martínez Martínez, R. (2013). Comunidades de blogs para la escritura académica en la enseñanza superior: Un caso de innovación educativa en México. *Revista Mexicana de Investigación Educativa* 18, 507–535.
- Rosas Prado, C. E., Zuloeta Salazar, J. F., Urbina Rosas, C. M., and Zuñe Chero, L. (2019). *Relationship Between Personality Factors and Learning Styles in Peruvian University Students*. Trujillo: Universidad César Vallejo
- Roux, R., González Quintero, E. F., and Mendoza Valladares, J. L. (2018). A pedagogical intervention experience to improve the disciplinary and ESL academic writing abilities of applied linguistics students. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana* 55, 1–13. doi: 10.7764/PEL.55.2.2018.6
- Selfa Sastre, M., and Falguera Garcia, E. (2022). From text on paper to digital poetry: creativity and digital literary Reading practices in initial teacher education. *Front. Psychol.* 13:882898. doi: 10.3389/fpsyg.2022.882898
- Soares Sito, L. R. (2018). Ensaio das estratégias das artes letradas nas zonas de contato: trajetórias de letramento acadêmico, ações afirmativas e políticas de conhecimento. *Revista Brasileira de Linguística Aplicada* 18, 821–852. doi: 10.1590/1984-6398201812992
- Soares Sito, L. R., Méndez Rendón, J. C., and Vásquez Ramírez, L. Y. (2019). Leos in practice -the educational experience in a center for readings, scribalities, and oralities. *Ikala* 24, 419–438. doi: 10.17533/udea.ikala.v24n02a12
- Solimine, G., Angel, M., and Garcia-Quismondo, M. (2020). Proposal of visual literacy indicators for competencies courses. An academic literacy perspective for academic excellence 11, 17–15. doi: 10.4403/jlis.it-12577,
- Suarez-Brito, P., Baena-Rojas, J. J., López-Caudana, E. O., and Glasserman-Morales, L. D. (2022). Academic literacy as a component of complex thinking in higher education: a scoping review. *Eur. J. Contemp. Educ.* 11, 931–945. doi: 10.13187/ejced.2022.3.931
- Taillefer, G. (2005). Reading for academic purposes: the literacy practices of British, French and Spanish law and economics students as background for study abroad. *J. Res. Read.* 28, 435–451. doi: 10.1111/j.1467-9817.2005.00283.x
- Thomas, G. (2020). *Case Study. SAGE Research Methods Foundations*. London: SAGE Publications Ltd.
- Torres, R., and Côrtes, D. (2021). “Wreadings: digital poetry and collaborative practice” in *The Community and the Algorithm: A Digital Interactive Poetics*. ed. A. Klobucar (Wilmington: Vernon Press), 3–26.
- Trujillo, C. C., Resino, D. A., Muñoz, I. A., and Gordo, P. M. (2022). Systematic integrative review of the final project: contextual and pedagogical aspects. [Revisión sistemática integrativa del Trabajo Fin de Grado: aspectos contextuales y pedagógicos]. *Revista Complutense De Educacion* 33, 543–553. doi: 10.5209/rced.74631
- Urzúa-Martínez, S., Riquelme-Yañez, R., and Micin-Carvalho, S. (2021). Impacto de un programa de lectoescritura en el rendimiento académico de estudiantes de primer año universitario en Chile. *Ikala, Revista de Lenguaje y Cultura*. 26, 283–302. doi: 10.17533/udea.ikala.v26n2a05,
- Watson, S. L., Watson, W. R., and Tay, L. (2018). The development and validation of the attitudinal learning inventory (ALI): a measure of attitudinal learning and instruction. *Educ. Technol. Res. Dev.* 66, 1601–1617. doi: 10.1007/s11423-018-9625-7
- Zashikhina, I. M. (2021). Academic writing: a discipline or disciplines? *Vyshee Obrazovanie v Rossii* 30, 134–143. doi: 10.31992/0869-3617-2021-30-2-134-143



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

John Mark R. Asio,
Gordon College, Philippines
Hui Helen Li,
Wuhan University of Technology, China
Javier Corbalán,
Universidad de Murcia, Spain

*CORRESPONDENCE

Sarfraz Aslam
✉ sarfraz.aslam@unitar.my
Nisar Abid
✉ nisar.abid@umt.edu.pk

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 16 August 2022

ACCEPTED 09 January 2023

PUBLISHED 27 January 2023

CITATION

Abid N, Aslam S, Alghamdi AA and Kumar T
(2023) Relationships among students' reading
habits, study skills, and academic achievement
in English at the secondary level.
Front. Psychol. 14:1020269.
doi: 10.3389/fpsyg.2023.1020269

COPYRIGHT

© 2023 Abid, Aslam, Alghamdi and Kumar. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Relationships among students' reading habits, study skills, and academic achievement in English at the secondary level

Nisar Abid^{1*}, Sarfraz Aslam^{2*}, Abdulelah A. Alghamdi³ and
Tribhuwan Kumar⁴

¹Department of Education, School of Social Sciences and Humanities, University of Management and Technology, Lahore, Pakistan, ²Faculty of Education and Humanities, UNITAR International University, Petaling Jaya, Malaysia, ³Faculty of Education, Umm Al-Qura University, Mecca, Saudi Arabia, ⁴Department of English Language and Literature, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

Introduction: Reading is an attempt to comprehend the writer's message for personal growth and success in the relevant fields. Thus, psychologists consider it a multifaceted cognitive process of constructing meanings from texts. The present study was conducted to determine the relationships among students' reading habits, study skills, and academic achievement in English at the secondary level in Punjab, Pakistan.

Methods: The ($n = 1614$) students enrolled in the science section for the academic year 2019–2020 participated in this descriptive correlational survey, selected from 40 high schools in Lahore, Punjab, Pakistan, through a non-proportionate stratified random sampling technique. The Reading Habits Questionnaire (RHQ) and the Study Skills Scale (SSS) were used to collect data about students' reading habits and study skills. At the same time, academic achievement was the students' grades obtained in the ninth class in the subject of English that were determined by the Board of Intermediate and Secondary Education (BISE) Lahore in 2019. Students' responses were analyzed through descriptive and inferential statistics.

Results: The results indicated that students have competent reading habits and study skills. The correlational findings showed a strong positive relationship among reading habits, study skills, and academic achievement in English, while moderate positive relationships between reading habits and academic achievement in English. However, regression analysis results were significant, while reading habits and study skills moderately predicted academic achievement.

Discussion: It is implicated that teachers should plan such assignments and tasks based on reflective thinking by considering the role of study skills in academic achievement. Moreover, teachers and school administrators could mutually create timetables for library lessons to build reading habits and study skills among learners.

KEYWORDS

reading, reading habits, study skills, academic achievement, secondary school students

1. Introduction

Knowledge gained through reading is vital for the cognitive, behavioral, and attitudinal development of learners (He, 2014; Baffoe and Okae-Anti, 2020; Hassan et al., 2021) because it is a person's ability to enhance information and comprehend the words effectively (Sabbah, 2016;

Al-Jarf, 2019). An individual reads for numerous reasons, i.e., knowledge development, recreation, joy, relaxation, and so on (Whitten et al., 2016). However, Erguvan (2016) and Mirza et al. (2021) directed that reading is an active part of life that is not just about pleasure when needed. However, Chotitham and Wongwanich (2014) conjectured that reading helps to develop critical and judgmental thinking abilities used to solve problems by conceptualizing context. Hence, Erdem (2015) and Pretorius and Klapwijk (2016) quantified that reading is essential to success because it starts from the commencement of school and continues throughout the lifetime.

Fischer et al. (2015), Oyewole (2017), Al-Jarf (2019) recognized that the importance of reading in learning could not be ignored because it is an emancipatory tool that releases students' academic frustration, ignorance, and destitution. Palani (2012) distinguished that reading is an instrument used to exchange information, while reading habit is an academic activity that enables students to benefit from reading materials. Therefore, Walia and Sinha (2014) specified that reading habits require complex skills, such as perceiving a message, skimming and scanning information, and understanding the context. Thus, compelling reading depends on readers' behaviors, known as study skills that enable them to conceptualize the new knowledge effectively (DiPerna and Elliott, 2000; Habibu and Ejembi, 2011; Gormley et al., 2018; Naqvi et al., 2018; Iheakanwa et al., 2021). While the effective study makes one narrate in their way using the stipulated meanings of the words and terms, the researchers take up for explanation and clarity (Biyik et al., 2017).

According to the available literature, students' reading habits and study skills have been of great importance for decades; while several deficiencies were found in previous studies, thus researchers considered few of them that are related to the study context. First, the researchers mainly focused on the influence of reading habits and study skills on academic achievement separately in Western countries (Bhan and Gupta, 2010; Sabbah, 2016). A few addressed Eastern countries restricted to the university level (Demir et al., 2012; Davarci, 2013; Dilshad et al., 2013; Erguvan, 2016; Alzahrani et al., 2018; Porkaew and Fongpaiboon, 2018; Thamarasseri, 2018; Ameyaw and Anto, 2019; Ehsan and Sultana, 2020; Tonka and Bakir, 2020; Mirza et al., 2021; Nguyen Thi Thu, 2022). However, students' reading habits and study skills may be initiated from the school level enabling the individuals to grow in competence, comfort, and understanding of the audience. At the same time, previous researchers focused on university level students' reading habits. Second, there are methodological identities that lead to dubious findings not confirming the influence of reading habits and study skills on academic achievement (Goel, 2014; Lawrence, 2014; Quadir and Chen, 2015; Sherifat and Murthy, 2016; Ameyaw and Anto, 2018; Silverrajo and Hassan, 2018; Balan et al., 2019; Hassan et al., 2021). In general, there is a scarcity of research aiming to determine the correlation between students' reading habits and achievement through the role of study skills at any academic level. Finally, in Pakistan, few studies could explore reading habits as a singular variable of different groups of students (Bajwa et al., 2011; Hussain and Munshi, 2011; Rasheed, 2012). Numerous researchers only examined the relationship between reading habits and academic achievement (Bashir and Mattoo, 2012; Bibi et al., 2020; Ehsan and Sultana, 2020). Moreover, Fazal et al. (2012) only investigated the association between study skills and achievement. Thus, this research examines the relationship among students' reading habits,

study skills, and academic achievement in English as practiced at the secondary level in Lahore, Punjab, Pakistan.

Lahore is the capital of Pakistan's Punjab province. In terms of population, this is the second largest city in Pakistan. It is located in the northeastern part of Pakistan's Punjab province. Lahore is one of the most cosmopolitan cities in Pakistan and is home to various cultures, traditions, and customs. Specifically, it provides researchers with opportunities to contextualize perspectives in light of academic processes and ethics.

2. Literature review

2.1. Reading habits

Rosli et al. (2018) suggested that reading is an attempt to comprehend the writer's message, while Alnahdi and Aftab (2020) stated that it is a gateway to all other information, which may lead to understanding the world outside the text. Hence, Al-Jarf (2021) and Dadzie (2008) asserted that reading is a multifaceted cognitive process of comprehending words written in a textual form that allows readers to enhance their knowledge for personal growth and academic success. Moreover, Ogeyik and Akyay (2009), Erguvan (2016), Mirza et al. (2021) stated that reading is just a method of communication between the writer and the reader. Thus, Bhan and Gupta (2010) and Baron (2017) assumed that reading is the art of decoding and interpreting messages from various written materials such as books, magazines, journals, newspapers, dictionaries, encyclopedias, pamphlets, and diaries. Hassan et al. (2021) stated that reading habits influence reading materials, activities, time duration, place of reading, and reader motivation. In this study, reading habits are considered to be the students' reading preferences, interest in reading, attitude toward reading, and reading problems during study at the secondary level.

2.2. Study skills

Study skills are the readers' inclination toward organizing, highlighting, reviewing, reciting, and using devices, flashcards, etc. to comprehend new knowledge effectively (DiPerna and Elliott, 2000; DiPerna, 2006; Rozalski, 2008; Madhavi et al., 2014; Sabbah, 2016). While reading habit is the frequency, a reader regularly reads (Winne, 2013). Moreover, study skills are the students' intellectual practices to process new information effectively and efficiently, while reading habits are considered a psychological trait of one's personality (Farrington et al., 2012; Pillai, 2012; Mansor et al., 2013; Shahidi et al., 2014; Ameyaw and Anto, 2018; Rosli et al., 2018). Thus, the concept of study skills is different from reading habits. This research defines study skills as secondary school students' approaches to comprehending new knowledge.

2.3. Reading habits and academic achievement

Horbec (2012) and Singh (2011) determined a significant positive relationship between students' reading habits and academic achievement. Hence, Issa et al. (2012) explored that students'

reading patterns vary and have a moderately significant influence on academic success, while Bashir and Mattoo (2012) examined that academic performance is dependent on the level of students' study habits; thus, reading habits influence on future success, which was confirmed by Owusu-Acheaw and Larson (2014) through quantitative measures. Chotitham and Wongwanich (2014) found a moderate positive relationship between students' study habits and achievement. However, Lawrence (2014) rejected the association between students' academic achievement and study habits, and Goel (2014) confirmed that study habits do not influence academic performance. Therefore, Schwabe et al. (2015), Quadir and Chen (2015) concluded through a quantitative correlational study that heavy reading habits significantly impact reading efficiency; the longer the reading time, the better the results. Malik and Parveen (2016) discovered significant differences in low- and high-academic achievers' attitudes toward study habits. They determined that high achievers are more concentrated and exhibited better study habits, good time management skills, and punctuality compared to low achievers. In the meantime, Sherafat and Murthy (2016) directed that study habits facilitate learners toward higher achievement because of their significant connections with academic achievement. Consequently, Silverrajoo and Hassan (2018) revealed divergent findings that students' reading methods have a negative and weak relationship with academic achievement.

Ameyaw and Anto (2018) recognized the importance of reading styles in students' learning and found that reading styles affect students' performance. Meanwhile, Alzahrani et al. (2018) verified that students' reading styles significantly impact their performance. Dolmaz and Kaya (2019) discovered that students' creative writing skills are affected by their reading styles. Moreover, Balan et al. (2019) determined that students' purpose of reading significantly affected their performance, as Annamalai and Muniandy (2013) suggested that academic performance is based on students' reading purpose. Hence, Whitten et al. (2016) and Fatiloro et al. (2017) discovered that reading habits significantly assist students in learning and enhancing their performance. Bibi et al. (2020) examined that students' study habits were significantly positively associated with achievement. Ehsan and Sultana (2020) predicted that reading habits significantly increase students' performance. Moreover, Hassan et al. (2021) found a significant correlation between secondary school students' reading habits and their reading achievement and concluded that reading habits significantly contribute to academic achievement. Thus, Nguyen Thi Thu (2022) revealed that reading habits have a significant role in the development of students writing performance.

On the other hand, by designing a correlational study, Tonka and Bakir (2020) found a negative relationship between reading anxiety and reading habits. Thus, they concluded that reading anxiety plays a role in students' performance and reading habits. Similarly, Alnahdi and Aftab (2020) found a significant negative association between study habits and academic stress, reading habits, and academic achievement. The researchers measured all the variables through a questionnaire consisting of four scales and 43 items.

2.4. Study skills and academic achievement

Nouhi et al. (2009) determined that study skills have a significant positive association with academic success measured through a

closed-ended questionnaire confirmed by Awang and Sinnadurai (2011) through an experimental study. Meanwhile, Hassanbeigi et al. (2011) and Sabbah (2016) verified that study skills are critical for academic success because they positively correlate with academic achievement found through a descriptive correlational survey using a study skills scale. Hence, Fazal et al. (2012) suggested that higher academic achievers use a wide range of study skills than low achievers, while there was a weak correlation between study skills and academic success. Furthermore, Demir et al. (2012) revealed through an experimental study that students' study skills had a considerable influence on performance which was also confirmed by Wernersbach et al. (2014) from an experimental study. In both of the studies, researchers measured study skills through closed-ended items. Moreover, they also discovered that study skills significantly impact students' academic self-efficacy. Nonetheless, Tahamtani et al. (2017) and Naqvi et al. (2018) revealed a weak negative link between achievement and study habits through quantitative measures that were rejected by Gormley et al. (2018), who found a significant positive impact of study skills on achievement.

Several gaps were found in already conducted studies; first, the researchers mainly focused on the influence of reading habits and study skills on academic achievement separately in Western countries, while few addressed this phenomenon in Eastern countries. However, the investigation was restricted to university level students. Second, methodological identities lead to dubious findings not confirming this phenomenon. Finally, in Pakistan, few studies could explore reading habits as a singular variable of different groups of students. At the same time, some researchers only examined the relationship between reading habits and academic achievement. Thus, this study aimed to develop our understanding of the relationship between students' reading habits, study skills, and academic achievement in English.

2.5. Summary

Reading is an attempt to comprehend the writer's message for personal growth and success. Thus, psychologists consider it a multifaceted cognitive process of constructing meaning from texts. Bhan and Gupta (2010) stated that reading is the art of decoding and interpreting messages from the content of the written material that is often carried out in magazines, journals, newspapers, books, dictionaries, encyclopedias, pamphlets, diaries, and so on. While reading habits are the degree to which a reader engages in reading while studying skills to gain knowledge. Reading habits assist students in learning more, whereas study skills encourage them to understand new information effectively. Both reading habits and study skills influence students' academic performance.

Based on literature insights, the following hypotheses are formulated:

Hypothesis (H₁): A significant relationship exists between students' reading habits and their academic achievement in English language comprehension.

Hypothesis (H₂): A significant relationship exists between students' study skills and their academic achievement in English language comprehension.

3. Methods

3.1. Design

A research design is comprised of numerous elements (i.e., research paradigm, research approach, research design, and data collection method that provide guidelines for carrying out the study (Creswell and Clark, 2017; Myers, 2019), while a correlational research design is used to determine the relationship between two or more than two variables (Cohen et al., 2018). Thus, a correlational research design of a quantitative approach (positivism paradigm) was used. At the same time, a cross-sectional survey method was applied to collect data about studied variables (i.e., reading habits, study skills, and academic achievement in English).

3.2. Sample

The sample comprised 10th-grade students enrolled in district Lahore's public sector high schools for the academic year 2019–2020. The inclusion criteria were those students who enrolled in the science section only. During the data collection, the total number of active students in both sections (i.e., science and arts) of 10th grade was 36,847 enrolled at 334 high schools in district Lahore (Government of Punjab [GOP], 2019). While in the science section, the active students were 17,028, considered an accessible population of this study. A total of 1,800 (900 boys and 900 girls) were selected from 40 high schools through a non-proportionate random sampling technique that was 10.57% of the accessible population, which shows the sample was normally distributed. Out of 1,800 selected students, 1,619 participated as respondents because 181 students had not passed the subject English in the ninth-grade annual examination conducted by the Board of Intermediate and Secondary Education (BISE) Lahore. Three students declined to participate in this survey, while two could not complete the questionnaires. Therefore, the final sample consisted of $n = 1,614$ secondary school students.

3.3. Instruments

Two instruments were used, i.e., the Reading Habits Questionnaire (RHQ) and Study Skills Scale (SSS), to collect data about students' reading habits and study skills.

Reading Habits Questionnaire (RHQ): The researchers developed a paper and pencil student self-report RHQ based on Ajzen's (1991) Theory of Planned Behavior (TPB) and the social-cognitive theory of self-regulated learning strategies (Pintrich et al., 1993; Duncan and McKeachie, 2005; Duncan et al., 2007). Ajzen's (1991) TPB suggests that socio-psychological characteristics of a person's behavior, such as reading, influence reader proximal behaviors (Stokmans, 1999; Miesen, 2003; Van Schooten et al., 2004), while the social-cognitive theory of self-regulating learning strategies suggested that students' reading habits are meta-cognitively and behaviorally active in a student's learning process to achieve goals (Eccles and Wigfield, 2002). The classical test theory model was utilized to develop RHQ, which initially consisted of 44 closed-ended items. Each item was constructed on a 5-point Likert-type agreement scale

ranging from 1 (strongly disagree) to 5 (strongly agree), which means developing level reading habits to advance level reading habits. However, content validity was ensured by five education and assessment experts to validate the content coverage, language appropriateness, and usability of RHQ at the secondary level. Moreover, a pilot study was conducted on 250 students selected purposively from the target population to confirm unidimensionality among items and scales through exploratory factor analyses (EFAs) using Statistical Package for Social Sciences (SPSS) Version 23 software. Four subscales of RHQ (i.e., preferences for reading, interest in reading, attitude toward reading, and reading problems) were constructed during EFA. In contrast, nine items (two to three from each subscale) were deleted because their factor loading values (λ) were less than 0.5. In an analysis of items, reliability was also determined through Cronbach's alpha ($\alpha = 0.821$) value which was statistically acceptable. Psychometric evidence shows that RHQ was reliable for determining students' reading habits. Improved RHQ consisted of 35 items based on four subscales, i.e., preferences of reading (10 items), interest in reading (nine items), attitude toward reading (nine items), and reading problems (seven items).

Study Skills Scale (SSS): The researchers adopted the SSS from Academic Competence and Evaluation Scale, developed by DiPerna and Elliott in 2000. The validity, as well as reliability of SSS, was confirmed by numerous researchers (Kettler et al., 2014; Strunk, 2014; Anthony and DiPerna, 2018) and concluded that SSS is a standardized scale to measure study skills. The SSS consisted of 11 items that were also constructed on a 5-point Likert-type frequency scale, ranging from 1 (never) to 5 (almost always) which means developing level to advance level skills (DiPerna and Elliott, 2000). The SSS was also administered to 250 students to ensure reliability through Cronbach's alpha tests and found a value of $\alpha = 0.874$ that was suitable to measure study skills in the local context (Pakistan).

Academic Achievement: Students' marks obtained in ninth grade in the subject of English were asked them that determined by the Board of Intermediate and Secondary Education (BISE) Lahore in 2019. Their obtained scores in the subject of English were considered an academic achievement of students.

3.4. Data collection and analysis

After getting consent from the district education administration officer, the researchers personally gained permission from the selected schools' principals and class teachers for data collection. All the selected students were informed in their classes about the study purpose and given the right to withdraw from the study at any time before data analyses. RHQ and SSS administration occurred over 8 weeks during mid of the September to mid of November 2019 academic year. Before administering the instruments, participants were informed about the confidentiality procedures. Moreover, the researchers encouraged them to respond honestly and told them to write about their obtained marks in ninth grade in English. After collecting the questionnaires, the researchers quickly scanned the participants' responses about missing answers, and the students were asked to complete the responses in the questionnaire. A total of 1,614 participants provided valuable responses about their reading habits and study skills. The response rate was 89.6%, acceptable in social sciences research

for quantitative data. Students' responses about reading habits and study skills were analyzed by applying descriptive statistics (i.e., mean, standard deviation, skewness, and kurtosis) and inferential statistics (i.e., Pearson r test and regression test) through SPSS version 25 software.

4. Results

4.1. Descriptive statistics on students' reading habits and study skills

The Kolmogorov–Smirnov test was used to check the normality of data. At the same time, the skewness and kurtosis values indicated that data were normally distributed because skewness and kurtosis were between -2 and $+2$, which was suitable for parametric statistics (George, 2011; Albers, 2017; Mishra et al., 2019). Table 1 indicates that students give more preferences to reading than their attitude toward reading, interest in reading, and reading problems as $M = 4.13$, $SD = 0.486$; $M = 3.91$, $SD = 0.616$; $M = 3.68$, $SD = 0.676$; and $M = 3.57$, $SD = 0.813$, respectively. Students also thought they were facing reading problems because English was not their native language as $M = 3.68$; $SD = 0.813$. Overall, results revealed that students have permissive and desired reading habits and study skills as $M = 3.88$, $SD = 0.455$ and $M = 3.85$, $SD = 0.602$.

4.2. Inferential statistics on students' reading habits, study skills, and academic achievement in English

The Pearson product–moment correlation analysis was applied to test null hypotheses, i.e., whether there are significant relationships among students' reading habits, study skills, and academic achievement in English.

4.3. Hypotheses testing

Refer to Table 2 for the intercorrelation of variables among reading habits subscales, overall reading habits, and study skills; there were moderate-to-high positive correlations among variables. For reading habits, study skills, and academic achievement in English, students' reading habits were significantly and positively correlated as $r = 0.314$ – 0.721 . A value of $r = (\pm) 0.3$ – 0.7 exhibits a moderate-to-high correlation between variables (Akoglu, 2018;

Schober et al., 2018). Results also reveal higher positive correlations between reading habits and study skills while moderate positive correlations between reading habits and academic achievement in English as $r = 0.848$ and 0.584 , respectively. Moreover, there was a higher positive correlation between study skills and academic achievement in English as $r = 0.721$. Thus, it is revealed that students' reading habits and study skills are positively associated with academic achievement in English.

Since the hypotheses “there is a significant relationship between students' reading habits and academic achievement in English, and there is a significant relationship between students' study skills and academic achievement in English” were accepted because moderate-to-strong positive relationships were found among students' reading habits, study skills, and academic achievement in English.

A regression analysis was conducted to explore whether students' reading habits and study skills predict academic achievement in English. Students' reading habits and study skills served as independent variables, while students' academic achievement in English served as dependent variables. The regression analysis results were significant. The unique individual predictor for students' academic achievement in English interested in reading and attitude toward reading. These two sub-factors of reading habits significantly predicted 42 and 43% of the variance, respectively.

In contrast, the numeric regression does not considerably reveal the remaining two sub-factors (preferences of reading and reading problems). However, students' reading habits accounted for 44% of the variance, and study skills accounted for 48% of the variance. Refer to Table 3 for unstandardized betas, standard errors, standardized betas, and adjusted R^2 . The independent variables in these analyses are moderately correlated and predict academic achievement because the variance inflation factor (VIF) estimation was below 5.0 in regression.

5. Discussion

Reading habit is a crucial aspect of creating a literate society because it helps to shape personality, develop creative and critical thinking abilities, and enhance knowledge (Palani, 2012; Mansor et al., 2013; Fischer et al., 2015; Bano et al., 2018; Rosli et al., 2018; Al-Jarf, 2019; Wu et al., 2019; Hassan et al., 2021). At the same time, study skills are the readers' strategies to process new information effectively (Kuterbach, 2012; Anthony and DiPerna, 2018; Abid et al., 2021). Both reading habits and study skills are

TABLE 1 Descriptive statistics of reading habits and study skills.

Scales	Mean	SD	Skewness	Kurtosis
Preferences of reading	4.13	0.486	−0.632	0.623
Interest in reading	3.68	0.676	−0.871	0.546
Attitude toward reading	3.91	0.616	−1.027	1.162
Reading problems	3.68	0.813	−1.024	1.219
Reading habits-total score	3.88	0.455	−0.987	1.047
Study skills	3.85	0.602	−0.893	1.016

$N = 1614$.

TABLE 2 Intercorrelations matrix and relationships among students' reading habits, study skills, and academic achievement in English.

	PR	IR	AR	RP	RHT	SS	AAE
PR	–						
IR	0.703**	–					
AR	0.761**	0.644**	–				
RP	0.781**	0.752**	0.734**	–			
RHT	0.862**	0.833**	0.819**	0.808**	–		
SS	0.883**	0.875**	0.867**	0.859**	0.848**	–	
AAE	0.591**	0.713**	0.418**	0.314**	0.584**	0.721**	–

PR, preferences of reading; IR, interest in reading; AR, attitude toward reading; RP, reading problems; RHT, reading habits total score; SS, study skills; AAE, academic achievement in English; N , 1614; and **correlation is significant at the 0.01 level (two-tailed).

TABLE 3 Summary of regression analyses, with 95% confidence intervals, of students' reading habits and study skills predicting academic achievement in English.

Dependent variable	Independent variables	<i>B</i>	Lower	Upper	<i>SE B</i>	β	Adjusted R^2	<i>p</i>
AAE	PR	0.015	−0.13	0.042	0.01	0.16	0.05	0.060
	IR	0.017	−0.010	0.039	0.01	0.18	0.42**	0.020
	AR	0.016	−0.014	0.044	0.01	0.19	0.43**	0.035
	RP	0.018	−0.16	0.038	0.01	0.25	0.05	0.087
	RHT	0.013	−0.010	0.038	0.01	0.21	0.44**	0.002
	SS	0.016	−0.013	0.043	0.01	0.27	0.48**	0.001

PR, preferences of reading; IR, interest in reading; AR, attitude toward reading; RP, reading problems; RHT, reading habits total score; SS, study skills; AAE, academic achievement in English; and ** $p < 0.001$.

interdependent and influence students' academic performance as well as future success (Demir et al., 2012; Wernersbach et al., 2014; Tahamtani et al., 2017; Alzahrani et al., 2018; Ameyaw and Anto, 2018; Gormley et al., 2018; Balan et al., 2019; Dolmaz and Kaya, 2019; Ehsan and Sultana, 2020). Therefore, this study is designed to examine relationships among students' reading habits, study skills, and academic achievement at the secondary level in Lahore, Pakistan. Lahore is one of the cosmopolitan cities of Pakistan and a hub of many cultures, traditions, and customs. Regarding the academic processes and ethics, it provides opportunities for researchers to contextualize the perspectives accordingly. Reading habits have been and are still being taught in schools, colleges, and universities through model reading by teachers, parents, or elders of the families. In addition, the reading and recitation of fold tales and poems get to gathers like at Pak Tea House, Lawrence Garden, Quaid-e-Azam Library, and so on, while formal schools books, extra reading exercises, and reading and writing competitions at the school level are prepared through a variety of book reading within the context of the particular objective. The results of normality tests indicated that the data were normality distributed and suitable to apply parametric statistics. The descriptive findings also showed that students have more preferences for reading than their attitude toward reading, interest in reading, and reading challenges. These results support the finding of numerous researchers (e.g., Pehlivan et al., 2010; Mansor et al., 2013; Owusu-Acheaw and Larson, 2014; Haliru et al., 2015; Erguvan, 2016; Krashen, 2016; Kulatunga, 2016; Loan and Shah, 2017; Ameyaw and Anto, 2018; Porkaew and Fongpaiboon, 2018; Mirza et al., 2021) who found that students give more preference to read academic content from textbooks and other reading materials (i.e., newspaper, storybooks, poetry, novel, magazines, cartoons, comics, sports, etc.). Rasheed (2012) determined that reading habits play a substantial essential role in developing positive attitudes toward reading. However, Maiyo and Siahi (2015) revealed that higher achievers had better reading habits than low achievers. Students prefer reading online because they can easily read content from the internet material in this technological age, so they prefer reading online (Dollah et al., 2017). Thus, Molotja and Themane (2018) found that students' reading habits may enhance through global reading strategies and problem-solving strategies. Moreover, it is found that students have competence in reading habits and study skills. These results are also in line with the findings of previous studies, e.g., Dadzie (2008), Ogeyik and Akyay (2009), Bhan and Gupta (2010), and Issa et al. (2012), and Sabbah (2016) revealed that the majority

of the students read books to pass the exams that why they have good reading habits. Furthermore, it is determined that students also possess the competence level of study skills that confirmed the study conducted by numerous researchers (i.e., DiPerna, 2004, 2006; Rozalski, 2008; Kuterbach, 2012; DuPaul and Stoner, 2014; Anthony and DiPerna, 2018; Abid et al., 2021).

Furthermore, researchers also concluded from correlational results that there were moderate-to-significant positive correlations among reading habits, study skills, and academic achievement in English. In the literature review, it is seen that these findings are consistent with the results of Singh (2011), Horbec (2012), Issa et al. (2012), Sabbah (2016), Ameyaw and Anto (2018), Hassan et al. (2021). They found a positive relationship between reading habits and academic achievement, while reading habits influence students' academic performance. Moreover, some researchers determined a moderate relationship between reading habits and academic success (e.g., Chotitham and Wongwanich, 2014; Kutay, 2014; Owusu-Acheaw and Larson, 2014; Alzahrani et al., 2018; Adigun et al., 2021; Nguyen Thi Thu, 2022). Sherafat and Murthy (2016) directed that study habits facilitate learners toward higher achievement because of their significant connections with academic achievement, that confirmed by Bibi et al. (2020). In contrast, few researchers found different results due to participants' different selection procedures and contextual differences (i.e., purposive sample method, content, reading material, culture, etc.). For example, Lawrence (2014), Goel (2014), Alnahdi and Aftab (2020) revealed no significant correlation between students' academic achievement and study habits. At the same time, Silverrajoo and Hassan (2018) found that students' reading styles have a negative, weak relationship with academic achievement. Findings regarding study skills: e.g., Nouhi et al. (2009), Awang and Sinnadurai (2011), Hassanbeigi et al. (2011), Maiyo and Siahi (2015), Gormley et al. (2018) revealed a significant positive connection between study skills and academic success, whereas Demir et al. (2012) and Wernersbach et al. (2014) found study skills have a considerable influence on performance that support the present study findings. However, few researchers found a negative correlation between study skills and academic performance (Fazal et al., 2012; Tahamtani et al., 2017; Naqvi et al., 2018). Furthermore, researchers determined in this study students' reading habits have positive correlations with study skills. Thus, reading habits and study skills directly correlate with their academic achievement in English. In addition, it is also revealed that reading habits and study skills moderately predict students'

academic achievement. Annamalai and Muniandy (2013) suggested that academic performance is based on students' reading habits. Whitten et al. (2016) and Fatiloro et al. (2017) revealed that reading habits significantly help students learn more to enhance their academic performance. Ehsan and Sultana (2020) predicted that reading habits significantly improve students' academic performance.

6. Conclusion

Reading habits and study skills differ in conceptual understanding. Reading habits are the degree to which readers regularly read, whereas study skills are the ability to comprehend new information effectively. Both reading and study habits influence students' academic performance. It is concluded that the collected data were normally distributed. The descriptive findings about reading habit sub-constructs indicated that students give more preferences to reading than their attitude toward reading, interest in reading, and reading problems. Simultaneously, they have competent reading habits and study skills. Furthermore, it is found that there are moderate-to-strong positive correlations among reading habits, study skills, and academic achievement in English. Thus, it is concluded that reading habits and study skills directly correlate with academic achievement in English. In addition, it is also revealed that reading habits and study skills moderately predict students' academic achievement.

7. Implications for practice

It is determined that students prefer reading to their attitude toward reading, interest in reading, and challenges and have competent reading habits and study skills. Thus, it is suggested that teachers plan such assignments and tasks based on reflective thinking (Aslam et al., 2021), so students have to visit the school library to read more academic material to accomplish assigned tasks through extensive reading. Students' reading habits and study skills have a moderate-to-strong connection with their academic achievement in English. So school administrations design a timetable by consulting with teachers, allowing students to spend at least an hour in the library regularly. In contrast, the library should have up-to-date reading material, exciting storybooks, and stock which attract students. In addition, parents can also engage their children in constant reading at home by providing related textbook materials and allowing them to watch educational television programs to gain the essence of reading habits and study skills.

8. Limitations and implications for future research

There are several limitations to this study. First, this study was conducted on secondary school students by selecting a sample from the Lahore district of Punjab, Pakistan. Therefore, future studies may include participants from other districts of Punjab and other provinces of Pakistan to increase the generalizability of results. Second, longitudinal studies are needed to explore the change in students reading habits and

study skills over time. To enhance reading habits and study skills, interventional studies may build lifelong reading habits and study skills among learners to make a scholarly society. Last but not least, future researchers may explore parent's role in developing their children's reading habits and study skills by selecting participants from diverse populations. Cultural factors would affect students' reading habits; thus, Pakistan's unique culture should be considered a potential theoretical explanation in future.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

Author contributions

NA presented the main idea and wrote the first draft of the manuscript. SA contributed to conducting the methodology. SA, AA, and TK were involved with the revisions and proofreading. All authors contributed to the article revisions and approved the submitted version.

Acknowledgments

The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this study by Grant Code: 22UQU4280253DSR01.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abid, N., Ali, R., and Akhter, M. (2021). Exploring gender-based difference towards academic enablers scales among secondary school students of Pakistan. *Psychol. Sch.* 58, 1380–1398. doi: 10.1002/pits.22538
- Adigun, I. O., Oyewusi, F. O., and Aramide, K. A. (2021). The impact of Covid-19 pandemic lockdown on reading engagement of selected secondary school students in Nigeria. *Interdiscip. J. Educ. Res.* 3, 45–55. doi: 10.51986/ijer-2021.vol3.01.05
- Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-T
- Akdoglu, H. (2018). User's guide to correlation coefficients. *Turkish J. Emerg. Med.* 18, 91–93. doi: 10.1016/j.tjem.2018.08.001
- Albers, M. J. (2017). *Introduction to quantitative data analysis in the behavioral and social sciences*. Hoboken, NJ: John Wiley & Sons, doi: 10.1002/9781119290384
- Al-Jarf, R. (2021). Collaborative mobile ebook reading for struggling EFL college readers. *IOSR J. Res. Method Educ.* 11, 32–42.
- Al-Jarf, R. (2019). Quality in teaching reading to high school students. *Eurasian Arabic Stud.* 5, 36–62.
- Alnahdi, A. S., and Aftab, M. (2020). Academic stress, study habits and academic achievement among university students in Jeddah [Special Issue]. *Int. J. Psychosoc. Rehabil.* 24, 97–104. doi: 10.37200/1JPR/V24SP1/PR201138
- Alzahrani, S. S., Soo Park, Y., and Tekian, A. (2018). Study habits and academic achievement among medical students: A comparison between male and female subjects. *Med. Teach.* 40, 1–9. doi: 10.1080/0142159X.2018.1464650
- Ameyaw, S. K., and Anto, S. K. (2018). Read or perish: Reading habits among students and its effect on academic performance: A case study of eastbank senior high school-accra. *Libr. Philos. Pract.* 1–23. Available online at: <https://digitalcommons.unl.edu/libphilprac/1748>
- Ameyaw, S., and Anto, S. K. (2019). Gender variation in reading habits in schools in Moland: A case study of Asantekwaa SDA Junior High School. *Eur. J. Educ. Stud.* 6, 250–264. doi: 10.5281/zenodo.3473923
- Annamalai, S., and Muniandy, B. (2013). Reading habit and attitude among Malaysian polytechnic students. *Int. Online J. Educ. Sci.* 5, 32–41.
- Anthony, C. J., and DiPerna, J. C. (2018). Piloting a short form of the academic competence evaluation scales. *Sch. Ment. Health* 10, 314–321. doi: 10.1007/s12310-018-9254-7
- Aslam, S., Hali, A. U., Zhang, B. H., and Saleem, A. (2021). *The Teacher Education Program's Impact on Preservice Teachers' Reflective Thinking in Pakistan*. Thousand Oaks, CA: Sage Open, doi: 10.1177/21582440211055724
- Awang, M., and Sinnadurai, S. K. (2011). A study on the development of strategic tools in study orientation skills towards achieving academic excellence. *J. Lang. Teach. Res.* 2, 60–67. doi: 10.4304/jltr.2.1.60-67
- Baffoe, G. A., and Okae-Anti, A. (2020). Reading habits of selected communication educators in Ghana. *J. Educ. Pract.* 11, 45–51.
- Bajwa, N., Gujjar, A., Shaheen, G., and Ramzan, M. (2011). A comparative study of the study habits of the students from formal and non-formal systems of education in Pakistan. *Int. J. Bus. Soc. Sci.* 2, 175–186.
- Balan, S., Katenga, J. E., and Simon, A. (2019). Reading habits and their influence on academic achievement among students at Asia pacific international university. *Abstr. Proc. Int. Scholars Conf.* 7, 1490–1516. doi: 10.35974/isc.v7i1.928
- Bano, J., Jabeen, Z., and Qutoshi, S. B. (2018). Perceptions of teachers about the role of parents in developing reading habits of children to improve their academic performance in schools. *J. Educ. Educ. Dev.* 5, 42–59. doi: 10.22555/joeed.v5i1.1445
- Baron, N. S. (2017). Reading in a digital age. *Phi Delta Kappan* 99, 15–20. doi: 10.1177/003172171773418
- Bashir, I., and Mattoo, N. H. (2012). A study on study habits and academic performance among adolescents (14–19) years. *Int. J. Soc. Sci. Tomorrow* 1, 1–5.
- Bhan, K. S., and Gupta, R. (2010). Study habits and academic achievement among the students belonging to scheduled caste and non-scheduled caste group. *J. Appl. Res. Educ.* 15, 1–9.
- Bibi, A., Naseer, N., and Habib, Z. (2020). Study habits of students and academic achievement: A correlational study. *Glob. Educ. Stud. Rev.* 5, 114–122. doi: 10.31703/gesr.2020(V-III).12
- Biyik, M. A., Erdogan, T., and Yildiz, M. (2017). The examining reading motivation of primary students in the terms of some variables. *Int. J. Prog. Educ.* 13, 31–49.
- Chothitham, S., and Wongwanich, S. (2014). The reading attitude measurement for enhancing elementary school students' achievement. *Procedia Soc. Behav. Sci.* 116, 3213–3217. doi: 10.1016/j.sbspro.2014.01.737
- Cohen, L., Manion, L., and Morrison, K. (2018). *Research methods in education*, 8th Edn. Milton Park: Routledge.
- Creswell, J. W., and Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Dadzie, P. S. (2008). Reading for education: The roles of libraries. *Ghana Libr. J.* 20, 1–14. doi: 10.4314/glj.v20i1.33978
- Davarci, N. (2013). *An investigation on the evaluation of the relationship between the reading habits of 8th grade elementary school students and their habits of computer-internet usage* [Unpublished master thesis]. Nigeria: University of Nigeria.
- Demir, S., Kilinc, M., and Dogan, A. (2012). The effect of curriculum for developing efficient studying skills on academic achievements and studying skills of learners. *Int. Electron. J. Elem. Educ.* 4, 427–440.
- Dilshad, M., Adnan, A., and Akram, A. (2013). Gender differences in reading habits of university students: An evidence from Pakistan. *Pak. J. Soc. Sci. (PJSS)* 33, 311–320.
- DiPerna, J. C. (2004). Structural and concurrent validity evidence for the academic competence evaluation scales-college edition. *J. Coll. Couns.* 7, 64–72. doi: 10.1002/j.2161-1882.2004.tb00260.x
- DiPerna, J. C. (2006). Academic enablers and student achievement: Implications for assessment and intervention services in the schools. *Psychol. Sch.* 43, 7–17. doi: 10.1002/pits.20125
- DiPerna, J. C., and Elliott, S. N. (2000). Academic competence evaluation scales. *Psychol. Corp.* doi: 10.1037/t14965-000
- Dollah, W. A. K. W., Fakeh, S. S. K. W., Kamal Rafedzi, E. R., Ibrahim, A., Rahim, H., Masron, M. Z. A., et al. (2017). Inculcating reading habits among secondary school students. *J. Sci. Eng. Res.* 4, 407–421.
- Dolmaz, M., and Kaya, E. (2019). The effect of 7th grade students' reading habits and their academic achievement in social studies and Turkish courses on their creative writing skills. *Int. Online J. Educ. Sci.* 11, 168–183. doi: 10.15345/iojes.2019.01.012
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Dev. Psychol.* 43, 1428–1446. doi: 10.1037/0012-1649.43.6.1428
- Duncan, T. G., and McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educ. Psychol.* 40, 117–128. doi: 10.1207/s15326985ep4002_6
- DuPaul, G. J., and Stoner, G. (2014). *ADHD in the schools: Assessment and intervention strategies*, 3rd Edn. New York, NY: Guilford Press.
- Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Ann. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153
- Ehsan, T., and Sultana, N. (2020). Predicting the role of study habits in academic achievement: A study of university students in Punjab. *Pak. J. Educ.* 37, 95–112. doi: 10.30971/pje.v37i1.1410
- Erdem, A. (2015). A research on reading habits of university students: Sample of Ankara University and Erciyes University. *Procedia Soc. Behav. Sci.* 174, 3983–3990. doi: 10.1016/j.sbspro.2015.01.1145
- Erguvan, D. (2016). 'Students' attitudes towards extensive and intensive reading and 'instructors' motivational strategies. *Arab World Eng. J. (AWEJ)* 7, 136–150.
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., et al. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.
- Fatilor, O. F., Adesola, O. A., Hameed, B. A., and Adewumi, O. M. (2017). A survey on the reading habits among colleges of education students in the information age. *J. Educ. Pract.* 8, 106–110.
- Fazal, S., Hussain, S., Majoka, M. I., and Masood, S. (2012). The role of study skills in academic achievement of students: A closer focus on gender. *Pak. J. Psychol. Res.* 27, 37–51.
- Fischer, L., Hilton, J., Robinson, T. J., and Wiley, D. A. (2015). A multi-institutional study of the impact of open textbook adoption on the learning outcomes of post-secondary students. *J. Comput. High. Educ.* 27, 159–172. doi: 10.1007/s12528-015-9101-x
- George, D. (2011). *SPSS for Windows step by step: A simple study guide and reference, 17.0 update*, 10th Edn. London: Pearson Education.
- Goel, U. (2014). Comparative study of study habits in relation to academic achievement of senior secondary school students. *Gyanodaya J. Prog. Educ.* 7, 18–25. doi: 10.5958/2229-4422.2014.00004.8
- Gormley, M. J., Pinho, T., Pollack, B., Puzino, K., Franklin, M. K., Busch, C., et al. (2018). Impact of study skills and parent education on first-year GPA among college students with and without ADHD: A moderated mediation model. *J. Atten. Disord.* 22, 334–348. doi: 10.1177/1087054715594422
- Government of Punjab [GOP] (2019). *Census of school education department*. Lahore: Government of Punjab.
- Habibu, I., and Ejembi, S. (2011). The role of schools and public libraries in promoting reading habit among children and adolescents in Nigeria. *Inf. Knowl. Manag.* 1, 33–40.
- Haliru, R. A., Abdulkarim, M., Mohammed, A. D., and Dangani, B. U. (2015). An assessment of reading habit among secondary school students in Kaduna metropolis. *J. Hum. Soc. Sci.* 20, 12–17.
- Hassan, I., Latif Azmi, M. N., Muhamad, S. N., and Abdullah, A. T. H. (2021). Reading habits and their correlation with reading achievement among ESL learners in selected

- Malaysian secondary schools. *Arab World Eng. J. (AWEJ)* 12, 385–399. doi: 10.24093/awej/vol12no3.27
- Hassanbeigi, A., Askari, J., Nakhjavani, M., Shirkhoda, S., Barzegar, K., Mozayyan, M. R., et al. (2011). The relationship between study skills and academic performance of university students. *Procedia Soc. Behav. Sci.* 30, 1416–1424. doi: 10.1016/j.sbspro.2011.10.276
- He, M. (2014). “Extensive reading and students’ academic achievement: A case study,” in *Exploring EFL fluency in Asia*, eds T. Muller, J. Adamson, P. S. Brown, and S. Herder (Berlin: Springer), 231–243. doi: 10.1057/9781137449405_14
- Horbec, D. (2012). The link between reading and academic success. *Eng. Aust.* 47, 58–67.
- Hussain, I., and Munshi, P. (2011). Identifying reading preferences of secondary school students. *Creat. Educ.* 2, 429–434. doi: 10.4236/ce.2011.25062
- Iheakanwa, J. U., Obro, S., and Akpochofa, W. P. (2021). Reading ability, study habits and students’ academic performance in social studies. *Libr. Philos. Pract.* 1–21. Available online at: <https://digitalcommons.unl.edu/libphilprac/5675>
- Issa, A. O., Aliyu, M. B., Akangbe, R. B., and Adedeji, A. F. (2012). Reading interests and habits of the federal polytechnic, OFFA, students. *Int. J. Learn. Dev.* 2, 470–486. doi: 10.5296/ijld.v2i1.1470
- Kettler, R. J., Elliott, S. N., DiPerna, J. C., Bolt, D. M., Reiser, D., and Resurreccion, L. (2014). Student and teacher ratings of academic competence: An examination of cross-informant agreement. *J. Appl. Sch. Psychol.* 30, 338–354. doi: 10.1080/15377903.2014.950442
- Krashen, S. (2016). “Compelling reading and problem-solving: The easy way (and the only way) to high levels of language, literacy, and life competence,” in *Proceedings of the epoch making in english language teaching and learning twenty-fifth international symposium on english teaching, english teachers*, Taipei, 115–125.
- Kulatunga, R. K. (2016). *A study on understanding the reading habits and library usage of under graduate students (2007/2008 batch) of Uva Wellasa University of Sri Lanka*. [Unpublished doctoral dissertation]. Badulla: Uva Wellasa University of Sri Lanka.
- Kutay, V. (2014). *A survey of the reading habits of Turkish high school students and an examination of the efforts to encourage them to read* [Doctoral dissertation]. Loughborough: Loughborough University.
- Kuterebach, J. M. (2012). *A model of academic enablers and academic performance among postsecondary learners* [Unpublished doctoral dissertation]. Pennsylvania: Penn State University.
- Lawrence, A. S. (2014). Relationship between study habits and academic achievement of higher secondary school students. *Online Submission* 4, 143–145. doi: 10.15373/2249555X/June2014/43
- Loan, F. A., and Shah, R. (2017). Survey of the literature reading habits and preferences of adolescents: A study of a public school in India. *Libr. Inform. Sci. Res. Electron. J.* 27, 80–96.
- Madhavi, S., Naidu, S., Krishnaveni, A., and Kiran, P. (2014). Study skills assessment among medical undergraduates-where they stand. *J. Dent. Med. Sci.* 13, 16–19. doi: 10.9790/0853-131031619
- Maiyo, J., and Siahi, E. A. (2015). Study of the relationship between study habits and academic achievement of students: A case of Spicer Higher Secondary School, India. *Int. J. Educ. Admin. Pol. Stud.* 7, 134–141. doi: 10.5897/IJEAPS2015.0404
- Malik, M., and Parveen, N. (2016). Study habits and academic achievement: A comparative analysis of the high and low academic achievers. *Bahria J. Prof. Psychol.* 15, 46–54.
- Mansor, A. N., Rasul, M. S., Abd Rauf, R. A., and Koh, B. L. (2013). Developing and sustaining reading habit among teenagers. *Asia Pac. Educ. Res.* 22, 357–365. doi: 10.1007/s40299-012-0017-1
- Miesen, H. W. J. M. (2003). Predicting and explaining literary reading: An application of the theory of planned behavior. *Poetics* 31, 189–212. doi: 10.1016/S0304-422X(03)00030-5
- Mirza, Q., Pathan, H., Khatoon, S., and Hassan, A. (2021). Digital age and reading habits: Empirical evidence from Pakistani Engineering University. *TESOL Int. J.* 16, 210–231.
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., and Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Ann. Card. Anaesth.* 22, 67–72. doi: 10.4103/aca.ACA_157_18
- Molotja, T. W., and Themane, M. (2018). Enhancing learners’ reading habits through reading bags at secondary schools. *Read. Writ.* 9, 1–9. doi: 10.4102/rw.v9i1.185
- Myers, M. D. (2019). *Qualitative research in business and management*. Thousand Oaks, CA: Sage.
- Naqvi, S., Chikwa, G., Menon, U., and Al Kharusi, D. (2018). Study skills assessment among undergraduate students at a private university college in Oman. *Mediterr. J. Soc. Sci.* 9, 139–147. doi: 10.2478/mjss-2018-0034
- Nguyen Thi Thu, H. (2022). The effects of reading habits on writing performance: A case study at Van Lang University. *Int. J. TESOL Educ.* 2, 105–133. doi: 10.54855/ijte.22247
- Nouhi, E., Shakoobi, A., and Nakhei, N. (2009). Study habits and skills, and academic achievement of students in Kerman University of medical sciences. *J. Med. Educ.* 12, 77–80. doi: 10.22037/jme.v12i13.4.1249
- Ogeyik, M. C., and Akyay, E. (2009). Investigating reading habits and preferences of student teachers at foreign language departments. *Int. J. Lang. Soc. Cult.* 28, 72–78.
- Owusu-Acheaw, M., and Larson, A. G. (2014). Reading habits among students and its effect on academic performance: A study of students of Koforidua polytechnic. *Libr. Philos. Pract.* 1, 1–22.
- Oyewole, O. (2017). Impact of poor reading culture among selected secondary school students in Owo local government area of Ondo state, Nigeria. *Dev. Country Stud.* 7, 88–101.
- Palani, K. K. (2012). Promoting reading habits and creating a literate society. *J. Arts Sci. Commerce* 3, 90–94.
- Pehlivan, A., Serin, O., and Serin, N. B. (2010). Determining reading interests and habits of candidate teachers (TRNC Sample). *Procedia Soc. Behav. Sci.* 9, 869–873. doi: 10.1016/j.sbspro.2010.12.251
- Pillai, S. K. (2012). An empirical study on study habits of X standard students in Nagarkovil district. *Res. Expo Int. Multidiscip. Res. J.* 2, 15–27.
- Pintrich, P. R., Smith, D. A., Garcia, T., and McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educ. Psychol. Meas.* 53, 801–813. doi: 10.1177/0013164493053003024
- Porkaew, K., and Fongpaiboon, A. (2018). Effects of extensive reading on Thai tertiary students’ reading attitudes. *Arab World Eng. J. (AWEJ)* 9, 207–219. doi: 10.24093/awej/vol9no1.15
- Pretorius, E. J., and Klapwijk, N. M. (2016). Reading comprehension in South African schools: Are teachers getting it, and getting it right? *Per Linguam* 32, 1–20. doi: 10.5785/32-1-627
- Quadir, B., and Chen, N. S. (2015). The effects of reading and writing habits on learning performance in a blog learning environment. *Asia Pac. Educ. Res.* 24, 635–644. doi: 10.1007/s40299-014-0210-5
- Rasheed, S. (2012). Children’s reading habits: A study of Lahore city. *Pak. J. Inform. Manag. Libr.* 5, 412–437.
- Rosli, N. A., Razali, N. F., Zamil, Z. U. A., Noor, S. N. F. M., and Baharuddin, M. F. (2018). The determination of reading habits among students: A concept. *Int. J. Acad. Res. Bus. Soc. Sci.* 7, 791–798. doi: 10.6007/IJARBS/v7-i12/3710
- Rozalski, M. E. (2008). Practice, practice, practice: How to improve students’ study skills. *Beyond Behav.* 17, 17–23.
- Sabbah, S. (2016). The effect of study habits on English language achievement. *Arab World Eng. J. (AWEJ)* 7, 238–257.
- Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesth. Analg.* 126, 1763–1768. doi: 10.1213/ANE.0000000000002864
- Schwabe, F., McElvany, N., and Trendtel, M. (2015). The school age gender gap in reading achievement: Examining the influences of item format and intrinsic reading motivation. *Read. Res. Q.* 50, 219–232. doi: 10.1002/rrq.92
- Shahidi, F., Dowlatkhan, H. R., Avand, A., Musavi, S. R., and Mohammadi, E. (2014). A study on the quality of study skills of newly-admitted students of Fasa University of Medical Sciences. *J. Adv. Med. Educ. Prof.* 2, 45–50.
- Sherafat, R., and Murthy, C. V. (2016). A study of study habits and academic achievement among secondary and senior secondary school students of Mysore city. *Int. J. Indian Psychol.* 3, 161–170. doi: 10.25215/0302.055
- Silverrajoo, P., and Hassan, A. (2018). Relationship between study habits and academic achievement among health science students. *Int. J. Acad. Res. Bus. Soc. Sci.* 8, 763–780. doi: 10.6007/IJARBS/v8-i7/4418
- Singh, Y. G. (2011). Academic achievement and study habits of higher secondary students. *Int. Ref. Res. J.* 3, 27–42.
- Stokmans, M. J. W. (1999). Reading attitude and its effect on leisure time reading. *Poetics* 26, 245–261. doi: 10.1016/S0304-422X(99)00005-4
- Strunk, T. A. (2014). *An exploration of the relationships between academic enablers and middle school achievement* [Unpublished doctoral dissertation]. Pennsylvania: Penn State University.
- Tahamtani, T., Jalil, K., Hosseini, M., and SoltaniArabshahi, K. (2017). Correlation of study habits with academic achievement among students attending the national medical science Olympiad. *J. Adv. Med. Educ.* 3, 19–23.
- Thamarasseri, I. (2018). Cognitive styles, study habits and academic achievement of students of Central University of Kashmir. *Stud. Home and Commun. Sci.* 12, 9–20. doi: 10.31901/24566780.2018.12.1-2.335
- Tonka, H., and Bakir, S. (2020). The examination of the relationship between the secondary school students’ habit of reading and their reading anxiety. *J. Educ. Issues* 6, 293–313. doi: 10.5296/jei.v6i1.16986
- Van Schooten, E., De Gloppe, K., and Stoel, R. D. (2004). Development of attitude toward reading adolescent literature and literary reading behavior. *Poetics* 32, 343–386. doi: 10.1016/j.poetic.2004.07.001

- Walia, P. K., and Sinha, N. (2014). Changing trend in reading habits of teenagers in Delhi: An impact assessment of demographic and environmental variables. *Libr Rev.* 63, 125–137. doi: 10.1108/LR-03-2013-0038
- Wernersbach, B. M., Crowley, S. L., Bates, S. C., and Rosenthal, C. (2014). Study skills course impact on academic self-efficacy. *J. Dev. Educ.* 37, 14–33.
- Whitten, C., Labby, S., and Sullivan, S. L. (2016). The impact of pleasure reading on academic success. *J. Multidiscip. Grad. Res.* 2, 48–64.
- Winne, P. H. (2013). “Learning strategies, study skills, and self-regulated learning in postsecondary education,” in *Higher education: Handbook of theory and research*, ed. M. Paulsen (Berlin: Springer), 377–403. doi: 10.1007/978-94-007-5836-0_8
- Wu, L., Valcke, M., and Van Keer, H. (2019). Factors associated with reading comprehension of secondary school students. *Educ. Sci. Theory Pract.* 19, 34–47. doi: 10.12738/estp.2019.4.003



OPEN ACCESS

EDITED BY

Evely Boruchovitch,
State University of Campinas,
Brazil

REVIEWED BY

Reza Pishghadam,
Ferdowsi University of Mashhad,
Iran
Federico Batini,
University of Perugia,
Italy

*CORRESPONDENCE

Manuel León Urrutia
✉ m.Leon-Urrutia@soton.ac.uk

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 31 October 2022

ACCEPTED 12 January 2023

PUBLISHED 06 February 2023

CITATION

Jiménez-Pérez EDP, de Vicente-Yagüe Jara MI,
León Urrutia M and García Guirao P (2023)
Emotions and reading: When reading is the
best way to improve skills in adolescents.
Front. Psychol. 14:1085945.
doi: 10.3389/fpsyg.2023.1085945

COPYRIGHT

© 2023 Jiménez-Pérez, de Vicente-Yagüe Jara,
León Urrutia and García Guirao. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Emotions and reading: When reading is the best way to improve skills in adolescents

Elena del Pilar Jiménez-Pérez¹, María Isabel de Vicente-Yagüe Jara²,
Manuel León Urrutia^{3*} and Pedro García Guirao⁴

¹Department of Didactics of Language, Art and Sports, University of Málaga, Málaga, Spain, ²Department of Didactics of Language, University of Murcia, Murcia, Spain, ³Department of Computer Science, University of Southampton, Southampton, United Kingdom, ⁴Department of Linguistics, WSB University in Wrocław, Wrocław, Poland

In recent years much research on reading competence in different languages has been published in parallel with the interest generated by the results of the PISA and PIRLS reports which were disseminated in the media and which have subsequently garnered the attention of public authorities. Studies that relate reading competence with emotional intelligence, however, are less frequent. This study aims to deepen the relationship between both constructs, using a quasi-experimental longitudinal approach that observes the evolution of 389 high school students in Spain from 16 to 18 years old. Evidence of a direct relationship between reading competence and emotional intelligence was obtained, particularly in the experimental group in which reading habits were stimulated.

KEYWORDS

emotional intelligence, reading competence, adolescents, quantitative research, reading habits, reading intervention

Introduction

Reading competence (RC) and emotional intelligence (EI) constructs have been widely studied in the scientific literature in recent years, as well as their importance in the field of education. Reading competence is a skill that can be specifically taught for better development, but it is a competence which is refined naturally through a person's reading habits since readers with stronger reading habits score better in reading skills (Jiménez-Pérez, 2014; OECD, 2016): training as a receiver qualifies as an issuer (Echeverría, 1998). In addition, it has been demonstrated that reading for pleasure provides students with a better ability to get good grades (Sullivan and Brown, 2015), that reading lengthens life (Bavishi et al., 2016), and that reading books outside of class when at the school stage of life predicts professional success in a positive way (Taylor, 2016). Reading also offers a long list of advantages such as reducing stress, improving behavior, and helping people to empathize and to make more creative decisions (Rodríguez, 2016).

With regard to reading competence and comprehension, there is still a long way to go in terms of research. For example, it should be noted that there is no scientific literature on the improvement of reading comprehension in Spanish since "not even serious attempts have been made to make this identification through reviews of the previous research" (Ripoll Salceda, 2014). Furthermore, there are other issues related to reading habits such as, for example, the impact of information and communication technologies (ICT) on people's ways of reading involves "the need to attend to new visual skills and to new reading skills" (Sánchez-Claros, 2016), or that motivating reading by awakening curiosity is more effective than using reading or motivation strategies (Gurning, 2017), such as through ICT, which facilitate the general acquisition of professional and personal

competences (Martínez-Cerdá and Torrent-Sellens, 2017) and thereby stimulate higher involvement on the part of students than they had before the digital world (Alonso-Mosquera et al., 2016). Therefore, if reading habits change, either in the devices for reading (Cordón-García, 2018), the typology of digital texts, or the brain's changing way of learning under technological impacts, reading competence must also be affected to a greater or lesser extent (Amiama-Spaillat and Mayor-Ruiz, 2017).

On the other hand, it has been demonstrated that emotions emerge during the reading of, for example, fictional literature. Thus, we agree with Vega et al. (1996, p. 304) that “readers are able to represent accurately the implicit emotional states of protagonists as a consequence of understanding their actions, goals, and relationships with other characters” (Gernsbacher et al., 1992; Gernsbacher and Robertson, 1992; Vega et al., 1996). Other works propose that texts of children's and youth literature be directed to the understanding, regulation, and management of emotions (Vicente-Yagüe Jara, 2018). Sometimes, reading also makes it possible to work on the emotions stemming from grief with students in early and primary education courses (Vicente-Yagüe Jara, 2018). Although when dealing with learning-related reading, emotions also appear, for example, “when the text becomes too difficult or easy for the reader and when conceptual obstacles create cognitive disequilibrium” (Graesser and D'Mello, 2012). It has even been shown that poor literacy is closely associated with states of anxiety (Stevenson et al., 2018).

There is also an inverse path whereby emotions influence reading: “when emotions are evoked during reading, they are likely to have an immediate influence on the reading process” (Mar et al., 2011). This is perhaps in part because “language as an expression of different personalities” (Vilches, 2004) converges in written texts through reading comprehension, and both self-concept and self-esteem also influence the degrees of reading comprehension (Asadi, 2014). In addition, there is evidence of a positive relationship between mood and comprehension performance (Mengual, 2017). To frame the concept, we agree that emotional intelligence or affective competence is defined as “the ability to recognize, express and channel emotional life. Personal balance, self-esteem and empathy acquire special importance. The meta-affectivity or ability of the subject to know and govern feelings is also important” (Martínez-Otero-Pérez and García-Lago, 2013).

Therefore, the proper management of emotions influences academic performance as well as influencing group performance (Suverbiola-Ovejas, 2011) with the result that this is better as a higher emotional intelligence score is obtained (Extremera and Fernández, 2004; Páez and Castaño, 2015; Pishghadam et al., 2022) according to the IE-stress relationship pointed out by Bermejo-Casado et al. (2017), who highlight the key role of reading competence as one of the main academic skills (Bermejo-Casado et al., 2017; Pishghadam et al., 2022). It is a skill which even influences cognitive aspects such as memory (Barreyro et al., 2009; Schammbach, 2016; Vernucci, 2017) since, not in vain, the term “intelligence” comes from Latin *intelligere*, and it is composed of *inter* + *legere* (reading) (Nouwens et al., 2018). In this way, the relationship between EI and reading is verified in numerous investigations (Freitas, 2012; Valdez, 2017), thereby demonstrating the direct relationship between a greater love of reading and a greater reading habit (De Nóbrega and Franco, 2016) and higher EI scores. Nevertheless, despite these investigations that defend the relationship between both constructs (Jiménez-Pérez et al., 2019; Akbari and Pishghadam, 2022), others have concluded that “although reading-comprehension proficiency was relatively associated with several dimensions of EQ (intrapersonal, interpersonal, and stress

management), the total EQ and its subscales were found to be poor predictors of reading comprehension” (Jiménez-Pérez et al., 2019). Further research is therefore necessary.

A better knowledge of the relationship between both constructs can facilitate changes in countries' educational regulations aimed at having a significant decrease in school failure rates, as it is already known that the development of personalized strategies according to student preferences improves reading comprehension (Akbari and Pishghadam, 2022). To contribute to this knowledge, and within this scheme of intervention strategies that can be put into practice to improve academic performance in the classroom, (Fernández Millán et al., 2021; Jiménez-Pérez, 2022) this study aims to analyze the effect of a motivational strategy for reading with a classroom program of intervention conducted by teachers and based on the constructs of reading competence and emotional intelligence. For this, two groups of students were selected to be followed from the beginning to the end of their high school studies (two academic years). In one group, students received specific instruction on reading from their teachers, and the other group followed the regular curricular program established in the official regulations.

The hypothesis to be tested is as follows: students who receive specific training in reading motivation will demonstrate better performance in both reading competence and have an improvement in their emotional intelligence scores.

Method

Participants

The study involved 389 students aged between 14 and 16 years of age, of which 45% were boys and 55% girls, belonging to 12 classes from six public high schools located in middle-class areas in Andalusia (Spain). The students were divided into centers at which they either did not receive any type of intervention (the control group or alpha group), and those at which the promotion of reading was carried out (omega group). In the omega group, teachers controlled the number of readings and type of reading after a brief approach through various media (viewing of films, getting to know the authors, and anecdotes about the literary movement of each work).

The alpha group consisted of 193 (86 male and 107 female, 44.6 and 55.4% respectively) and the omega group consisted of 196 (89 male and 107 female, 45.4 and 54.6% respectively).

Instruments

Two evaluation instruments were selected, both scientifically guaranteed for their reliability and validity from a psychometric point of view: the Wong and law emotional intelligence scale (WLEIS) scale (Pishghadam and Shayesteh, 2017) to measure emotional intelligence, and the CompLEC reading competence test (Ghabanchi and Rastegar, 2014).

Wong and law emotional intelligence scale

This scale provides a measure of basic self-reporting of emotional intelligence, and was initially developed for the organizational field. The recommended age to perform this test is from 16 years onwards, so it is

valid for the target group chosen for this study. The version of the WLEIS in Spanish of Vila and Pérez-González (2007) validated by Merino et al. (2016) has been used (Wong and Law, 2002; Balci, 2017), although recently a new Spanish version WLEIS-S (Wong and Law, 2002) has become made available. It is composed of 16 items measured using a 7-point Likert-type scale from strongly agree to strongly disagree where, in general, a higher score corresponds to greater emotional intelligence. These items are distributed according to four dimensions (with four items per dimension): (1) self-emotion appraisal; (2) others' emotion appraisal; (3) uses of emotion, and (4) regulation of emotion. This allows for adequate internal consistency values for each subscale to be obtained.

The CompLEC reading competence test

This test consists of five texts, three continuous and two discontinuous ones, and a total of 20 questions prepared following the PISA 2000 (OECD, 2016) parameters Llorens et al. (2011). The extent of the continuous texts varies from 274 to 426 words, and the written texts are mainly expository and argumentative. The discontinuous texts contain paragraphs, graphs and diagrams, with a smaller scope of 130 words.

The 20 questions are divided, following PISA guidelines, into three categories that represent the three basic aspects of reading competence that PISA evaluates: (a) 5 information retrieval questions; (b) 10 integration questions; (c) and 5 questions that reflect on the content and form of the text. The response format for 17 items is multiple choice with four alternatives, and three are open questions that require a brief response from the student. The psychometric properties of this test have been adequately contrasted in terms of reliability, homogeneity, difficulty and validity.

Although this test has been designed for adolescents between 11 and 14 years of age, the texts and questions used are perfectly valid for higher age ranges since the type of texts and questions suggested by PISA have been used. The CompLEC test can be obtained at the Ministry of Education of the Government of Spain's webpage, www.lee.es.

Research design and procedure

In this research, a post-descriptive quasi-experimental and longitudinal design was used. The control group followed, without alteration, the current official curriculum of the Junta de Andalucía government established for high school studies. Four start and end measurements (pre-test—post-test) framed in the two intervention periods (one in each course) were carried out in the group (omega) in which the experiment was conducted. Thus, each student completed both the IE test and the reading test at the beginning and at the end of the first year and the second year of high school, with the same tests being used in all cases and for both groups, alpha and omega. *The study, which involved minor human participants, was reviewed and approved by the Institutional Review Board of Asociación Española de Comprensión Lectora the Spanish association of reading comprehension. Written informed consent to participate in this study was provided by the parents or legal guardian of the minor participants.*

To ensure homogeneity in the collection of data, in both groups the individual follow-up for each student was conducted anonymously during class time, in the usual classroom with the collaboration of each

center's language and literature teachers, who had received the same instructions with respect to the intervention group. Likewise, the ethical premises required for research with human beings were respected with regard to informed consent, the guaranteed right to information and protection of personal data, confidentiality and freedom from discrimination for any reason, and the freedom to withdraw at any time.

A descriptive analysis (means and standard deviations) was proposed in each construct (IE and RC), for both groups and sexes, as well as a mixed ANOVA for repeated measures, in which the initial-final evaluation factors were included (P1–P4) according to group and sex typology.

Intervention

In the alpha group, the official regular teaching-learning process was not altered. The intervention project was launched in the omega group. For the students in the omega group, 30 books were chosen at random as a reading proposal from lists of the best books in the world in Spanish from a Google random query. The lists were from Wikipedia, Quelibroleo, culturacolectiva, 20 min, etc. and excluded those that could present explicit adult content (violence and sex).

The selected books (all in Spanish) were: Exupéry's El Principito (The Little Prince), Boyne's El Niño del Pijama de Rayas (The Boy in the Striped Pajamas), Ende's La Historia Interminable [The Endless Story], Bécquer's Rimas (Rhymes), Wilde's El Retrato de Dorian Gray (The Picture of Dorian Gray), Poe's El Escarabajo de Oro (The Gold-Bug), Galdós' Tormento (Torment), Asimov's Amigos Robots (Robbie), Mihur's Tres Sombreros de Copa (Three Top-Hats), Reverte's La Carta Esférica (The Spherical Letter), Ovid's La Metamorfosis (The Metamorphoses), Lorca's La Zapatera Prodigiosa (The Prodigious Shoemaker), Jiménez's Platero y Yo (Platero and I), Orwell's Rebelión en la Granja (Animal Farm: A Fairy Story), Delibes' El Príncipe Destronado (The Deposed Prince), Bazán's Insolación (Insolation), Borges' El Aleph (The Aleph and Other Stories), Bécquer's Leyendas (Legends), Shakespeare's Mucho Ruido y Pocas Nueces (Much Ado About Nothing), Ábalos' Grimpow, Hurley's Ghostgirl, Rowling's Harry Potter 7 (Harry Potter and the Deathly Hallows), Tolkien's El Hobbit (The Hobbit), Zafón's Marina, Mendoza's El Misterio de la Cripta Embrjada (The Mystery of the Haunted Crypt), Galeano's Mitos (Myths), Fences' Soldados de Salamina (Soldiers of Salamina), Sierra i Fabra's Las Chicas de Alambre (The Wire Girls), Moccia's A Tres Metros Sobre el Cielo (Three Meters Above the Sky), and Espinosa's El Mundo Amarillo (The Yellow World)].

The intervention was carried out in 30 sessions of 60 min each, giving a total of 30 readings (five per quarter, fifteen per course) since it has been proven that planning the actions improves reading comprehension (Vila and Pérez-González, 2007). The motivation to read was implemented through tabs with data from the films (when there was a film version of the book), curious facts about the authors and anecdotes about the literary movements to which they belong (for example, that Galdós and Bazán maintained an idyll out of wedlock, that Ende abhorred the film that was made of his book and denounced the production company, or that the causes of Poe's premature death are yet to be resolved).

As an example, we summarize an intervention with "Leyendas," a book by the Spanish writer Gustavo Adolfo Bécquer. At the beginning of the session, for 5 min, several anecdotes are told, the most interesting

being that the author himself predicted his posthumous fame with the phrase “I have a feeling that I will be more and better known dead than alive” since it is the best way to introduce a romantic reading about superstitions, souls and deaths.

Also that he was orphaned and found his love for reading in his godmother’s library, that he suffered from sexually transmitted diseases, that he died very young or that his writings were mysteriously lost. The teacher begins reading aloud for another 5 min, to focus attention on the reading. The remaining 50 min pass in silence because we have already verified that if they read aloud they are more concerned with appearing to read better than they actually do. At the end of the session, the teacher asks several random students to find out how far they have progressed (13 pages on average, 3.5 min/page). The readings are continued at home at a rate of 18 days maximum to conclude it. The trailer Bécquer and the witches of Decine21 is uploaded to the teaching platform.¹

Results

Tables 1, 2 show the mean and standard deviation of the variables emotional intelligence (EI) and reading competence (RC) in the four evaluation phases corresponding to the control group and the intervention group, as well as the differences between the sexes of the students participating in each phase.

Reading competence

The results for the participants in the intervention program show improvement over the course of the 2 years and along the four evaluations. The mixed ANOVA indicates a significant effect [$F_{(1,41)} = 29.11, p < 0.001$], as well as a positive intragroup impact for the evaluation factor [$F_{(1,41)} = 35.58, p < 0.001$].

Regarding the *post hoc* tests, in the omega group (intervention) there were significant differences between the ranges of P1 to P2 ($p < 0.001$) and of P3 to P4 ($p < 0.001$). However, in the alpha group (control) the differences were significant only in the range P1 to P2 ($p < 0.001$), while they were not significant in the range P3 to P4 ($p > 0.001$). In contrast, the total scores (P1 to P4) differed positively in both groups (alpha and omega), and were significant in the two groups ($p < 0.001$). Similarly, significant differences were found between the intervention (omega) and control (alpha) groups where the main effect of the group factor was found [$F_{(1,63)} = 15.69, p < 0.01$]; between both groups there is a difference of two points in favor of the omega intervention group. The effects of intervention and group in favor of the omega group were also significant, which together with the *ad hoc* tests confirm these significant differences. However, the size of the effect is moderate, so the magnitude of the differences is not highlighted by its amplitude ($d = 0.58$). Finally, the differences were not significant in the sex factor, neither with respect to the group ($p = 0.285$) nor the comparative evaluation/group/sex.

In general, reading competence improved in both groups (alpha: 12.1–13.6, omega: 11.1–15.6), with the biggest difference being in the omega group with a total of four points out of 20 against a point and a

TABLE 1 Descriptive statistics.

		Control group (alpha)		Intervention group (omega)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
EI	P1	4.12	0.22	4.05	0.21
	P2	4.31	0.18	4.36	0.40
	P3	4.32	0.25	4.79	0.30
	P4	4.54	0.24	5.13	0.16
RC	P1	12.1	0.28	11.1	0.29
	P2	13.7	0.65	12.3	0.62
	P3	13.5	0.18	13.8	0.51
	P4	13.6	0.22	15.6	0.45

TABLE 2 Mean per group and sex.

		Control group (alpha)		Intervention group (omega)	
		Female	Male	Female	Male
EI	P1	4.18	4.05	4.12	4.19
	P2	4.32	4.30	4.33	4.39
	P3	4.39	4.25	4.81	4.75
	P4	4.58	4.49	5.14	5.12
RC	P1	12.2	12.0	11.5	10.6
	P2	13.7	11.7	12.7	11.7
	P3	12.3	12.7	13.9	12.7
	P4	13.6	13.6	15.7	15.5

half from alpha, and starting from an initial level of one point less in the omega group.

Emotional intelligence

Participants in the intervention group (omega) also improved their emotional intelligence results over the 2 years and the four evaluations. The ANOVA analysis shows evidence of major effects in the intervention [$F_{(1,41)} = 61.42, p < 0.001$]. A positive effect can also be observed in intragroup analysis with evidence of significant differences in the omega group from the P1 to P2 ranges ($p < 0.001$) and from P3 to P4 ($p < 0.001$). However, in the control group (alpha), the differences were not significant between the ranges P1–P2 and P3–P4, although they were in the totals P1–P4. In both cases there was a significant overall improvement ($p < 0.001$) which, in the case of the omega group, reached an improvement of 1.08 total of a maximum score of seven.

Finally, as with reading competence, emotional intelligence did not show significant differences in the sex factor or in combination with the other comparatives (group and intervention).

Discussion

In this research the objective was to analyze how an intervention project to promote reading can influence an improvement in students’

¹ <https://www.youtube.com/watch?v=0Kxajyp4v60>

reading competence and emotional intelligence since a review of the literature points to a direct relationship between both constructs.

The results obtained show that the intervention group (omega) scored significantly above the control group both in reading competence as well as in emotional intelligence, offering evidence of compliance with the hypothesis proposed in this research. These data highlight the fact that specific instruction in the field of reading through the three chosen aspects (parallelism with the cinema, anecdotes about the author, and curiosities about the literary movement to which the work belongs) results in an improvement in the score of both the emotional intelligence and reading competence of the students in an intervention program developed over 2 years. The adoption of this hypothesis has special repercussions in combination with the theses on group work and teaching exposed by Vidal Raméntol and Fuertes Camacho (2013) relative to the teacher taking the emotions of a student into consideration and to the horizontal communication in the classroom (Merino et al., 2016).

In the evaluated phases it was observed that, although both groups increased their scores in general, the intervention group (omega) improved more than the control group (alpha) in both reading competence and emotional intelligence. This indicates that although the curriculum provided by the Junta de Andalucía fulfills its functions, the results can be significantly improved by planning a basic program to promote reading as proposed in this study. With regard to sex, the differences were positively inclined toward the female sex, which scored significantly better in both reading and emotional intelligence in the last phase of evaluation, as evidenced by the improvement not only in the intervention group (omega) but also in the control (alpha).

It is necessary to highlight that among the limitations of research designs with repeated measures there exists the possible effect of learning by practice since students have been able to improve their responses with repetition, and data obtained in later stages may seem better without really being so. In this case, however, it can be considered that the potential effect of this bias was limited both by the age of the participants and by the time difference over which the measurement instruments were administered. In addition, after each administration there was no feedback to students, so it is unlikely that they could “learn” which options or measures were more correct than others and try to improve their scores. Therefore, it is reasonable to consider that the improvements obtained in both reading competence and emotional intelligence are due more to the reading stimulus intervention program than to a bias derived from the experience effect with the measuring instruments.

This research provides a basis from which to propose a generalized program of intervention in reading promotion in Spanish educational centers in order to improve two of the most studied skills in recent years in the field of education: reading competence and emotional intelligence. On the one hand, a program could seek to stimulate students’ own resources to optimize their level of reading competence (Vidal Raméntol and Fuertes Camacho, 2013; Extremera et al., 2019) within the existing seven levels: literal (objective and subjective), representative, inferential, critical, emotional, creative and meta-cognitive (Jiménez-Pérez, 2015). And on the other hand, it could seek to perfect personal skills in emotional intelligence (Alfonso et al., 2017): the conception of one’s own emotions, the ability to control them, the ability to self-motivate (tremendously useful in the field of teaching), the recognition of others’ emotions and the control of relationships.

It would be interesting to extend this research to students from other Spanish-speaking countries, such as those in Latin America, in order to evaluate reading competence comparatively across the broad

spectrum that the Spanish language and its cultural nuances present in the world. In addition, it would be interesting to carry out a comparative analysis of the curricular programs in the different countries where Spanish is the official language according to the parameters analyzed in this study. It is also an interesting line of research to check the opposite effect, that is, to test whether implementing a program for improvement in emotional intelligence in both primary and secondary school classrooms can improve academic performance in reading competence so that in the long run it would not be necessary to train such competence at higher educational levels, such as at university.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The study, which involved minor human participants, was reviewed and approved by the Institutional Review Board of Asociación Española de Comprensión Lectora, the Spanish association of reading comprehension. Written informed consent to participate in this study was provided by the parents or legal guardian of the minor participants.

Author contributions

EJ-P came up with the idea and focused on the theoretical framework. ML focused on the methodology. MV-YJ focused on the discussion and conclusions. PG focused on the writing and the translation into English. All authors contributed to the article and approved the submitted version.

Acknowledgments

This research has been carried out thanks to the sponsorship of the Spanish Association of Reading Comprehension (Asociación Española de Comprensión Lectora), the UMA Hum-1048 research team, and the research stays that the authors have carried out.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Akbari, M., and Pishghadam, R. (2022). Developing new software to analyze the emotional load of language. *BCT* 1, 1–13. doi: 10.56632/bct.2022.1101
- Alfonso, S., Deaño, M., Ramos, A., Conde, A., Gayo, M. E., Iglesias, V., et al. (2017). Prevención e intervención de las dificultades de aprendizaje en lectura. *Rev. INFAD Psicol.* 3, 71–82. doi: 10.17060/ijodaep.2017.n1.v3.976
- Alonso-Mosquera, M. H., González-Vallés, J. E., and Muñoz-de-Luna, A. B. (2016). Ventajas e inconvenientes del uso de dispositivos electrónicos en el aula: percepción de los estudiantes de grados en comunicación/Advantages and disadvantages of using electronic devices in the classroom. *SEECI* 41, 136–154. doi: 10.15198/seeci.2016.41.136-154
- Amiama-Spaillat, C., and Mayor-Ruiz, C. (2017). Digital reading and reading competence: the influence in the Z generation from the Dominican Republic. *Comunicar* 25, 105–114. doi: 10.3916/C52-2017-10
- Asadi, N. (2014). The relationship between self-concept, self-efficacy, self-esteem and reading comprehension achievement: evidence from Iranian EFL learners. *Int. J. Soc. Sci. Educ.* 5, 58–66.
- Balci, O. (2017). The effects of learning-style based activities on students' reading comprehension skills and self-efficacy perceptions in English foreign language classes. *High. Educ. Stud.* 7, 35–54. doi: 10.5539/hes.v7n4p35
- Barreiro, J., Burin, D., and Aníbal, D. (2009). Capacidad de la memoria de trabajo verbal. Validez y fiabilidad de una tarea de amplitud de lectura. *Interdisciplinaria* 26, 207–228.
- Bavishi, A., Slade, M. D., and Levy, R. B. (2016). A chapter a day: association of book reading with longevity. *Soc. Sci. Med.* 164, 44–48. doi: 10.1016/j.socscimed.2016.07.014
- Bermejo-Casado, L., Campos-García-de-Quevedo, G., and Sánchez-Bayón, A. (2017). Estrés y síndrome de desgaste profesional en la organización de eventos en el sector de las agencias de comunicación. *Rev. Lat. Comun. Soc.* 72, 1394–1417. doi: 10.4185/RLCS-2017-1225.Spanish
- Cordón-García, J. A. (2018). Combates por el libro: inconclusa dialéctica del modelo digital. *EPI* 27, 467–481. doi: 10.3145/epi.2018.may.02
- De Nóbrega, M., and Franco, G. (2016). Inteligencia Emocional y actividad lectora en una escuela de 1.º ciclo. *Rev. INFAD Psicol.* 5, 159–165. doi: 10.17060/ijodaep.2014.n1.v5.658
- Echeverría, MC. (1998). Expresión Verbal en la Comunicación Mediática. Revista Latina de Comunicación Social. Available at: www.revistalatinacs.org/a/95mirta.htm
- Extremera, N., and Fernández, P. (2004). El papel de la inteligencia emocional en el alumnado: evidencias empíricas. *RIEOEI* 34, 1–9. doi: 10.35362/rie3334005
- Extremera, N., Rey, L., and Sánchez-Álvarez, N. (2019). Validation of the Spanish version of the Wong Law emotional intelligence scale (WLEIS-S). *Psicothema* 31, 94–100.
- Fernández Millán, G., García Guirao, P., and López Martínez, O. (2021). El pensamiento crítico aplicado a "El Lazarillo de Tormes" a través del debate en 3º de la ESO. *Investig. Sobre Lect.* 16, 32–50. doi: 10.24310/isl.vi16.12870
- Freitas, L. (2012). A inteligência emocional e a leitura: um estudo com crianças do 1.º ciclo do ensino básico. Bachelor's Degree Final Project, Universidade da Madeira.
- Gernsbacher, M. A., Goldsmith, H. H., and Robertson, R. R. W. (1992). Do readers mentally represent characters' emotional states? *Cogn. Emot.* 6, 89–111. doi: 10.1080/02699939208411061
- Gernsbacher, M. A., and Robertson, R. R. W. (1992). Knowledge activation versus sentence mapping when representing fictional characters' emotional states. *Lang. Cogn. Process.* 7, 353–371. doi: 10.1080/01690969208409391
- Ghabanchi, Z., and Rastegar, R. (2014). The correlation of IQ and emotional intelligence with reading comprehension. *Read. Matrix* 14, 135–144.
- Graesser, A. C., and D'Mello, S. (2012). Moment-to-moment emotions during reading. *Read. Teach.* 66, 238–242. doi: 10.1002/TRTR.01121
- Gurning, B. (2017). The effect of teaching strategies and curiosity on students' achievement in Reading comprehension. *ELT* 10, 191–198. doi: 10.5539/elt.v10n11p191
- Jiménez-Pérez, E. (2014). *Comprensión Lectora vs Competencia Lectora: Qué Son y Qué Relación Existe Entre Ellas. Investigaciones Sobre Lectura.* 63–74. Málaga: Asociación Española de Comprensión Lectora.
- Jiménez-Pérez, E. (2015). Niveles de comprensión lectora. *Leng. Textos* 41, 19–25.
- Jiménez-Pérez, E. (2022). Critical competence and critical thinking. *Investig. Sobre Lect.* 17, 1–13.
- Jiménez-Pérez, E., Alarcón, R., and Vicente-Yagüe, M. I. (2019). Reading intervention: correlation between emotional intelligence and reading competence in high school students. *Rev. Psicodidact.* 24, 24–30. doi: 10.1016/j.psicoe.2018.10.001
- Llorens, A. C., Gil, L., Vidal-Abarca, E., Martínez, T., Mañá, A., Gilabert, R., et al. (2011). Prueba de competencia lectora para educación secundaria (CompLEC). *Psicothema* 23, 808–817.
- Mar, R., Oatley, K., Djikic, M., and Mullin, J. (2011). Emotion and narrative fiction: interactive influences before, during, and after reading. *Cogn. Emot.* 25, 818–833. doi: 10.1080/02699931.2010.515151
- Martínez-Cerdá, J. F., and Torrent-Sellens, J. (2017). Empoderamiento mediático mediante e-learning. Diseño y validación de una escala. *EPI* 26, 43–54. doi: 10.3145/epi.2017.ene.05
- Martínez-Otero-Pérez, V., and García-Lago, V. (2013). Aspectos conceptuales sobre la inteligencia afectiva y su fomento a través de la lectura en la escuela. *Rev. Comun. SEECI* 30, 106–120. doi: 10.15198/seeci.2013.30.106-120
- Mengual, E. Z. (2017). Metacomprensión e inteligencia emocional: relación e influencia en la comprensión lectora en alumnado de 5º y 6º de educación primaria. Doctoral dissertation. Madrid: Universidad Complutense de Madrid.
- Merino, C., Lunahuaná, M., and Pradhan, R. K. (2016). Validación estructural del Wong-Law Emotional Intelligence Scale (WLEIS): estudio preliminar en adultos. *Liberabit* 22, 103–110. Spanish. doi: 10.24265/liberabit.2016.v22n1.09
- Nouwens, S., Groen, M. A., Kleemans, T., and Verhoeven, L. (2018). The role of semantic retrieval in children's reading comprehension development in the upper primary grades. *J. Res. Read.* 41, 597–614. doi: 10.1111/1467-9817.12128
- OECD (2016). *PISA 2015 Results (Volume 1): Excellence and Equity in Education*, PISA. Paris: OECD Publishing.
- Páez, M. L., and Castaño, J. J. (2015). Inteligencia emocional y rendimiento académico en estudiantes universitarios. *Psicol. Caribe* 32, 268–285. doi: 10.14482/psdc.32.2.5798
- Pishghadam, R., Faribi, M., Kolahi Ahari, M., Shadloo, F., Gholami, M. J., and Shayesteh, S. (2022). Intelligence, emotional intelligence, and emo-sensory intelligence: which one is a better predictor of university students' academic success? *Front. Psychol.* 13, 1–11. doi: 10.3389/fpsyg.2022.995988
- Pishghadam, R., and Shayesteh, S. (2017). Emo-sensory expression at the crossroads of emotion, sense, and language: a case of color-emotion associations. *Int. J. Soc. Cult. Lang.* 5, 15–25.
- Ripoll Salceda, J. C. (2014). ¿Existen métodos de mejora de la comprensión lectora en español y basados en evidencias? *Investigaciones Sobre Lectura* 2, 44–52. doi: 10.24310/revistaisl.vi2.10962
- Rodríguez, E. (2016). Leer nos Hace más Felices. El País. Available at: https://elpais.com/elpais/2016/01/22/eps/1453483676_726569.html (Accessed June 3, 2019).
- Sánchez-Claros, J. P. (2016). Nuevas alfabetizaciones en un entorno multimodal: nuevas necesidades lectoras para un entorno textual múltiple. *ISL* 6, 51–57. doi: 10.37132/isl.v0i6.153
- Schamallbach, A. (2016). Influencia de la comprensión verbal y la memoria de trabajo en el rendimiento escolar en la asignatura de lengua castellana en infantil. M.Sc. Thesis, UNIR.
- Stevenson, J., Pimperton, H., Kreppner, J., Worsfold, S., Terleksi, E., Mahon, M., et al. (2018). Language and reading comprehension in middle childhood predicts emotional and behaviour difficulties in adolescence for those with permanent childhood hearing loss. *J. Child Psychol. Psychiatry* 59, 180–190. doi: 10.1111/jcpp.12803
- Sullivan, A., and Brown, M. (2015). *Reading for Pleasure*. 1st ESRC: London.
- Suvertola-Ovejas, I. (2011). Competencia emocional y rendimiento académico en el alumnado universitario. VA 117-E, 1–17. doi: 10.15178/va.2011.117E.1-17
- Taylor, M. (2016). The impact of reading on career prospects. *Books Keeps* 189, 6–9.
- Valdez, W. A. (2017). La inteligencia emocional y el desarrollo de habilidades en los estudiantes del quinto año de formación profesional en la facultad de ciencias de la educación de la Universidad Nacional de San Agustín Arequipa. Doctoral dissertation. Arequipa: Universidad Nacional de San Agustín de Arequipa de Perú.
- Vega, M., León, I., and Díaz, J. M. (1996). The representation of changing emotions in reading comprehension. *Cogn. Emot.* 10, 303–322. doi: 10.1080/026999396380268
- Vernucci, S., Canet-Juric, L., Andrés, M. L., Burin, D. I. (2017). Comprensión lectora y cálculo matemático: El rol de la memoria de trabajo en niños de edad escolar. *Psykhe* 26, 1–13. doi: 10.7764/psykhe.26.2.1047
- Vicente-Yagüe Jara, M. I. (2018). "Función de la literatura infantil y juvenil para la regulación de las emociones" in *Investigación e Innovación en Educación Literaria*. eds. M. I. Vicente-Yagüe Jara and E. Jiménez Pérez (Madrid: Síntesis), 15–29.
- Vicente-Yagüe Jara, M. I. (2018). "Elaboración y análisis de un repertorio literario sobre muerte y duelo para su uso didáctico en Educación Infantil y Educación Primaria" in *Contenidos Docentes de Vanguardia*. eds. R. Chao, A. Amaro and M. Rodorigo (Barcelona: Gedisa), 131–146.
- Vidal Raméntol, S., and Fuertes Camacho, M. T. (2013). La dinámica de grupos para el trabajo cooperativo facilita la comunicación. *Vivat Acad. Rev. Comun.* 123, 1–12. doi: 10.15178/va.2013.123.1-12
- Vila, E., and Pérez-González, J. C. (2007). Madurez para la carrera e inteligencia emocional en alumnado de secundaria: una aproximación correlacional. In F. Etxebarria, I. Sarasola, J. F. Lukas, J. Etxebarria and A. Martxueta, Eds. Proceedings of the 13rd Congreso Nacional de Modelos de Investigación Educativa; 2007 June 27–29; Universidad de Donostia-San Sebastián. Donostia: Erein/AIDPE. 712–717.
- Vilches, I. (2004). Inteligencia emocional en la asignatura de lenguaje y comunicación. *Contextos* 12, 77–86.
- Wong, C. S., and Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: an exploratory study. *Leadersh. Q.* 13, 243–274. doi: 10.1016/S1048-9843(02)00099-1



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga,
Spain

REVIEWED BY

Macarena Becerra,
University of Malaga,
Spain
Weiwei Zhang,
Quzhou University,
China
Fareeha Javed,
Government of Punjab,
Pakistan
Sara Kashefian-Naeeni,
Shiraz University of Medical Sciences,
Iran
Andrew Philominraj,
Universidad Católica del Maule,
Chile

*CORRESPONDENCE

Xuecong Miao
✉ xuecong_phd@163.com
Yang Dong
✉ yangdong3-c@my.cityu.edu.hk

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 15 November 2022

ACCEPTED 16 February 2023

PUBLISHED 13 March 2023

CITATION

Zheng H, Miao X, Dong Y and Yuan D-C (2023)
The relationship between grammatical
knowledge and reading comprehension: A
meta-analysis.
Front. Psychol. 14:1098568.
doi: 10.3389/fpsyg.2023.1098568

COPYRIGHT

© 2023 Zheng, Miao, Dong and Yuan. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

The relationship between grammatical knowledge and reading comprehension: A meta-analysis

Haoyuan Zheng¹, Xuecong Miao^{2,3*}, Yang Dong^{4*} and Daniel-Chongbo Yuan⁵

¹Faculty of Teacher and Education, Guangzhou Huashang College (廣州華商學院), Guangzhou, China, ²Research Center for Overseas Studies and Media Reports on Hainan, Hainan University, Haikou, China, ³Faculty of Sociology, Huazhong University of Science and Technology, Wuhan, China, ⁴Department of English, Hainan University, Haikou, China, ⁵Department of Literacy, Hong Kong Metropolitan University, Kowloon, Hong Kong SAR, China

The study aimed to examine the cohesive tie effect on reading comprehension through the grammatical knowledge cognition process. The present meta-analysis examined the correlation between grammatical knowledge and reading comprehension based on empirical results published between 1998 and 2021. This study selected 86 studies with a total of 14,852 readers whose grades were grouped from primary school to university. The results showed that the overall correlation effect size between grammatical knowledge and reading comprehension was large, and the significant interaction effect of the grade group was confirmed through moderator analysis. The results suggested that the grammatical knowledge's function of the cohesive tie has a transfer effect across different text comprehension scripts.

KEYWORDS

grammatical knowledge, reading comprehension, reading stage, learning cognitive condition, mental model, meta-analysis

Introduction

Reading comprehension refers to an ability to acquire literal or inferential meaning from the given text through the interaction between sentences' cognition process and situation image construction (Snow, 2002; Cain et al., 2004; Dong et al., 2021, 2022). Regardless of whether it is an Eastern or a Western country, reading comprehension plays a vital role in human career development, knowledge base construction, and language ability development and even in printed script-based communication (Carretti et al., 2009; García and Cain, 2014; del Pilar Jiménez et al., 2019). As one kind of foundational knowledge and linguistic skill, grammatical knowledge could influence reading fluency and reading accuracy and further influence reading comprehension through discourse comprehension (e.g., Alderson, 2000, p. 37; Muter et al., 2004; Chik et al., 2012). *Cohesive tie theory* (Cain and Nash, 2011) suggests that grammatical knowledge generates a necessary inference for discourse comprehension through information integration. *Reading stages theory* (Chall, 1996) suggests that grammatical knowledge should have different effects on reading comprehension across different grade groups. Hjetland et al. (2019) and Zhang (2012) reported that the grade of primary and secondary school readers positively moderated the correlation between GK and reading comprehension. Moreover, *cognitive condition theory* (Tagarelli et al., 2011) suggested that the function of grammatical

knowledge on the text comprehension process ought to be different between the first-language (L1) and second-language (L2) scripts due to literacy exposure. Zhang and Koda (2018) and Shibasaki et al. (2015) used the structural equation model to report that the GK contributed more to unfamiliar language scripts than to familiar scripts in text comprehension in school-age students. However, the interaction effects between grammatical knowledge and reading comprehension across developed relations (grade group) and literacy exposure (language type) remain unclear. Moreover, extensive studies reported the inconsistent correlations between grammatical knowledge and reading comprehension, and whether the variance could be explained by the interaction effect of reading stages or learning cognitive conditions requires further investigation. Therefore, the current study considered the meta-analytic approach to synthesize empirical studies, investigating the rule of cohesive tie function of grammatical knowledge on text comprehension.

Literature review

Grammatical knowledge and reading comprehension

Grammatical knowledge, in general, refers to knowing the rule of grammar application in reading text construction of verb form, word order, and sentence structure (Shiotsu, 2010; Purpura, 2013). Under the framework of the cohesive tie, past studies have confirmed grammatical knowledge as a prerequisite of efficient reading for reading both L1 and L2 texts (Grabe, 2009; Jeon and Yamashita, 2014). This is because grammatical knowledge determines the sentence structure and further impacts reading comprehension through the word function sequence (e.g., Gahl and Garnsey, 2004; Vinyals et al., 2015), which means that grammatical knowledge may assist individual word comprehension or semantic chunk comprehension when readers use the rule of sentence structure for decoding sentence meanings (Tunmer, 1989; Rego and Bryant, 1993); that is, the better the proficiency in grammatical knowledge application, the better the text reading comprehension performance. Readers could take the grammatical rule to assist text comprehension through the indirect contribution of text semantic meaning inference (Rego and Bryant, 1993).

A consensus has been reached on the fact that fully developed grammatical knowledge could allow readers to monitor syntactic information through text comprehension progress. The relationship between grammatical knowledge and reading comprehension is associated with the reader's ability to generate coherence in the text and monitor meaning acquisition during the reading process (Fender, 2001; Grabe and Stoller, 2012). For example, the grammatical clues offered coherence hints or information to aid readers in constructing text and understanding discourses (Zwaan and Rapp, 2006; Aryadoust and Baghaei, 2016), which indicates that the grammatical knowledge could be regarded as an ability to process syntactic structures and complicated sentences in text comprehension (Zwaan and Rapp, 2006; Aryadoust and Baghaei, 2016).

Previous studies, however, reported various results regarding the role of grammatical knowledge in text reading comprehension, from a strong correlation to a weak correlation. Some studies reported a strong correlation between grammatical knowledge and reading

comprehension (Bentin et al., 1990; Muter et al., 2004; Shiotsu and Weir, 2007; Kim and Cho, 2015). Conversely, other empirical surveys and experimental research reported a weak association due to the development effect of lexical inference (Kern, 1989; Kleeck, 2008; Zhang, 2012), knowledge base (Schatschneider et al., 2004; Silva and Cain, 2015; Brimo et al., 2017), and phonology knowledge (Demont and Gombert, 1996; Blackmore and Pratt, 1997; Farnia and Geva, 2013). Therefore, the correlation between grammatical knowledge and text reading comprehension was unclear.

Potential moderator selection

The current study selected grade group and language type as potential moderators for the following reasons.

Grade group

The Reading stages theory (Chall, 1996) claims that readers start learning to read in early primary school to become professional in reading to learn at university. Higher reading stages match more complicated requirements of text reading comprehension and proficiency in grammatical knowledge application. Previous studies also showed various results on the correlation between reading comprehension and grammatical knowledge. For example, Farnia and Geva (2013) reported that younger readers (grade 1) performed better in English grammatical knowledge tests than older readers (grades 4 and 6). However, it was unknown whether the variance could be explained by the grade group. Based on the reading stage statement (Chall, 1996) and previous reading ability development research (e.g., Mol and Bus, 2011), this study regarded grades 1–3 of primary school as group L; grades 4–6 as group H; and grades 7–12 as grade S. This was informed by García and Cain's (2014) study which showed that readers had an independent reading ability development period in secondary school, and regarded university learning as grade U, respectively.

Language type

Learning cognitive condition statement suggested that the grammatical effect on text comprehension was involvement in noticing and processing the discourse structures (Tagarelli et al., 2011). Due to the familiarity effect on text scripts, more contribution of grammatical knowledge on text comprehension might be found in L2 than in L1 text comprehension tasks (Koda, 1992; Waltzman and Cairns, 2000). The language type might be an explanation of the various empirical correlation findings between reading comprehension and grammatical knowledge.

Past relevant meta-analytic review

In the past two decades, a few studies (e.g., Jeon and Yamashita, 2014) reviewed the contribution of grammatical knowledge to reading comprehension. For example, Rodd et al. (2015) used a meta-analytic approach to report the effects of localizing semantic grammatical knowledge and written comprehension on brain function. Jeon and Yamashita (2014) reported the correlation between grammatical knowledge and general L2 acquisition. However, past meta-analysis studies failed to solve the significant heterogeneity problem through the

moderator analysis, which means that the theoretical contribution to the correlation between reading comprehension and grammatical knowledge requires further investigation. In addition, previous studies only selected a small number of empirical studies for correlation calculation, meaning that the selected studies were not representative. The current picture of the interaction effect with reading stage development and learning cognitive condition statement thus remains unclear.

The current study

This article extends the current literature concerning the influence of the developmental relations of grammatical knowledge on reading comprehension, investigating the overall correlation between grammatical knowledge and reading comprehension from the recent research published in the past 20 years. This study further examines the interaction effects of the reading stage and learning cognitive condition on the correlation between reading comprehension and grammatical knowledge.

Method

The official guideline of the meta-analytic approach The Preferred Reporting Items for Systematic reviews and Meta-Analyses (*PRISMA*) was applied to this study, including the literature base, inclusion criteria, coding process, and meta-analytic procedure. The *PRISMA* is an official meta-analytic procedure guideline on meta-analysis implementation.

Literature base

To avoid the comprehension system mistake (García and Cain, 2014), this study only selected materials written in Chinese and English. The Chinese materials were selected from the CNKI database, which included almost all possible academic resources written in Chinese, and the English materials were searched from Google Scholar, PsycINFO, ERIC, and Pre-Quest. Relevant studies were identified through two groups of keywords; the first group of words was related to grammatical knowledge (grammar*, grammatical*, syntax*, syntactic*, oral cloze test*, sentence completion*, error recognition*, implicit knowledge*, and explicit knowledge*) and the second group of words was related to reading comprehension (sentence comprehension*, paragraph comprehension*, passage comprehension*, text comprehension*, reading comprehension*, reading ability*, comprehension ability*, reading acquisition*, reading performance*, comprehension ability*, and comprehension performance*). This study attempted to select all possible materials with publication dates from 1 January 1998 to 1 December 2022, including published or unpublished research articles, dissertations, and conference articles. Thus, 1,104 articles were selected from the database.

Inclusion criteria

Materials eligible for inclusion in this meta-analysis were as follows: (a) empirical studies with a minimum of 30 participants, (b)

the selected studies should have correlation scores on the correlation between grammatical knowledge and reading comprehension, (c) the participants should be students from grade 1 of primary school to undergraduate students, (d) students were not diagnosed with deaf or blind, (e) studies investigated the concurrent correlation between grammatical knowledge and reading comprehension, (f) studies should provide enough indicators which could be transformed into Fisher's *z*-transformation, and (g) grammatical knowledge and reading comprehension should come from the same scripts (e.g., L1 grammatical knowledge and L1 reading comprehension, L2 grammatical knowledge and L2 reading comprehension). If the grammatical knowledge and reading comprehension came from different scripts (e.g., L1 grammatical knowledge and L2 reading comprehension), the related study was removed. Thus, 128 articles fulfilled the inclusion criteria.

Coding process

Studies were coded according to the characteristics of the participants and the measurements by two independent coders. Coding was based on (a) study number, (b) first author's name, (c) publication year, (d) material type (journal article or dissertation), (e) sampling area, (f) sample size, (g) grade group (L refers to grades 1–3 of primary school, H refers to grades 4–6 of primary school, S refers to secondary school, U refers to university), (h) target language scripts, and (i) correlation effect size. Any unclear information was emailed to the article's author for clarification. Because this study did not investigate the transfer effect in grammatical knowledge and attempt to examine the hypothesis from the learning cognitive condition theory, the two coders were asked to remove those articles in which the grammatical knowledge and reading comprehension did not come from the same scripts, for example, the correlation between L1 grammatical knowledge and L2 reading comprehension. If one article provided more than one correlation that came from different grade groups, this study treated them as an independent study (Mol and Bus, 2011). If one article provided more than one available correlation indicator in the same grade group, first, this study preferred to select the correlation indicator from the standardized measurements than from the research-developed test. After this step, if the article still had more than one available indicator, Hedges et al. (2010) cluster regression was applied to calculate the final Fisher's *z*-transformation for this study, ensuring that each study only provided one correlation indicator for further meta-analysis (Mol and Bus, 2011). The internal reliability for two independent coders was 0.96 in the first round of coding. The difference came from sampling area coding; however, this problem was solved by using the country's name as the sampling area. Thus, 65 articles remained after the coding. Detailed information can be found in Table 1. All materials selected for the meta-analysis have been listed in references with “*.”

Meta-analytic procedure

This study removed two articles conducted by Zhang (2011, 2012), and one study conducted by Brimo et al. (2017) due to the effect size, which was over 3.5 standard deviation (García and Cain, 2014). As a result, 62 articles with 86 studies were included in the final analysis.

TABLE 1 Moderators and outcomes for the matched set of self-report studies in meta-analysis.

Study No	Materials type	First author	Publication time	Sample size	Area	Grade group ^a	Language type	Fisher's z	SE
1	Journal Article	Roth	2002	66	USA	L	L1	0.36	0.13
2	Journal Article	Catts	2002	268	USA	L	L1	0.32	0.06
3	Journal Article	Chamberlain	2008	31	Canada	U	L1	0.79	0.19
4	Journal Article	Gottardo	2009	79	Mexical	L	L2	0.42	0.12
5	Journal Article	Kim	2011	242	USA	L	L1	0.45	0.07
6a	Journal Article	Vellutino	2007	297	USA	L	L1	0.34	0.06
6b	Journal Article	Vellutino	2007	171	USA	H	L1	0.35	0.08
7	Journal Article	Chung	2013	78	HK	S	L1	0.51	0.12
8a	Journal Article	Geva	2012	390	Canada	S	L2	0.66	0.05
8b	Journal Article	Geva	2012	149	Canada	S	L1	0.68	0.08
9	Book Chapter	Rescorla	2000	22	USA	S	L1	0.58	0.23
10	Journal Article	Aryadoust	2016	825	Iran	U	L2	0.65	0.04
11a	Journal Article	Cutting	2009	56	USA	S	L2	0.58	0.14
11b	Journal Article	Cutting	2009	56	USA	S	L2	0.60	0.14
12	Journal Article	Nassaji	1999	60	Canada	U	L2	0.56	0.13
13a	Dissertation	Brimo	2011	193	USA	S	L2	0.42	0.07
13b	Dissertation	Brimo	2011	193	USA	S	L2	0.24	0.07
14	Journal Article	Leider	2013	51	USA	H	L2	0.57	0.14
15	Journal Article	O'Connor	2004	72	USA	S	L1	0.60	0.12
16	Journal Article	Bowey	2005	97	USA	L	L1	0.41	0.10
17a	Journal Article	Farnia	2013	400	Canada	L	L2	0.24	0.05
17b	Journal Article	Farnia	2013	400	Canada	H	L2	0.44	0.05
17c	Journal Article	Farnia	2013	153	Canada	L	L1	0.37	0.08
17d	Journal Article	Farnia	2013	153	Canada	H	L1	0.60	0.08
18	Journal Article	Whyte	2013	26	USA	H	L1	0.54	0.21
19	Journal Article	Lesaux	2006	480	Canada	H	L2	0.50	0.05
20a	Journal Article	Schatschneider	2004	384	USA	L	L1	0.32	0.05
20b	Journal Article	Schatschneider	2004	189	USA	L	L1	0.21	0.07
21	Journal Article	Adlof	2010	433	USA	S	L1	0.62	0.05
22a	Journal Article	Rescorla	2002	34	USA	L	L1	0.34	0.18
22b	Journal Article	Rescorla	2002	34	USA	L	L1	0.34	0.18
22c	Journal Article	Rescorla	2002	34	USA	H	L1	0.34	0.18
22d	Journal Article	Rescorla	2002	34	USA	H	L1	0.60	0.18
23	Journal Article	Silverman	2015	377	USA	H	L1	0.48	0.05
24a	Journal Article	Van Gelderen	2004	397	Netherlands	S	L1	0.62	0.05
24b	Journal Article	Van Gelderen	2004	397	Netherlands	S	L2	0.59	0.05
25	Journal Article	Potocki	2013	131	France	L	L1	0.37	0.09
26	Journal Article	Chaney	1998	41	USA	L	L1	0.34	0.16
27a	Journal Article	Kim	2015	200	Korea	S	L2	0.65	0.07
27b	Journal Article	Kim	2015	200	Korea	S	L2	0.66	0.07
28a	Journal Article	van Gelderen	2003	397	Netherlands	S	L1	0.56	0.05
28b	Journal Article	van Gelderen	2003	397	Netherlands	S	L2	0.56	0.05
29	Journal Article	Chik	2012	274	HK	L	L1	0.54	0.06

(Continued)

TABLE 1 (Continued)

Study No	Materials type	First author	Publication time	Sample size	Area	Grade group ^a	Language type	Fisher's z	SE
30	Journal Article	Oakhill	2003	102	UK	L	L1	0.42	0.10
30	Journal Article	Oakhill	2003	102	UK	S	L1	0.58	0.10
31	Journal Article	Lasagabaster	2001	126	Spain	H	L2	0.58	0.09
32	Journal Article	Jafari	2016	50	Iran	U	L2	0.80	0.15
33a	Journal Article	Oakhill	2012	102	UK	H	L1	0.42	0.10
33b	Journal Article	Oakhill	2012	92	UK	H	L1	0.58	0.11
33c	Journal Article	Oakhill	2012	83	UK	H	L1	0.51	0.11
34	Journal Article	Goff	2005	180	Australia	H	L1	0.65	0.08
35a	Journal Article	Shiotsu	2007	107	UK	U	L1	0.73	0.10
35b	Journal Article	Shiotsu	2007	182	UK	U	L2	0.71	0.08
35c	Journal Article	Shiotsu	2007	591	UK + Japan	U	L2	0.73	0.04
36a	Journal Article	Proctor	2011	294	USA	H	L1	0.45	0.06
36b	Journal Article	Proctor	2011	294	USA	H	L1	0.55	0.06
37	Journal Article	Foorman	2015	218	USA	L	L1	0.44	0.07
38	Journal Article	Park	2012	28	Korea	U	L2	0.81	0.20
39	Journal Article	Zhang	2012	190	China	U	L2	0.22	0.07
40	Journal Article	Mokhtari	2012	32	USA	H	L1	0.52	0.19
41	Journal Article	Gui	2018	181	China	U	L2	0.60	0.08
42	Journal Article	Xiang	2016	168	China	S	L2	0.59	0.08
43a	Journal Article	Gong	2009	68	China	S	L2	0.68	0.12
43b	Journal Article	Gong	2009	68	China	S	L2	0.56	0.12
44	Journal Article	Lu	2015	106	China	S	L2	0.69	0.10
45	Journal Article	Wu	2016	70	Japan	U	L2	0.66	0.12
46	Journal Article	Shen	2014	68	China	U	L2	0.68	0.12
47	Dissertation	Liao	2012	44	China	S	L2	0.50	0.16
48	Dissertation	Zhang	2011	31	China	U	L2	0.84	0.19
49a	Dissertation	Luo	2010	63	China	U	L2	0.71	0.13
49b	Dissertation	Luo	2010	63	China	U	L2	0.72	0.13
50	Dissertation	Tang	2013	188	China	S	L2	0.60	0.07
51a	Dissertation	Deng	2014	35	China	U	L2	0.62	0.18
51b	Dissertation	Deng	2014	35	China	U	L2	0.74	0.18
52	Journal Article	Shen	2011	68	China	U	L2	0.81	0.12
53a	Journal Article	Chen	2008	58	China	L	L1	0.40	0.14
53b	Journal Article	Chen	2008	58	China	L	L1	0.31	0.14
54	Journal Article	Li	2008	50	China	U	L2	0.79	0.15
55	Dissertation	Wang	2012	110	China	S	L2	0.48	0.10
56	Journal Article	Zhang	2011	190	China	U	L2	0.22	0.07
57	Journal Article	Gong	2010	72	China	S	L2	0.66	0.12
58	Dissertation	Yan	2012	91	China	U	L2	0.60	0.11
59	Journal Article	Li	2016	199	China	L	L1	0.34	0.07
60a	Dissertation	Jiang	2003	188	China	S	L2	0.61	0.07
60b	Dissertation	Jiang	2003	188	China	S	L2	0.60	0.07
61	Journal Article	Zhang	2017	77	China	U	L2	0.83	0.12

(Continued)

TABLE 1 (Continued)

Study No	Materials type	First author	Publication time	Sample size	Area	Grade group ^a	Language type	Fisher's z	SE
62	Journal Article	Guan	2007	188	China	S	L2	0.60	0.07
63	Journal Article	Chen	2005	41	China	U	L2	0.57	0.16
64a	Dissertation	Wang	2001	176	China	S	L2	0.56	0.08
64b	Dissertation	Wang	2001	176	China	S	L2	0.60	0.08

^aL = lower grade, from kindergarten to grade 2, H = higher grade, from grade 3 to grade 6, S = secondary school, grade 7 to grade 12. U = undergraduate students and master students.

The researchers input the correlation indicator through comprehension meta-analysis 3.0 and transformed it into Fisher's z for further analysis. This study selected Fisher's z -transformation because z is approximately constant, and z has asymmetrical distribution (Borenstein et al., 2010). The values of Fisher's z were 0.10 ($r=0.10$), 0.31 ($r=0.30$), and 0.50 ($r=0.50$), denoting small effect size, moderate effect size, and large effect size, respectively (Borenstein et al., 2010).

This study used the random-effects model to report the effect size (Borenstein et al., 2010) and also reported a 95% confidence interval (CI). The effect size was interpreted as significant if the CI did not include zero. Next, the Q value was reported to examine the heterogeneity within materials. If the Q value reached a significant level ($p<0.05$), then a meta-regression analysis was carried out to further analyze the effects of selected moderators (Hedges and Pigott, 2004; Borenstein et al., 2010). To compare the effect sizes across factors, this study applied the following equation to examine the difference of $Teta$ (Durlak, 2009; Lenz, 2013): $Teta = Diff / SE$, $Diff = z_1 - z_2$, $SE = \sqrt{Variance\ z_1 + Variance\ z_2}$.

If $|Teta| \geq 2.58$, the factor difference was interpreted as significant ($p<0.01$). To examine publication bias, this study examined Rosenthal's fail-safe number, a funnel plot through the trim-and-fill approach, the rank correlation test, and Egger's regression test (Borenstein et al., 2010).

Results

Descriptive statistics

A total of 62 articles with 88 studies ($N=14,852$) were included in this meta-analysis. In total, 16 studies were master's dissertations and doctoral theses, and the other 72 studies were published in peer-reviewed journals. Specifically, 20 studies ($n=3,324$) investigated the correlation between grammatical knowledge and reading comprehension in group L, 16 studies ($n=2,749$) reported the correlation in group H, 29 studies ($n=5,652$) investigated the correlation in secondary school students (group S), and 21 studies ($n=2,747$) examined the correlation in higher education students (group U). For language type, 41 studies ($n=6,505$) investigated the correlation between grammatical knowledge and reading comprehension in L1 scripts, and 47 studies ($n=8,347$) focused on the relationship between L2 grammatical knowledge and L2 reading comprehension.

Meta-analysis

Table 2 provides the results of the correlation effect size between grammatical knowledge and reading comprehension. The overall

effect size was large (Fisher's $z=0.54$); however, the heterogeneity analysis showed that the materials' heterogeneity was significant ($Q=271.48$, $p<0.001$). Moderator analysis through meta-regression showed that grade group explained 59% variance of materials' heterogeneity and language type explained 13% variance of materials' heterogeneity. Publication bias examination showed the effect size-distributed symmetry (see Figure 1): Rosenthal's fail-safe number was 8,824, the tau value in the rank correlation test was insignificant ($\tau=0.04$, $p>0.05$), and the intercept value in Egger's regression test was not significant (intercept = 0.26, $p>0.05$). The publication bias examination showed that the current study did not have significant publication bias.

Because the majority of materials' heterogeneity resulted from the grade group, this study further investigated the correlation effect size between grammatical knowledge and reading comprehension in four grade groups. Regarding group L, the results showed that the correlation effect size was moderate (Fisher's $z=0.36$). The heterogeneity analysis showed that the materials' heterogeneity was not significant ($Q=23.29$, $p>0.05$). Publication bias examination showed the effect size-distributed symmetry (see Figure 2): Rosenthal's fail-safe number was 1,935, the tau value in rank correlation test was insignificant ($\tau=0.06$, $p>0.05$), and the intercept value in Egger's regression test was not significant (intercept = 0.41, $p>0.05$). The results showed that the current study did not have significant publication bias.

Regarding group H, the results showed the correlation effect size was nearly large (Fisher's $z=0.49$), and heterogeneity analysis showed that materials' heterogeneity was not significant ($Q=11.52$, $p>0.05$). Publication bias examination showed the effect size-distributed symmetry (see Figure 3): Rosenthal's fail-safe number was 2,255, the tau value in rank correlation test was insignificant ($\tau=0.03$, $p>0.05$), and the intercept value in Egger's regression test was not significant (intercept = 0.37, $p>0.05$). The results showed that the current study did not have significant publication bias.

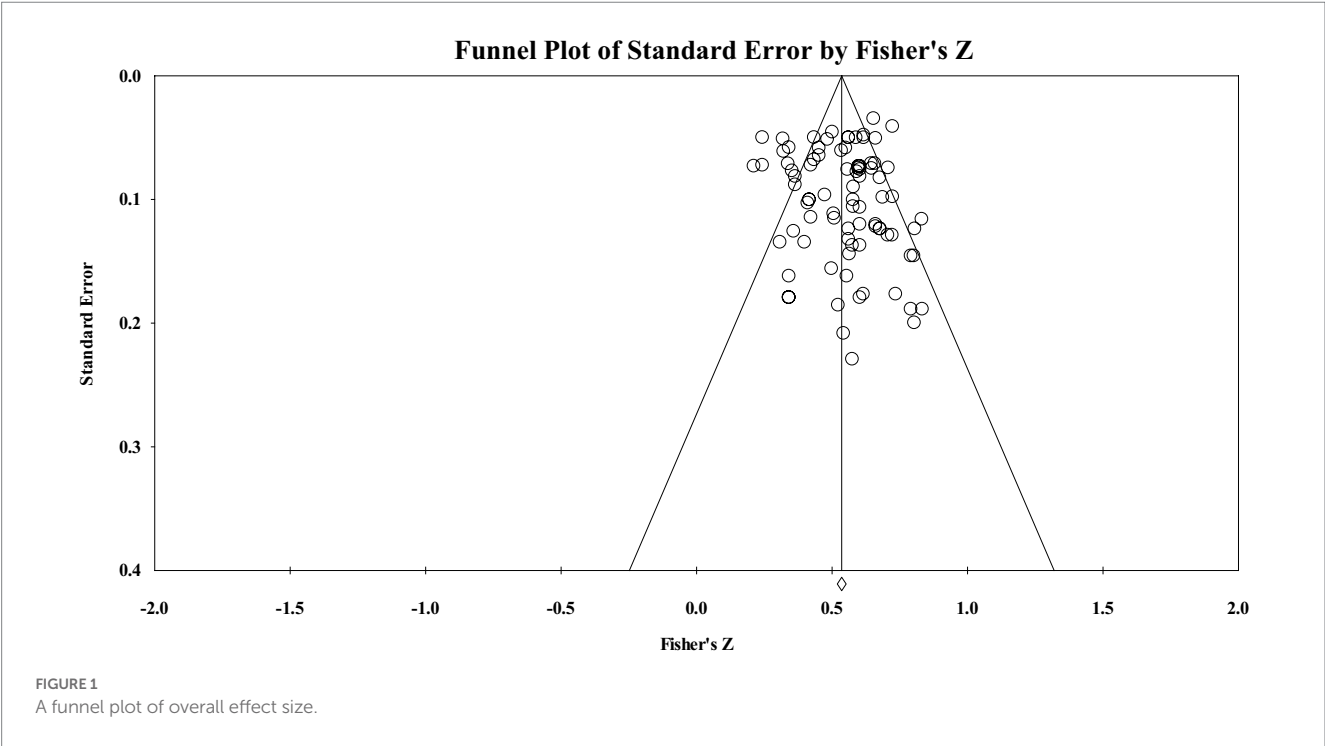
Regarding group S, the results showed that the correlation effect size was large (Fisher's $z=0.61$), and heterogeneity analysis showed that materials' heterogeneity was not significant ($Q=9.81$, $p>0.05$). Publication bias examination showed the effect size-distributed symmetry (see Figure 4): Rosenthal's fail-safe number was 2,555, the tau value in rank correlation test was insignificant ($\tau=-0.09$, $p>0.05$), and the intercept value in Egger's regression test was not significant (intercept = -0.16, $p>0.05$). The results showed that the current study did not have significant publication bias.

Regarding group U, the results showed the correlation effect size was large (Fisher's $z=0.69$), and heterogeneity analysis showed that materials' heterogeneity was not significant ($Q=10.77$, $p>0.05$). Publication bias examination showed the effect size-distributed symmetry (see Figure 5): Rosenthal's fail-safe number was 5,188, the

TABLE 2 Correlation of overall effect between grammatical knowledge and reading comprehension with subgroup analysis.

Variable	<i>k</i>	Fisher's <i>z</i>	Variance	95% <i>CI</i>	<i>Q</i>	<i>N</i> fail-safe	<i>Teta</i> (between grade group)
Overall	86	0.54	0.0003	[0.51, 0.57]	271.48***	855	<i>Teta</i> (H & L) = 4.60, <i>Teta</i> (H & S) = 4.90, <i>Teta</i> (S & U) = 3.27. <i>Teta</i> (L & S) = 10.21, <i>Teta</i> (L & U) = 11.67, <i>Teta</i> (H & U) = 7.07
Group L	20	0.36	0.0004	[0.32, 0.40]	23.29	124	
Group H	16	0.49	0.0004	[0.45, 0.53]	11.52	142	
Group S	29	0.61	0.0002	[0.58, 0.63]	9.81	323	
Group U	21	0.69	0.0004	[0.65, 0.73]	10.77	269	

****p* < 0.001.



tau value in rank correlation test was insignificant ($\tau = 0.14, p > 0.05$), and the intercept value in Egger's regression test was not significant (intercept = 0.36, $p > 0.05$). The results showed that the current study did not have significant publication bias.

Effect size comparison examination showed that the difference between each of the two groups' effect size was significant ($Teta_{H \& L} = 4.60, p < 0.01$; $Teta_{H \& S} = 4.90, p < 0.01$; $Teta_{S \& U} = 3.27, p < 0.01$).

Discussion

The current study showed that the overall correlation effect size was large between grammatical knowledge and reading comprehension. Effect size increased significantly from the lower grade group to the higher grade group and from moderate to large. The results indicated that the cohesive tie on text comprehension interacted significantly with the reading stage. At each reading stage, the interaction effect of learning cognitive condition was not significant on the correlation between reading comprehension and grammatical knowledge (Figure 6).

Grammatical knowledge on reading comprehension

A large effect size was found in the overall correlation between grammatical knowledge and reading comprehension. Jeon and Yamashita (2014) reported that the correlation between grammatical knowledge and reading comprehension in L2 scripts was extremely high ($r = 0.85$) in groups S and U, which was very different from the current results. Jeon and Yamashita (2014) had some obvious limitations. First, they only included 18 articles in their meta-analysis, and the results showed that the *Q* value from heterogeneity analysis was significant. Second, Jeon and Yamashita (2014) did not remove over 3.5 standard deviation effect sizes as the outlier, which showed that the correlation was unclear and did not generate convincing conclusions due to the significant heterogeneity problem. Through the literature search, the current results were consistent with most empirical studies, which showed that the average correlation between grammatical knowledge and reading comprehension was nearly large (Mecartty, 2000; Gersten et al., 2001; Perfetti et al., 2005), indicating that the grammatical knowledge significantly determined the comprehension process.

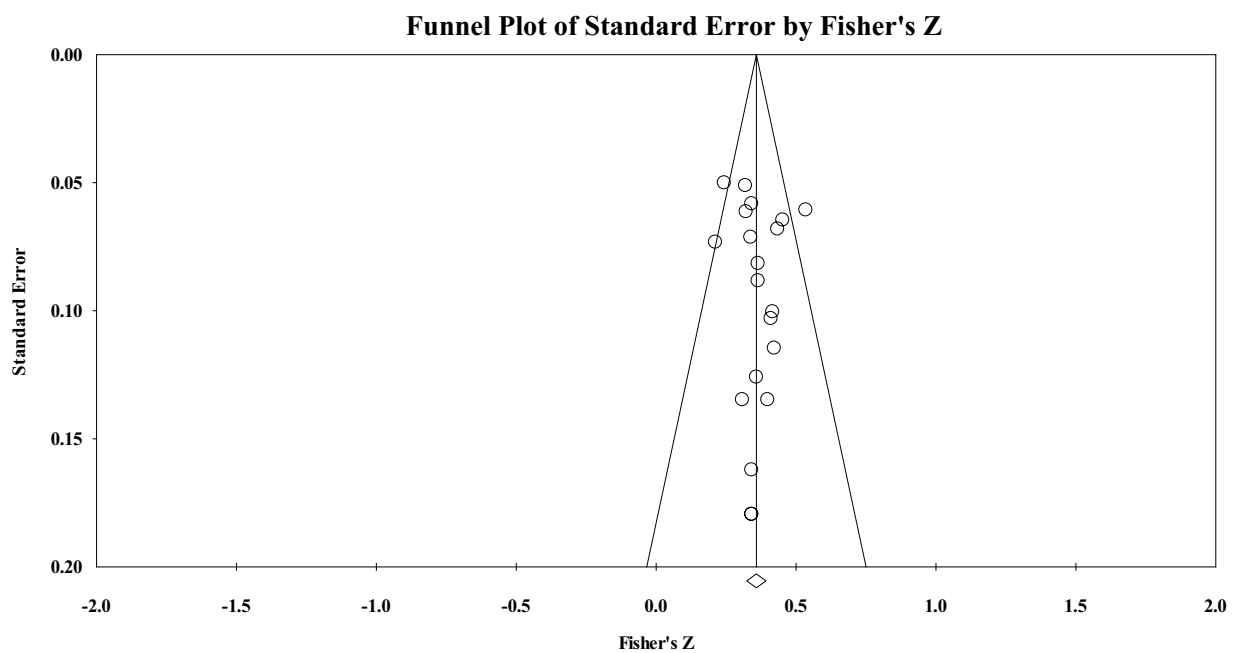


FIGURE 2
A funnel plot of group L effect size.

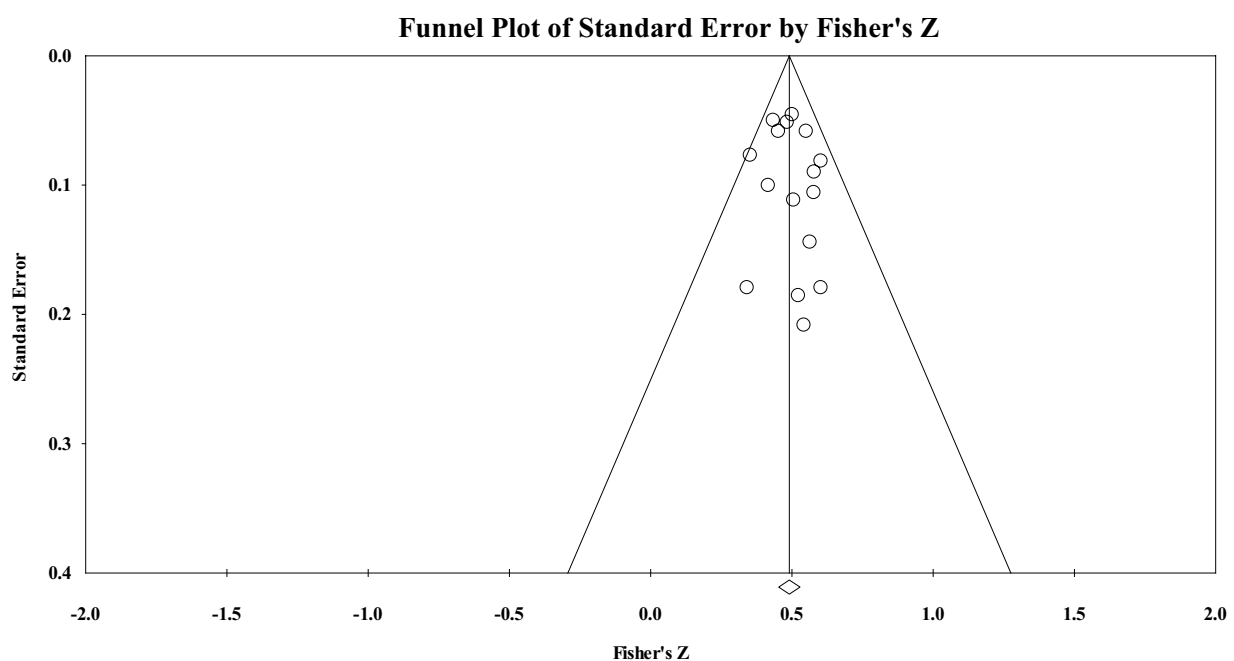


FIGURE 3
A funnel plot of group H effect size.

Moreover, this result informed how the cohesive tie function of grammatical knowledge might have great determining power on reassembling decoded words into phrases and clauses, allowing readers to be more efficient in detecting and correcting reading errors to enhance their comprehension process indirectly (e.g., Slobin, 1966; Bowey, 2005; Kempen et al., 2012).

Grade group effect

The current results showed that the higher grade group had a significantly larger effect size than the lower grade group. This result was consistent with the majority of past longitudinal studies (e.g., Oakhill et al., 2003; Farnia and Geva, 2013). The reasons for the

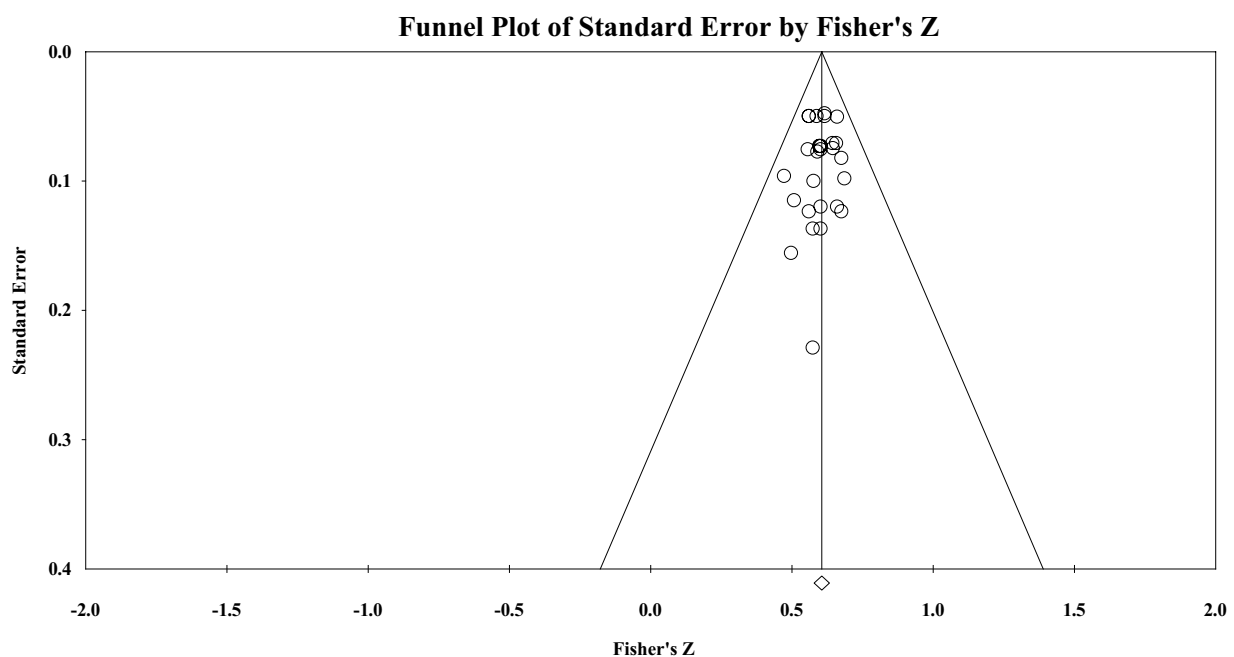


FIGURE 4
A funnel plot of group S effect size.

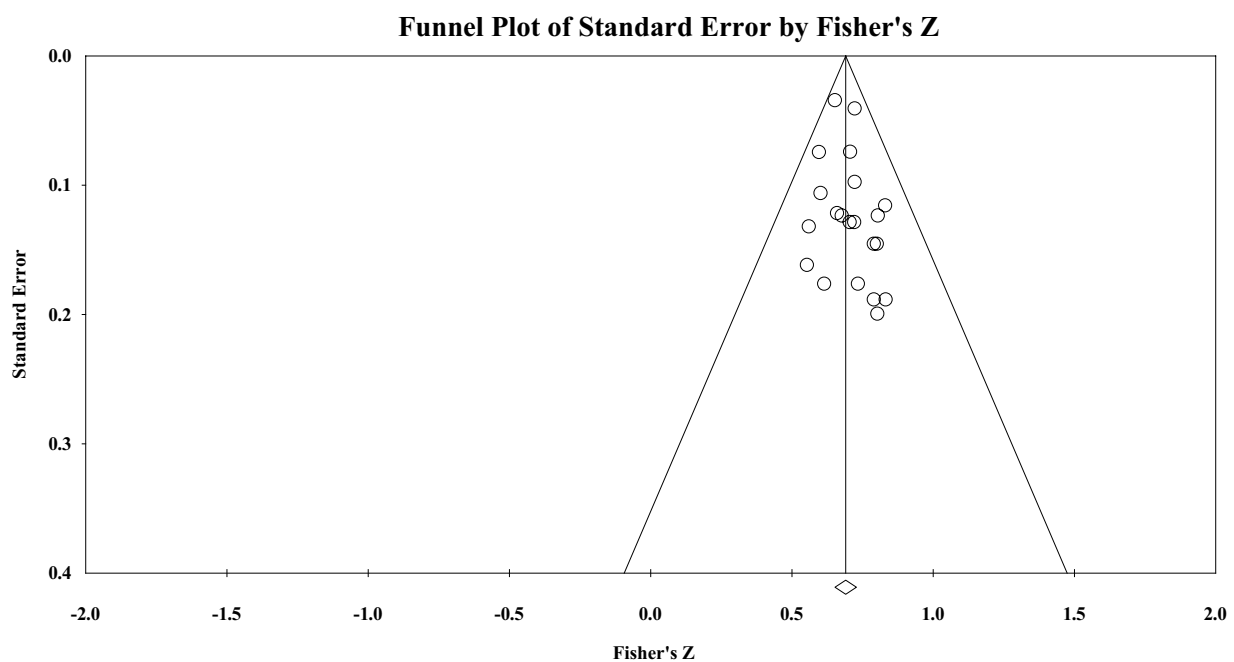
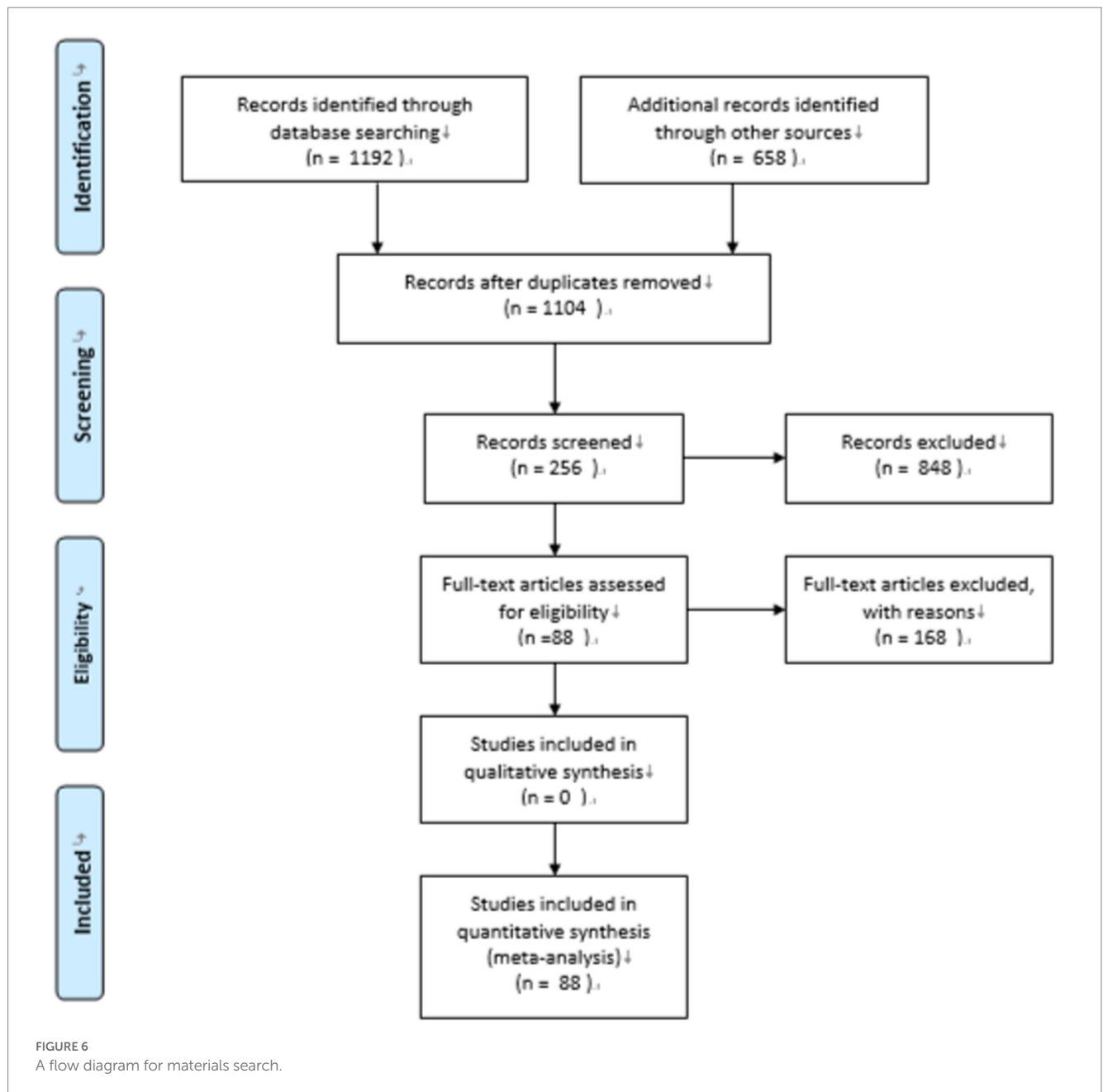


FIGURE 5
A funnel plot of group U effect size.

consistency could be that the requirements, the complicated structure, and the organization level of text comprehension were higher in the higher grade group, resulting in the fact that the requirement of the cohesive tie function was higher in the text coherence clues process (Cain, 2003; Westerveld et al., 2008; Nuske and Bavin, 2015) and organization structure cognition (Oakhill and Cain, 2000; Vaughan

and Dillon, 2006; Kendeou et al., 2007). This result informed that the function of the cohesive tie might continue developing after the age of 16, which was different from the majority of comprehension factors, such as decoding ability, vocabulary knowledge, and metalinguistic knowledge (Mol and Bus, 2011; García and Cain, 2014; Trapman et al., 2014). The current result also echoed reading stage theory regarding



grammatical knowledge function on text reading comprehension, where the more complicated-level text comprehension needed a higher application of grammatical knowledge for coherence inference.

Language type effect

In each grade group, the language type was not a significant moderator on the correlation between grammatical knowledge and reading comprehension. This result echoed the previous study, which showed that linguistic skills could work similarly across languages' scripts (Cummins, 1986; Gottardo and Mueller, 2009; Perfetti and Stafura, 2014). In other words, grammatical knowledge could have a transfer effect across all languages, suggesting that grammatical knowledge had an independent cognition function on text

comprehension. Readers received text information and delivered target comprehension information into the situation model-building process. The grammatical knowledge worked together with other linguistic skills on text meaning coherence, which indicates that this process encompassed the integrated process and the synthesizing process on word reading and sentence meaning judgment in situation model construction. According to the current results, grammatical knowledge may have an independent effect on reading comprehension across various scripts' text reading, which means that, for bilingual learners' reading comprehension, a transfer or compensation effect would help readers to apply either L1 or L2 grammatical knowledge to other language scripts' comprehension (Sparks et al., 2008). Therefore, readers only need to fully develop grammatical knowledge in one language script, and then, the grammatical knowledge can work in other scripts' text comprehension processes.

Limitations

The current study has several limitations. First, this study selected materials that were only written in Chinese or English. Materials written in other languages were not considered. Second, the selected materials only reported the correlation between grammatical knowledge and reading comprehension in Chinese reading, Dutch reading, English reading, and French reading. Next, although past studies mentioned that grammatical knowledge should involve morphological knowledge and verb forms, due to inconclusive findings, this study excluded these two variables. In future studies, if the researcher has an agreed definition in the grammatical knowledge category, it should be further explored. Finally, meta-regression only presented how the extent of the heterogeneity of the materials could be explained by the moderator, while the interaction effect mechanism of each moderator was not identified. For example, this study reported that the grade group and language types were two significant moderators on the overall correlation but did not investigate the internal working interaction effect between these two moderators.

Conclusion

The study's conclusions were drawn from the combined results of 86 studies conducted with more than 14,000 readers. The reading comprehension measures only included Chinese, Dutch, English, and French reading comprehension test in the current study's database. To summarize, this meta-analysis confirmed that the overall correlation between grammatical knowledge and reading comprehension was large (Fisher's $z = 0.54$). The cohesive tie of the grammatical knowledge on reading comprehension had a significant interaction effect with the reading stage, and the higher requirement of reading comprehension showed higher function application of cohesive tie. The results indicated that the cohesive tie function of grammatical knowledge had

a transfer effect across different scripts' reading comprehension. The correlation pattern between grammatical knowledge and text reading comprehension was similar across Chinese, Dutch, English, and French scripts.

Author contributions

XM and YD had similar contribution, including data analysis and data collection. All authors contributed to the article and approved the submitted version.

Funding

This paper received financial support by intramural project of Guangzhou Huashang College (2021HSDS36).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- *Adlof, S. M., Catts, H. W., and Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *J. Learn. Disabil.* 43, 332–345. doi: 10.1177/0022219410369067
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- *Aryadoust, V., and Baghaei, P. (2016). Does EFL Readers' lexical and grammatical knowledge predict their Reading ability? Insights from a perceptron artificial neural network study. *Educ. Assess.* 21, 135–156. doi: 10.1080/10627197.2016.1166343
- Bentin, S., Deutsch, A., and Liberman, I. Y. (1990). Syntactic competence and reading ability in children. *J. Exp. Child Psychol.* 49, 147–172. doi: 10.1016/0022-0965(90)90053-B
- Blackmore, A. M., and Pratt, C. (1997). Grammatical awareness and reading in grade 1 children. *Merrill-Palmer Quarterly* (1982-) 43, 567–590. doi: 10.1023/A:1024538123804
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* 1, 97–111. doi: 10.1002/jrsm.12
- *Bowie, J. A. (2005). *Predicting individual differences in learning to read. The science of reading: A handbook*. 155–172, Hoboken, NJ: Blackwell Publishing Ltd..
- *Brimo, D., Apel, K., and Fountain, T. (2017). Examining the contributions of syntactic awareness and syntactic knowledge to reading comprehension. *J. Res. Read.* 40, 57–74. doi: 10.1111/1467-9817.12050
- Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *Br. J. Dev. Psychol.* 21, 335–351. doi: 10.1348/02615100322277739
- Cain, K., and Nash, H. (2011). The influence of connectives on young readers' processing and comprehension of text. *J. Educ. Psychol.* 103, 429–441. doi: 10.1037/a0022824
- Cain, K., Oakhill, J., and Bryant, P. (2004). Children's reading comprehension ability: concurrent prediction by working memory, verbal ability, and component skills. *J. Educ. Psychol.* 96, 31–42. doi: 10.1037/0022-0663.96.1.31
- Carretti, B., Borella, E., Cornoldi, C., and De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: a meta-analysis. *Learn. Individ. Differ.* 19, 246–251. doi: 10.1016/j.lindif.2008.10.002
- *Catts, H. W., Fey, M. E., Tomblin, J. B., and Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *J. Speech Lang. Hear. Res.* 45, 1142–1157. doi: 10.1044/1092-4388(2002/093)
- Chall, J. (1996). *Stages of reading development (2nd ed.)*. Fort Worth, TX: Harcourt Brace.
- *Chamberlain, C., and Mayberry, R. I. (2008). American sign language syntactic and narrative comprehension in skilled and less skilled readers: bilingual and bimodal evidence for the linguistic basis of reading. *Appl. Psycholinguist.* 29, 367–388. doi: 10.1017/S014271640808017X
- *Chaney, C. (1998). Preschool language and metalinguistic skills are links to reading success. *Appl. Psycholinguist.* 19, 433–446. doi: 10.1017/s0142716400010250
- *Chen, B. G., and Chen, Y. L. (2008). The relationships of primary school students' syntactic awareness and phonological awareness with their performance of Reading comprehension. *Psychol. Sci.* 31, 892–895. doi: 10.16719/j.cnki.1671-6981.2008.04.022
- *Chik, P. P. M., Ho, C. S. H., Yeung, P. S., Chan, D. W. O., Chung, K. K. H., and Luan, H., ... & Lau, W. S. Y. (2012). Syntactic skills in sentence reading comprehension among Chinese elementary school children. *Read. Writ.*, 25, 679–699. doi: 10.1007/s11145-010-9293-4
- *Chung, K. K., Ho, C. S. H., Chan, D. W., Tsang, S. M., and Lee, S. H. (2013). Contributions of syntactic awareness to reading in Chinese-speaking adolescent readers with and without dyslexia. *Dyslexia* 19, 11–36. doi: 10.1002/dys.1448

- Cummins, J. (1986). Empowering minority students: a framework for intervention. *Harv. Educ. Rev.* 56, 18–37. doi: 10.1080/0013191860380107
- *Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., and Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Ann. Dyslexia* 59, 34–54. doi: 10.1007/s11881-009-0022-0
- del Pilar Jiménez, E., Alarcón, R., and de Vicente-Yague, M. I. (2019). Reading intervention: correlation between emotional intelligence and reading competence in high school students. *Revista de Psicodidáctica (English Ed.)* 24, 24–30. doi: 10.1016/j.pscoc.2018.10.001
- Demont, E., and Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *Br. J. Educ. Psychol.* 66, 315–332. doi: 10.1111/j.2044-8279.1996.tb01200.x
- *Deng, J. (2014). The influence of syntactic knowledge, vocabulary breadth and vocabulary depth on Reading comprehension-evidence from Chinese EFL students. Master dissertation. Fujian: Fujian Normal University.
- Dong, Y., Chow, B. W. Y., Mo, J., Zheng, H. Y., Wu, S. X. Y., and Yang, J. (2022). Comprehension of quantifier sentences in poor readers with different levels of arithmetic proficiency. *Read. Writ.* 1–22. doi: 10.1007/s11145-022-10344-9
- Dong, Y., Chow, B. W. Y., Wu, S. X. Y., Zhou, J. D., and Zhao, Y. M. (2021). Enhancing poor readers' reading comprehension ability through word semantic knowledge training. *Read. Writ. Q.* 37, 348–364. doi: 10.1080/10573569.2020.1820410
- Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *J. Pediatr. Psychol.* 34, 917–928. doi: 10.1093/jpepsy/jsp004
- *Farnia, E., and Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *J. Res. Read.* 36, 389–421. doi: 10.1111/jrir.12003
- Fender, M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Lang. Learn.* 51, 319–396. doi: 10.1111/0023-8333.00157
- *Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A., and Trueman, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th–10th grades. *J. Educ. Psychol.* 107, 884–899. doi: 10.1037/edu0000026
- Gahl, S., and Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language* 80, 748–775. doi: 10.1353/lan.2004.0185
- García, J. R., and Cain, K. (2014). Decoding and reading comprehension: a meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Rev. Educ. Res.* 84, 74–111. doi: 10.3102/0034654313499616
- Gersten, R., Fuchs, L. S., Williams, J. P., and Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: a review of research. *Rev. Educ. Res.* 71, 279–320. doi: 10.2307/3516086
- *Geva, E., and Farnia, E. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Read. Writ.* 25, 1819–1845. doi: 10.1007/s11145-011-9333-8
- *Goff, D. A., Pratt, C., and Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Read. Writ.* 18, 583–616. doi: 10.1007/s11145-004-7109-0
- *Gong, S. Y., Xu, X. C., and Liu, H. S. (2010). The relationship between English syntactic awareness, Reading comprehension monitoring and contextual information use. *Psychol. Explor.* 30, 34–38. doi: 10.3969/j.issn.1003-5184.2010.03.008
- *Gong, S. Y., Xu, X. C., Ye, J., and Han, Y. P. (2009). The relationship between English phonological and syntactic awareness working memory and English Reading in grade one of junior middle school. *J. Educ. Sci. Hunan Norm. Univ.* 8, 91–95.
- *Gottardo, A., and Mueller, J. (2009). Are first-and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *J. Educ. Psychol.* 101, 330–344. doi: 10.1037/a0014320
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2012). *Teaching and researching reading*. Foreign Language Teaching and Research Press. Beijing.
- *Guan, F. (2007). The influence of meta-language awareness on Students' English Reading. *J. Liaoning Teach. Coll.* 2007, 75–76. doi: 10.3969/j.issn.1008-3898.2007.01.040
- *Gui, M. (2018). The impact of component skills on Chinese foreign language learners' Reading comprehension under component skills analysis. *J. PLA Univ. Foreign Lang.* 41, 99–107.
- Hedges, L. V., and Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychol. Methods* 9, 426–445. doi: 10.1037/1082-989x.9.4.426
- Hedges, L. V., Tipton, E., and Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* 1, 39–65. doi: 10.1002/jrsm.5
- Hjetland, H. N., Lervåg, A., Lyster, S. A. H., Hagtvet, B. E., Hulme, C., and Melby-Lervåg, M. (2019). Pathways to reading comprehension: a longitudinal study from 4 to 9 years of age. *J. Educ. Psychol.* 111, 751–763. doi: 10.1037/edu00000321
- *Jafari, A. M., and Rad, N. F. (2016). The influence of phonological and grammatical awareness on EFL students' reading performance. *J. Lang. Teach. Res.* 7, 1164–1173. doi: 10.17507/jltr.0706.14
- Jeon, E. H., and Yamashita, J. (2014). L2 reading comprehension and its correlates: a meta-analysis. *Lang. Learn.* 64, 160–212. doi: 10.1111/lang.12034
- *Jiang, X. F. (2003). The study on the influence of phonological awareness and syntactic awareness in Student's English Reading. Master dissertation. Chongqing: Southwest China Normal University.
- Kempen, G., Olsthoorn, N., and Sprenger, S. (2012). Grammatical workspace sharing during language production and language comprehension: evidence from grammatical multitasking. *Lang. Cogn. Process.* 27, 345–380. doi: 10.1080/01690965.2010.544583
- Kendeou, P., Van Den Broek, P., White, M. J., and Lynch, J. (2007). "Comprehension in preschool and early elementary children: skill development and strategy interventions" in *Reading comprehension strategies: theories, interventions, and technologies*. ed. D. S. McNamara (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 27–45.
- Kern, R. G. (1989). Second language reading strategy instruction: its effects on comprehension and word inference ability. *Mod. Lang. J.* 73, 135–149. doi: 10.2307/326569
- *Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., and Wagner, R. K. (2011). Componential skills of beginning writing: an exploratory study. *Learn. Individ. Differ.* 21, 517–525. doi: 10.1016/j.lindif.2011.06.004
- *Kim, J., and Cho, Y. (2015). Proficiency effects on relative roles of vocabulary and grammar knowledge in second language Reading. *English Teach.* 70, 75–96.
- Kleeck, A. V. (2008). Providing preschool foundations for later reading comprehension: the importance of and ideas for targeting inferencing in storybook-sharing interventions. *Psychol. Sch.* 45, 627–643. doi: 10.1002/pits.20314
- Koda, K. (1992). The effects of lower-level processing skills on FL reading performance: implications for instruction. *Mod. Lang. J.* 76, 502–512. doi: 10.1111/j.1540-4781.1992.tb05400.x
- *Lasagabaster, D. (2001). The effect of knowledge about the L1 on foreign language skills and grammar. *Int. J. Biling. Educ. Biling.* 4, 310–331. doi: 10.1080/13670050108667735
- *Leider, C. M., Proctor, C. P., Silverman, R. D., and Harring, J. R. (2013). Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Read. Writ.* 26, 1459–1485. doi: 10.1007/s11145-013-9427-6
- Lenz, A. S. (2013). Calculating effect size in single-case research: a comparison of nonoverlap methods. *Meas. Eval. Couns. Dev.* 46, 64–73. doi: 10.1177/0748175612456401
- *Lesaux, N. K., Lipka, O., and Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Read. Writ.* 19, 99–131. doi: 10.1007/s11145-005-4713-6
- *Li, J. D., Li, K., and Chen, Q. Z. (2008). A research on the relationship between the second and third language of minority students in Xinjiang Normal university. *J. Xinjiang Norm. Univ.* 29, 129–132.
- *Li, R. B., Li, G. Z., Su, Y. K., and Wang, B. C. (2016). Effects of language experience on dialect(s)-speaking Children's metalinguistic awareness and linguistic competence. *Appl. Linguis.* 1, 26–35.
- *Liao, R. Y. (2012). Research into factors influencing students' Reading ability in English. Master dissertation. Nanjing: Nanjing Normal University.
- *Lu, M., and Zhang, H. (2015). The relationship between English meta-linguistic awareness, vocabulary knowledge and Reading comprehension of junior middle school students. *J. PLA Univ. Foreign Lang.* 38, 74–80.
- *Luo, Y. P. (2010). Relationship between English phonological and syntactic awareness and English Reading ability among college students. Master dissertation. Guangxi: Guangxi Normal University.
- Mecartty, F. H. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Appl. Lang. Learn.* 11, 323–348.
- *Mokhtari, K., and Niederhauser, D. S. (2013). Vocabulary and syntactic knowledge factors in 5th grade students' reading comprehension. *Int. Electron. J. Element. Educ.* 5, 157–170.
- Mol, S. E., and Bus, A. G. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychol. Bull.* 137, 267–296. doi: 10.1037/a0021890
- Muter, V., Hulme, C., Snowling, M. J., and Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Dev. Psychol.* 40, 665–681. doi: 10.1037/0012-1649.40.5.665
- Nuske, H. J., and Bavin, E. L. (2015). Narrative comprehension in 4–7-year-old children with autism: testing the weak central coherence account. *Int. J. Lang. Commun. Disord.* 46, 100824014249025–100824014249119. doi: 10.3109/13682822.2010.484847
- Oakhill, J., and Cain, K. (2000). Children's difficulties in text comprehension: assessing causal issues. *J. Deaf. Stud. Deaf. Educ.* 5, 51–59. doi: 10.1093/deaf/5.1.51
- *Oakhill, J. V., and Cain, K. (2012). The precursors of reading ability in young readers: evidence from a four-year longitudinal study. *Sci. Stud. Read.* 16, 91–121. doi: 10.1080/10888438.2010.529219

- *Oakhill, J. V., Cain, K., and Bryant, P. E. (2003). The dissociation of word reading and text comprehension: evidence from component skills. *Lang. Cogn. Process.* 18, 443–468. doi: 10.1080/01690960344000008
- *O'Connor, I. M., and Klein, P. D. (2004). Exploration of strategies for facilitating the reading comprehension of high-functioning students with autism spectrum disorders. *J. Autism Dev. Disord.* 34, 115–127. doi: 10.1023/b:jadd.0000022603.44077.6b
- *Park, H., Choi, S., and Lee, M. (2012). Visual input enhancement, attention, grammar learning, & Reading comprehension: an eye movement study. *English Teach.* 67, 241–265. doi: 10.15858/engtea.67.4.201212.241
- Perfetti, C. A., Landi, N., and Oakhill, J. (2005). “The acquisition of reading comprehension skill” in *The science of reading: a handbook*. eds. M. J. Snowling and C. Hulme (Hoboken, NJ: Blackwell Publishing), 227–247.
- Perfetti, C., and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Sci. Stud. Read.* 18, 22–37. doi: 10.1080/10888438.2013.827687
- *Proctor, C. P., Silverman, R. D., Harring, J. R., and Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Read. Writ.* 25, 1635–1664. doi: 10.1007/s11145-011-9336-5
- Purpura, J. E. (2013). *Assessing grammar. The companion to language assessment*. Hoboken, NJ: John Wiley & Sons, Inc.
- Rego, L. L. B., and Bryant, P. E. (1993). The connection between phonological, syntactic and semantic skills and children's reading and spelling. *Eur. J. Psychol. Educ.* 8, 235–246. doi: 10.1007/BF03174079
- *Rescorla, L. (2000). Do late-talking toddlers turn out to have reading difficulties a decade later? *Ann. Dyslexia* 50, 85–102. doi: 10.1007/s11881-000-0018-2
- *Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *J. Speech Lang. Hear. Res.* 45, 360–371. doi: 10.1044/1092-4388(2002/028)
- Rodd, J. M., Vitello, S., Woollams, A. M., and Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: an activation likelihood estimation meta-analysis. *Brain Lang.* 141, 89–102. doi: 10.1016/j.bandl.2014.11.012
- *Roth, F. P., Speece, D. L., and Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *J. Educ. Res.* 95, 259–272. doi: 10.1080/00220670209596600
- *Schatzschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., and Foorman, B. R. (2004). Kindergarten prediction of reading skills: a longitudinal comparative analysis. *J. Educ. Psychol.* 96, 265–282. doi: 10.1037/0022-0663.96.2.265
- *Shen, Y. L. (2014). The relative significance of vocabulary breadth and syntactic knowledge in the prediction of Reading comprehension test performance. *J. Hefei Norm. Univ.* 4, 87–92. doi: 10.3969/j.issn.1674-2273.2014.04.018
- *Shen, Y. L., and Tao, W. (2011). The relative significance of vocabulary breadth and syntactic knowledge in the prediction of Reading comprehension test performance. *Chinese J. Appl. Linguist.* 34, 113–126.
- Shibasaki, H., Tokimoto, S., Ono, Y., Inoue, T., and Tamaoka, K. (2015). English reading comprehension by Japanese high school students: structural equation modeling including working memory and L1 literacy. *Open J. Modern Linguist.* 05, 443–458. doi: 10.4236/ojml.2015.55039
- Shiotsu, T. (2010). *Components of L2 reading: Linguistic and processing factors in the reading test performances of Japanese EFL learners*. Stuttgart: Ernst Klett Sprachen.
- *Shiotsu, T., and Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Lang. Test.* 24, 99–128. doi: 10.1177/0265532207071513
- Silva, M., and Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *J. Educ. Psychol.* 107, 321–331. doi: 10.1037/a0037769
- *Silverman, R. D., Proctor, C. P., Harring, J. R., Hartranft, A. M., Doyle, B., and Zelinke, S. B. (2015). Language skills and reading comprehension in English monolingual and Spanish–English bilingual children in grades 2–5. *Read. Writ.* 28, 1381–1405. doi: 10.1007/s11145-015-9575-y
- Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *J. Verbal Learn. Verbal Behav.* 5, 219–227. doi: 10.1016/S0022-5371(66)80023-3
- Snow, C. E. (2002). *Reading for understanding: Toward a research and development program in reading comprehension*. Santa Monica, CA: RAND.
- Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., and Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *J. Educ. Psychol.* 100, 162–174. doi: 10.1037/0022-0663.100.1.162
- Tagarelli, K. M., Borges Mota, M., and Rebuschat, P. (2011). “The role of working memory in the implicit and explicit learning of languages” in *Proceedings of the 33rd annual conference of the cognitive science society*. eds. L. Carlson, C. Hölscher and T. F. Shipley (Austin, EX: Cognitive Science Society), 2061–2066.
- *Tang, Y. (2013). A study of the function of explicit and implicit grammatical knowledge in senior grade one students' Reading comprehension and the implication on grammar instruction. Master dissertation. Tianjin: Tianjin Normal University.
- Trapman, M., Van Gelderen, A., Van Steensel, R., Van Schooten, E., and Hulstijn, J. (2014). Linguistic knowledge, fluency and meta-cognitive knowledge as components of reading comprehension in adolescent low achievers: differences between monolinguals and bilinguals. *J. Res. Read.* 37, S3–S21. doi: 10.1111/j.1467-9817.2012.01539.x
- Tunmer, W. E. (1989). The role of language-related factors in reading disability. in D. Shankweiler and I. Y. Liberman (Eds.), *Phonology and reading disability: solving the reading puzzle*, 91–131. Ann Arbor, MI: The University of Michigan Press.
- *Van Gelderen, A., Schoonen, R., De Gloppe, K., Hulstijn, J., Simis, A., Snellings, P., et al. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first-and second-language Reading comprehension: a componential analysis. *J. Educ. Psychol.* 96, 19–30. doi: 10.1037/0022-0663.96.1.19
- *Van Gelderen, A., Schoonen, R., De Gloppe, K., Hulstijn, J., Snellings, P., Simis, A., et al. (2003). Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: a structural equation modeling approach. *Int. J. Biling.* 7, 7–25. doi: 10.1177/13670069030070010201
- Vaughan, M. W., and Dillon, A. (2006). Why structure and genre matter for users of digital information: a longitudinal experiment with readers of a web-based newspaper. *Int. J. Human-Comput. Stud.* 64, 502–526. doi: 10.1016/j.ijhcs.2005.11.002
- *Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., and Chen, R. (2007). Components of reading ability: multivariate evidence for a convergent skills model of reading development. *Sci. Stud. Read.* 11, 3–32. doi: 10.1207/s1532799xssr1101_2
- Vinyals, O., Kaiser, L., Koo, T., Petrov, S., Sutskever, I., and Hinton, G. (2015). *Grammar as a foreign language*. arXiv.org. Retrieved February 28, 2023, from <https://arxiv.org/abs/1412.7449>.
- Waltzman, D. E., and Cairns, H. S. (2000). Grammatical knowledge of third grade good and poor readers. *Appl. Psycholinguist.* 21, 263–284. doi: 10.1017/S014217640000206X
- *Wang, K. Y. (2001). The relationships between phonological, syntactic and semantic skills in English Reading performance of Chinese middle school students. Master dissertation. Gansu: Northwest Normal University.
- *Wang, C. Y. (2012). A study of the correlation between explicit grammatical knowledge, implicit grammatical knowledge and Reading comprehension ability of senior high school students. Master dissertation. Anhui: Anhui University.
- *Wei, Y. Z. (2005). Correlation between grammatical sensitivity and reading comprehension in Chinese English learners. *Foreign Lang. World* 1, 40–46.
- Westerveld, M. F., Gillon, G. T., and Moran, C. (2008). A longitudinal investigation of oral narrative skills in children with mixed reading disability. *Int. J. Speech Lang. Pathol.* 10, 132–145. doi: 10.1080/14417040701422390
- *Whyte, E. M., Nelson, K. E., and Scherf, K. S. (2014). Idiom, syntax, and advanced theory of mind abilities in children with autism spectrum disorders. *J. Speech Lang. Hear. Res.* 57, 120–130. doi: 10.1044/1092-4388(2013/12-0308)
- *Wu, S. N. (2016). Influence of vocabulary, syntactic and metacognitive awareness of Japanese students on their Chinese Reading comprehension. *Lang. Teach. Linguist. Stud.* 2, 59–66.
- *Xiang, T. C. (2016). The influence of syntactic awareness on the English Reading comprehension of students with learning disabilities in junior middle school and an intervention. Master dissertation. Shaanxi: Shaanxi Normal University.
- *Yan, L. (2012). A comparative study on the effects of lexical chunks recognition and syntax identification on Reading comprehension. Master dissertation. Chengdu: Southwest Petroleum University.
- *Zhang, Y. X. (2011). Vocabulary knowledge, grammatical knowledge, working memory and EFL Reading comprehension. Master dissertation. Shanghai: Shanghai Jiaotong University.
- Zhang, D. (2012). Vocabulary and grammar knowledge in second language reading comprehension: a structural equation modeling study. *Mod. Lang. J.* 96, 558–575. doi: 10.1111/j.1540-4781.2012.01398.x
- *Zhang, Y. H. (2017). The effects of metalinguistic awareness on college English Reading comprehension. Master dissertation. Jinan: Shandong University.
- Zhang, H., and Koda, K. (2018). Vocabulary knowledge and morphological awareness in Chinese as a heritage language (CHL) reading comprehension ability. *Read. Writ.* 31, 53–74. doi: 10.1007/s11145-017-9773-x
- *Zhang, D. B., and Zhao, S. H. (2011). Implicit and explicit grammatical knowledge in second language reading comprehension: a study of advanced Chinese EFL learners. *Foreign Lang. Teach. Res.* 43, 387–399.
- Zwaan, R., and Rapp, D. (2006). “Discourse comprehension” in *Handbook of psycholinguistics*. eds. M. Traxler and M. A. Gernsbacher. 2nd ed (Burlington, MA: Academic Press), 725–764.



OPEN ACCESS

EDITED BY

María Isabel De Vicente-Yagüe Jara,
University of Murcia, Spain

REVIEWED BY

Moisés Selfa Sastre,
Universitat de Lleida, Spain
José Manuel De Amo Sánchez-Fortún,
University of Almería, Spain
Nieves Gutiérrez Ángel,
University of Almería, Spain
Inmaculada Clotilde Santos Díaz,
University of Malaga, Spain

*CORRESPONDENCE

Rosa Tabernero-Sala
✉ rostab@unizar.es

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 30 January 2023

ACCEPTED 13 March 2023

PUBLISHED 06 April 2023

CITATION

Tabernero-Sala R and Colón-Castillo MJ
(2023) The promotion of critical reading
through the digital environment: A study on
the virtual epitexts used to promote children's
picturebooks.
Front. Psychol. 14:1154513.
doi: 10.3389/fpsyg.2023.1154513

COPYRIGHT

© 2023 Tabernero-Sala and Colón-Castillo.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

The promotion of critical reading through the digital environment: A study on the virtual epitexts used to promote children's picturebooks

Rosa Tabernero-Sala* and María Jesús Colón-Castillo

Department of Specific Didactics, University of Zaragoza, Zaragoza, Spain

The most recent research studies in the field of reading describe a new cultural ecosystem in which analog and digital reading coexist and contribute to transform what is read, either through the way reading is performed or by promoting reading. In this context, the training of critical readers is particularly important, an aspect emphasized by UNESCO and the curriculum frameworks based on its premises. In order to provide data for reflection on this question, this paper presents an essentially qualitative and interpretive documentary study of a sample of 836 virtual epitexts that promote children's picturebooks. The selected documents consist of the postings by 45 publishing houses between 2020 and 2022 on their YouTube and Vimeo channels. The results of the content analysis present the current tendencies in digital promotion of children's books and the strategies most likely to encourage critical reading. The insistence on the author's presence, the emphasis on the materiality of the book as an object, the strengthening of artistic discourse and the hybridization of reality and fiction, among other aspects, all propose a type of reading that favors the development of critical thinking. The results are complemented by a selection of virtual epitexts suggested to mediators and readers as resources of interest in promoting critical reading in socio-educational contexts.

KEYWORDS

virtual epitexts, book trailer, critical reading, critical reader, children's picturebooks, promotion by publishers, promotion of reading, non-fiction book

1. Introduction

The coexistence of analog and digital paradigms in today's information society has established a new reading ecosystem built on hybridization. This is evidenced in research into reading habits cited by scholars such as [Cordón-García \(2018, 2020\)](#) or [Lluch \(2018\)](#). The paradigm shift in reading has led to significant changes in the way children and young people approach information, knowledge and literature ([Lluch, 2018](#)). Social media have further influenced access to reading by throwing literate practices into an environment of constant and rapid modifications ([Cruces, 2017](#); [Lluch, 2017, 2018, 2021](#); [Cordón-García, 2018](#)). Contrary to expectations, the progress of digital publishing has strengthened the material aspect of reading to such an extent that the physicality of the book as an object

has become one of the key elements in the new ecosystem, since, as Littau (2006) states, the format determines both the content and the layout.

In light of the challenges the digital society poses to reader training, the research conducted by Wolf (2018) is of particular interest. In her study, she defends the construction of a biliterate reading brain as the way we read influences the way we think. Wolf (2018) understands that the different digital experiences modern readers use—social media, games, interactive platforms—distance them from the in-depth reading connected to the development of critical thinking. The author, therefore, suggests a paradigm that combines analog and digital cultures to develop biliteracy from the beginning, with an emphasis on analog means in the early years to ensure access to the digital environment with guarantees at a later stage. The 2020 report by the Centro Regional para el Fomento del Libro en América Latina y el Caribe (Cerlalc) *Lectura en papel vs. lectura en pantalla* (Kovač and Van Der Weel, 2020) follows along the same lines. In one of its chapters, Støle (2020) develops the concept of in-depth reading by insisting on the conditions this entails, since reading on a screen, in her opinion, is inferior in terms of comprehension as digital media require less attention. In fact, as the Cerlalc Report (Kovač and Van Der Weel, 2020) indicates, difficulty in reading extensive texts rises as the number of personal digital devices available to children increases (Støle, 2020). This idea is further confirmed by the results of the report *Developing Literacy Skills in a Digital World* OECD (2021), presented by the OECD and based on the 2018 Pisa Report. Likewise, after reviewing 54 studies, Delgado et al. (2018) state that printed text is associated with better comprehension as paper reading requires greater concentration compared to screen reading (Gil-Pelluch et al., 2020). Furthermore, Schilhab et al. (2020) emphasize the sense of stability that printed paper bestows on a physical book, making it the most suitable medium for the development of in-depth reading. The European Commission report on the creation of a Work Plan for Culture (2015–2018), *Promoting Reading in the Digital Environment* (Unión Europea, 2016), and the report by the Federación del Gremio de Editores en España (Millán, 2017), among others, agree on this idea. Both reports provide robustness to this line of research, which emphasizes one of the challenges posed by the new ecosystem regarding the training of readers who are capable of selecting and organizing information to convert it into knowledge. In short, the digital society offers a new definition of the agents involved in the new reading ecosystem. Thus, the concept of the reader as prosumer (García-Roca and De-Amo, 2019; De Amo and García-Roca, 2021), the new role played by mediation (Lluch, 2017, 2021; Zafra, 2017; De Amo Sánchez-Fortún, 2021), reflection on the authorial paradigm (Unsworth, 2015; Tabernero-Sala, 2019; Tabernero-Sala et al., 2022) and the entity that virtual epitexts acquire in the promotion of reading (Lluch et al., 2015; Lluch, 2018) define the need to experiment with reading models that help to train critical readers who are able to cope with the tension between digital and analog cultures, as recommended in the studies by Baron (2015, 2021) and the exploratory research by Mizrachi et al. (2018) and Mangen et al. (2019), among others.

In this context, one of the challenges that the information society faces regarding reader training is none other than how to incorporate critical citizens into the new cultural ecosystem—citizens who possess the necessary strategies to express themselves

using their own judgment in the midst of digital and analog paradigms (De Amo Sánchez-Fortún, 2021).

In the field of education, the reports by UNESCO insist on the need to develop critical thinking as a cornerstone in the construction of democratic, participatory societies capable of collectively meeting the challenges of the twenty-first century (Caro-Valverde, 2018). This is also expressed in the report *Reimagining our Futures Together: A New Social Contract for Education*—commissioned by UNESCO and prepared by the *Comisión Internacional sobre los Futuros de la Educación*, 2022. This report aims to promote lifelong education as a collective project, backed by the commitment to human rights, democratic participation and care for the planet. These objectives advocate for the questioning of absolute truths by citizens and, in doing so, they establish an education system that guarantees the access to accurate information as a basis for a commitment to truth.

In Spain, Organic Law 3/2020, of 29 December, which amends Organic Law 2/2006, of 3 May, on Education (España, Cortes Generales, 2020) expands on UNESCO indications and supports training that encourages critical thinking in primary and secondary education. Furthermore, Royal Decree 157/2022, of 1 March, which establishes the organization and minimum teaching requirements for primary education (España, Ministerio de Educación y Formación Profesional, 2022), introduces the concept of critical thinking linked to the training of competent, autonomous, critical readers who can understand and interpret multimodal texts that help them to meet the challenges of the modern information society. Media and information literacy is considered for this purpose, since its teaching focuses on information search strategies, acknowledgment of authorship, handling different documentary sources and the comparison, organization, critical evaluation and creative communication of information. In this context, the studies by Ennis (2011, 2018) and Kuhn (2018, 2019), among others, are of particular interest.

Ennis (2011, 2018) defines critical thinking as the reasonable, reflective thinking a person adopts regarding a subject on which they need to take some kind of decision. In this sense, citizens should be required to possess the abilities of a critical thinker, such as being well-informed, supporting beliefs on truths, justifying decisions and presenting reasons including points of view that differ from their own, while considering others' feelings.

Kuhn (2018, 2019) understands argumentation as the axis around which critical thinking gravitates and as a fundamentally collective and contextualized social and cultural practice that should be reflected in the socio-educational context, as suggested by Guzmán-Cedillo and Flores-Macías (2020) in their review of 73 studies conducted between 2000 and 2016 in educational settings. In this regard, Bezanilla et al. (2019) have reviewed the main methodologies used to foster critical thinking and suggest that teachers in these settings should employ strategies such as posing questions, choosing activities with a real-life context, using different information sources, stimulating reflective dialogue, referring to visual displays and analyzing the arguments formulated by the community and research. Likewise, Witarsa, and Muhammad (2023) propose a learning model based on inquiry, discovery and problem-solving.

Children's picturebooks are particularly appropriate for developing reading literacy in the early years, the stage when the analog paradigm should be prioritized, since, as the main research

studies show (Wolf, 2018), in-depth reading, which is essential for developing critical thinking, is linked to physical books. From this perspective, an analysis of the evolution of the picturebook toward what is called ‘non-fiction’ (Grilli, 2020, 2021; Tabernero-Sala, 2022; Tabernero-Sala and Laliena, 2023) in reader training is required, as the non-fiction book occupies one of the most vital sectors in the publishing market (Jan, 2021) and has influenced the development of children’s books as a whole (Tabernero-Sala, 2022). In this respect, researchers are beginning to investigate the natural pairing of critical reading and the non-fiction picturebook (Trigo-Ibáñez et al., 2022). This type of children’s book is now the most interesting genre in children’s picturebooks, both nationally and internationally (Smith and Robertson, 2019; Dindelli, 2021; Jan, 2021), due to its artistic nature (Grilli, 2021); it has also become the reading model presented in fictional discourse aimed at children (Goga, 2020; Goga et al., 2021; Kümmerling-Meibauer and Meibauer, 2021).

In keeping with the line begun in the studies by Sander’s (2018) relating to the analysis of non-fiction picturebooks in connection to critical reading, focusing on the analysis of virtual epitexts as one of the strategies to promote reading is appropriate inasmuch that we consider it important to analyze whether mediation fosters critical reading and, if applicable, which strategies the mediator should identify in order to recognize the promotional discourses that favor the development of the critical reader.

Social media have prompted the appearance of public, virtual epitexts that promote reading and have been generated by an inherent need for dissemination of publishing projects. Research studies such as those by Grøn (2014), Lluch et al. (2015), Basaraba (2016), Tabernero-Sala (2016b, 2018), Tabernero-Sala and Calvo-Valios (2016), Ibarra-Rius and Ballester-Roca (2017), Rovira-Collado (2017), Dimova et al. (2018), Lebrun et al. (2018), Romero-Oliva et al. (2019, 2020, 2023), and Bilushchak et al. (2020) emphasize the entity acquired by virtual epitexts used in the promotion of reading, both in disseminating and in proposing reading models that they defend by identifying a new type of social reader who makes essential changes to the construction of discourse. The twenty-first century reader receives information on multimedia devices that involve multimodal discourse. Similarly to Gray (2010), Lluch et al. (2015) emphasize the need to reflect on the importance of these types of promotional epitexts that, far from being auxiliary, paratextual elements—in the sense of the definition of paratexts suggested by Genette (2004)—actually create the text, are part of its identity and determine the meaning that the reader assigns to it by proposing the reading strategy as a type of ‘textual consumption’ (Lluch et al., 2015). In earlier research works (Tabernero-Sala, 2016b), we analyzed in more detail the nature of the book trailer as one of the most significant virtual epitexts, both from the perspective of the promotion of reading and regarding its potential in training the readers of the twenty-first century, in such a way that it may even have its own artistic entity, in accordance with what Unsworth (2015) terms ‘multimodal literary narratives’. In some cases, it may, in fact, become a way of bringing interpretive rewriting closer to the reader (Tabernero-Sala, 2021). Therefore, analysis of virtual epitexts is vitally important to make a detailed study of the new methods of book promotion and how these methods—insofar as they create meaning and guide reading—determine the receiver the discourse

requires. Thus, as a means of promoting reading and books—insofar as they create meaning—virtual epitexts offer the possibility of mediated reading that, in the case of children’s books, involves a reading model with links to the context in which it occurs, so much so that in previous studies (Tabernero-Sala et al., 2022) we have investigated the connection of the book trailer as an epitext for book promotion that defines a reading model in the case of non-fiction picturebooks. We understand that virtual epitexts may have changed in recent years, particularly since the COVID-19 health crisis, when, as several reports prove (Cencerrado Malmierca and Yuste Tuero, 2020; ERI-Lectura, 2020; Sánchez-Muñoz, 2022), reading habits changed and the use of virtual environments reached an unprecedented magnitude.

On this basis, we believed that an updated study on the virtual epitexts of children’s picturebooks would be of interest. Our initial proposal was to investigate and reflect on the new tendencies in the digital promotion of children’s books and to analyze to what extent, in the new cultural ecosystem, the promotion of reading in the area of children’s publishers shows a commitment to training critical citizens, in keeping with the UNESCO reports enshrined in different educational laws. Therefore, using the theoretical framework described and, in line with the studies by Gray (2010), we hereby present this study on the virtual epitexts used to promote children’s picturebooks, with a particular focus on the book trailer as the most widely used and firmly established tool for virtual promotion in the digital environment. The research concentrates on the following objectives:

Objective 1. To define the current tendencies of the virtual epitexts used in the promotion of children’s picturebooks.

Objective 2. To identify the construction strategies in virtual epitexts that are likely to encourage critical reading.

Objective 3. To select a corpus of examples of virtual epitexts as a resource for the promotion of reading and the training of critical readers in socio-educational contexts.

2. Materials and methods

To achieve the research objectives, we conducted a documentary study based on the content analysis of a sample of 836 pieces of digital audio-visual material. These audio-visual documents or videos were published by 45 publishing houses specializing in children’s literature from 1 January 2020—an important year due to the COVID-19 pandemic—to 31 December 2022. The main function of this digital content was to promote the reading of picturebooks aimed at children, so they are considered virtual epitexts of the books (Tabernero-Sala, 2016b).

The process of selecting the 45 publishing houses and the 836 virtual epitexts consisted of three stages:

- Firstly, 68 publishers were selected as being of interest for the study. The main sources consulted in this stage were the directory of Spanish publishers specializing in children’s and young people’s books published by the *Federación de Gremios de Editores de España*, n.d., the publishing houses belonging to *Asociación Álbum*, n.d. and the *Biblioteca Nacional de España*, n.d. of specialist publishers. For this initial selection, two essential assessment criteria were considered: publishing

in Spanish and being well-known publishers in the area of children's literature.

- The second stage consisted of a review of the digital promotion methods used by the publishers in 2020, 2021 and 2022. The selection of the years responded to the first objective of the study—to define the current tendencies in digital promotion—and we were also interested in observing how these tendencies have evolved since the beginning of the COVID-19 pandemic. Different studies (Cencerrado Malmierca and Yuste Tuero, 2020; Sánchez-Muñoz, 2022) have explored how the pandemic has hindered access to books due to the closure of bookshops and libraries during 2020 or to social distancing, which affected reading in classrooms. We decided to analyze how publishers developed digital promotion of children's picturebooks at that time, since, as shown in Figure 1, the number of publications on YouTube and Vimeo in 2020 increased. As for the promotion methods, we confirmed that the main, most stable method, which acted as content repository, was YouTube (Tabernero-Sala et al., 2022), and, to a lesser extent, Vimeo. Furthermore, publishers generally ran promotion on different social media, such as Instagram or Facebook, due to the immediate nature of these media and their great capacity for dissemination. The publications uploaded onto these media, however, posed problems related to their registration, such as frequent repetitions within the same document and instability. At the same time, we also observed that videos previously hosted on YouTube and Vimeo were usually disseminated through these media, so the study eventually focused on these two platforms. In this way, the final selection of publishers was adjusted according to the presence of the publisher on YouTube or Vimeo between 2020 and 2022. Table 1 shows the quantitative data of the selection process and the adjustment of the publishers selected. The sample was reduced from 68 to 45 publishing houses as some of the 68 publishers initially selected did not have a YouTube or Vimeo channel, others that did have one had not used it in the period from 2020 to 2022. These were, primarily, small, independent publishers, representing projects run by one or two people, usually involving a more limited capacity for media outreach. Additionally, certain large publishing houses had, in fact, used these channels but they had no audio-visual content aimed at promoting children's picturebooks; in other words, some of the publishers selected had also published books for adult readers, so, in some cases, their promotional videos focused solely on this sector of the population, and consequently, they were excluded from the selection process if they did not have any promotional videos for children's books. Table 2 presents the 45 publishing houses that were finally included in the research.
- The third stage of the document selection process consisted of a review of all the audio-visual publications on the official YouTube and Vimeo channels belonging to the 45 publishing houses—3,358 documents in total. This review identified the virtual epitexts used in the promotion of children's picturebooks. Through this process, other types of content by some publishing houses were dismissed. The videos that were discarded from the study were those that promoted textbooks intended for the context of formal education or

books aimed at young adult or adult audiences; we also discarded videos that presented pedagogical and educational conferences or videos that promoted other merchandising products, such as toys or videogames. In turn, videos aimed solely at the promotion of children's picturebooks, whether they were works of fiction or non-fiction, were included in the selection, excluding textbooks; we selected book trailers with a cinematographic style, videos displaying the book as an object that only showed the book or part of the book, videos in which writers, editors, book-sellers, librarians, readers or other mediators presented, commented on or read the book, or videos that showed activities or the publishers celebrating an event about this type of book (Figure 2); in other words, this study included audio-visual content aimed at promoting children's books, although we found that the promotional styles and the strategies used were varied and they merged together, thus revealing some significant tendencies. The analysis of this aspect was of interest to this study and the results are presented below. Lastly, considering these criteria, 836 audio-visual documents were selected for study. Table 3 lists the audio-visual documents published each year by the group of 45 publishers and the number of virtual epitexts that were selected for the study.

Once the selection process had been completed, each publisher's digital channel and the videos published were systematically registered arranged by date. This stage involved the registration of the quantitative data of interest for the research objectives and for the content analysis (Krippendorff, 2019). These data included the date of creation of the channel, the number of subscribers, the number of total views on the channel, the date each video was uploaded to the internet, the duration of each video, the number of views up to the date of the analysis and the corresponding links. Table 4 shows the registration sheet used for each of the 45 publishers.

Following the registration of the audio-visual documents, a mainly qualitative content analysis was conducted focusing on the constructive and rhetorical strategies in the audio-visual discourse typical of digital book promotion (Basaraba, 2016). The content analysis was based on a process of categorization, coding, analysis, refining of categories and qualitative interpretation, using a narrative approach based on the postulates of the school of critical theory, as indicated by Ruiz-Olabuénaga (2009). The results of the analysis according to the research objectives are presented below.

3. Results

3.1. Current tendencies in digital promotion

As regards the first objective, the analysis of the audio-visual documents detected five emergent macrocategories. These macrocategories identify the most frequent components currently emphasized by the virtual epitexts analyzed and help to classify the videos according to their constructive model and to determine the main tendencies in digital promotion of children's picturebooks. The macrocategories detected are listed and described in Table 5.

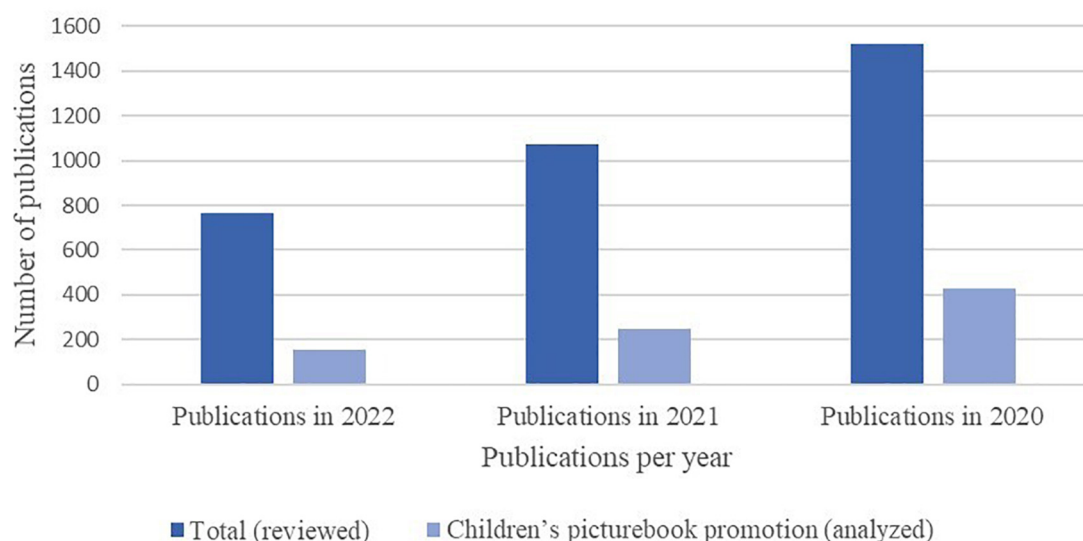


FIGURE 1

Audio-visual documents published on YouTube or Vimeo by publishers.

Figure 3 presents, in absolute values, the incidence of these macrocategories in the virtual epitexts analyzed. The macrocategories are not mutually exclusive, as several may be present in the same video.

3.1.1. Emphasis on the materiality of the book

Virtual epitexts that affect the materiality of the book is as a dominant tendency on some publishing channels. This type of video presents the book as an object, with an emphasis on its paratexts, mainly illustrations or other elements such as book covers, foldouts and pop-up components. Therefore, the picturebook aimed at children tends to be displayed as a promotional strategy, with an emphasis on its physical features, in contrast to other genres, such as the novel, predominantly focused on protecting the work. Promotion is based, therefore,

on encouraging child readers to want to interact with the book and enjoy touching it, handling it and discovering what is hidden in its pages. The study identifies publishers that base promotion on a display of the book, such as Akiara Books, Apila, Edelvives and Zahorí Books. Other publishers have opted for this form of promotion at specific times, as was the case with Andana in 2021 or Penguin España in 2022.

Among the virtual epitexts that focus on the book as an object, differences can also be observed regarding their construction strategies. In this respect, videos usually concentrate on the turning of the pages: as the fast-motion turning of all the pages in the book employed by Apila or Nuevo Nueve; the leisurely page-turning used by Zahorí Books, Andana or some of the videos by Maeva or Tecolote; the page-turning that pauses to explore and handle the book, with zoom effects to emphasize the paratexts or certain content, such as in the videos by Edelvives; or some less frequent artistic approaches, such as some examples by A buen paso or Kókinos. Other publishers attempt to develop their own style, such as the recent publications by Penguin España, very short videos that present the book and some of its pages in a clearly edited format or the virtual epitexts from Akiara Books, which are contextualized in a natural environment and carefully edited as far as esthetics and the senses are concerned (Table 6).

3.1.2. Emphasis on video editing and the esthetic component

Animated book trailers, characterized by their cinematographic style (Tabernero-Sala, 2016b) and their emphasis on the esthetic component are among the virtual epitexts analyzed. The most outstanding examples stand out for replacing the still image, typical of a book, with a moving image, typical of the cinema, including a soundtrack and creating a story with its own artistic entity. These elements are usually linked to a storyline synopsis and address the reader directly through questions about the continuity of the story. These resources aim to create an intriguing context, arouse curiosity and thereby encourage reading of the book. These videos

TABLE 1 Process of selecting publishers.

Publisher selection	Number of publishers
Initial selection of publishers specializing in children's literature	68
Publishers without a YouTube or Vimeo channel	15
Publishers with a YouTube or Vimeo channel without any activity between 2020 and 2022	7
Publishers with a YouTube channel with activity between 2020 and 2022	44
Publishers with a Vimeo channel with activity between 2020 and 2022	2
Publishers with a YouTube or Vimeo channel with activity between 2020 and 2022, but without any publications of interest for the purpose of this study	1
Final selection of publishers (YouTube or Vimeo channel and publications relevant to the purpose of this study between 2020 and 2022)	45

TABLE 2 Publishers in the study sample.

1	A buen paso	16	Edelvives	31	Litera
2	A fin de cuentos	17	Ediciones Castillo	32	López
3	Akiara books	18	Ediciones El Naranjo	33	Maeva
4	Algar	19	Ekaré	34	Nórdica
5	Amanuta	20	Flamboyant	35	Nube Ocho
6	Anaya	21	Fondo de Cultura Económica	36	Nuevo Nueve
7	Andana	22	Galimatazo	37	Océano Travesía
8	Apila	23	Iamiqué	38	Penguin España
9	Babel Libros	24	Impedimenta	39	Pequeño Editor
10	Bookolia	25	Juventud	40	Pintar-Pintar
11	Carambuco	26	Kalandraka	41	Silonia
12	Coco Books	27	Kókinos	42	Takatuka
13	Combel	28	Lata de sal	43	Tecolote
14	Cuatro Azules	29	Libre Albedrío	44	Wonder Ponder
15	Diego Pun	30	Libros del Zorro Rojo	45	Zahorí Books

require careful, professional design and editing, which involves financial investment. Therefore, few publishing houses maintain the quality in this type of digital promotion, as it is increasingly being replaced by ‘homemade’ videos and live recordings without any editing. Publishers still opting for book trailers include A Fin de Cuentos, Kalandraka, Libre Albedrío and Libros del Zorro Rojo, whose videos are defined by their artistic quality. Other publishers such as Amanuta also used to publish professional animated book trailers, but have almost ceased to do so or use them exceptionally for certain outstanding works, as this type of video is not usually the main focus of their digital channel. This is the case of Akiara Books, Combel, Edelvives, or Zahorí Books. As for their duration, they last one minute on average, although there was a general tendency in 2022 to reduce the length of these videos—except for a few significant exceptions—possibly due to their dissemination on social media such as Instagram—where speed and brevity are the priority—and to their cost. [Table 7](#) presents some animated and cinematic-style book trailers.

3.1.3. Emphasis on the interpretation or construction process of the book

The presence of authors and editors in virtual publications is a tendency shared Most of the publishers’ channels analyzed. The reason for their appearance may be to read the book, as we will see but, most usually, they appear to present books, promote reading, provide an interpretive commentary or reflect on the creative process. There are several different formats of virtual epitexts that emphasize these aspects according to, above all, the degree of complexity of the editing. On one hand, there are extensive formats, such as interviews, conferences and events. These videos are frequently recorded live with the participation of different guests or they even offer open access to the audience. Solo recordings of the authors commenting on their work are also common. The channels of the publishing houses A buen paso, Babel Libros, Combel, Diego Pun, Ediciones El Naranjo, Fondo de Cultura Económica, Nórdica or Wonder Ponder, among others, contain

this type of publication, mainly intended for adult mediators, although content aimed at children—such as some videos by Combel—can also be found. The publisher Takatuka launched the initiative ‘Las librerías recomiendan,’ a series of videos in which different booksellers recommend and comment on a book by this publisher. The frequency of these publications is remarkable, once again, in 2020. They commonly contain references to the lockdown and the benefits of reading in this context. On the other hand, more elaborate formats requiring editing are also published, showing the authors’ workspaces or following the development of their creative process. This type of video is less common but publishers such as Andana, Edelvives, Pequeño Editor, Penguin España or Takatuka have some noteworthy titles ([Table 8](#)).

3.1.4. Emphasis on the reading of the book

A marked tendency, mainly in 2020, is that of digital publications aimed at the reading of the book. In them, authors, editors, youtubers and other adult mediators read, sing or tell a story. A significant example is the ‘cuentacuentos’ (storyteller) initiative by the publisher Juventud, with 32 publications between April and July 2020. In these videos, the pages are displayed while a voiceover reads the text. Other publishing houses take the same line. For example, Carambuco provides publications both in an oral format and in sign language while displaying animated scenes from the book. In 2020 as well, Combel launched ‘cuentacombel,’ a project that offered, among others, readings of chapters of *Las aventuras de Pinocho* with the participation of different writers. Publishing houses such as López, Océano Travesía, Anaya, Fondo de Cultura Económica or Pintar-Pintar issued publications with the same purpose. On many occasions, the authors of the works or other mediators record the reading at home, as occurred during the lockdown. These initiatives did not continue after 2020, but they helped to boost a type of promotional video—characterized by its spontaneity—that has been used since then by many of the publishers analyzed. Furthermore, there are also edited virtual epitexts, of a more artistic nature, in which animation or the

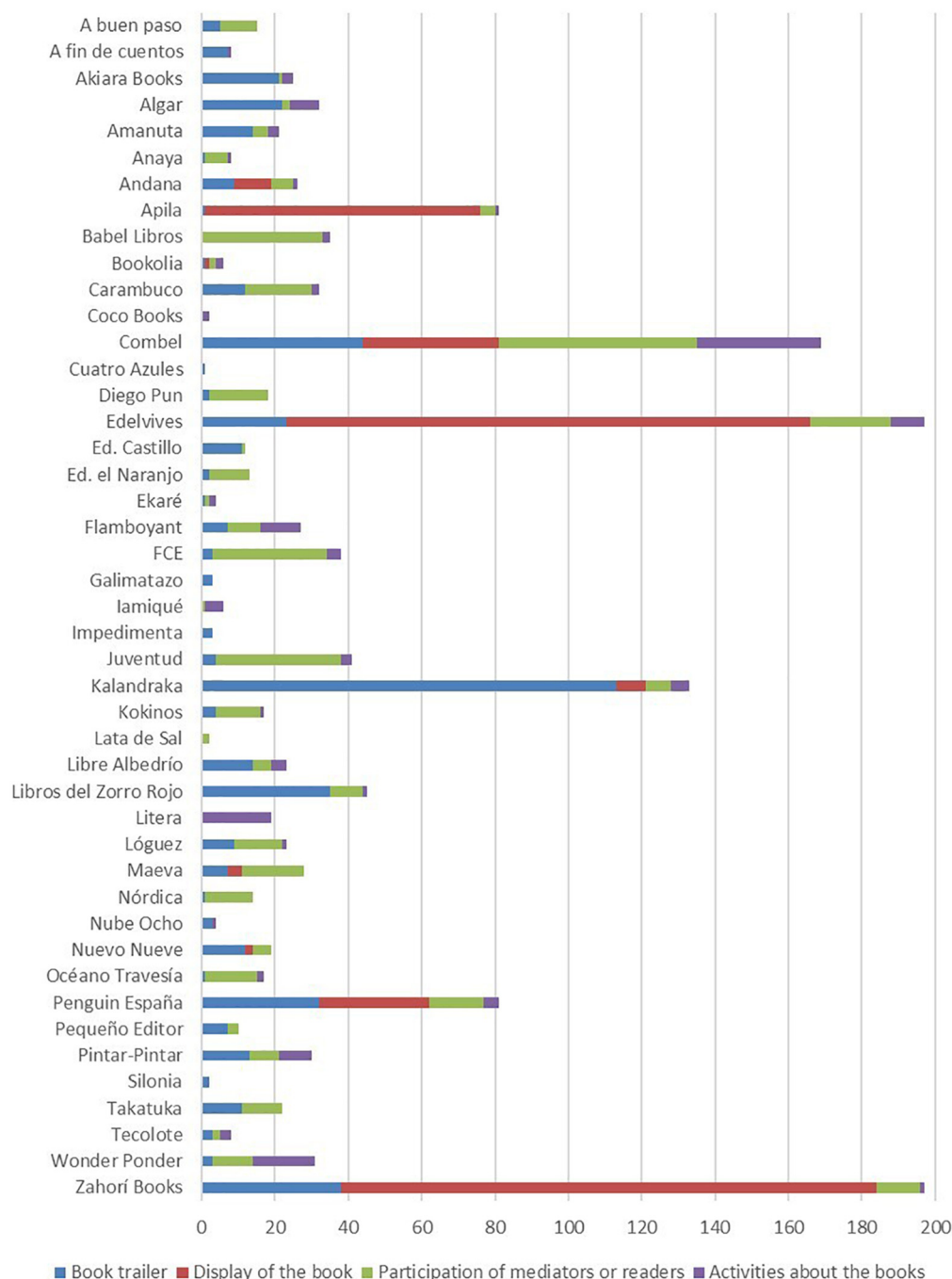


FIGURE 2

Type of videos for the promotion of children's picturebooks selected, by publisher.

contemplation of the book are complemented by a voice telling part of the story, as is the case in some of the videos by Kalandraka or Ediciones Castillo. **Table 9** provides examples of formats that emphasize the reading of the book.

3.1.5. Emphasis on the reception of the book and its receiver

Another emerging tendency, a minority in quantitative terms but still of interest to this study, is the presence of child mediators who read, give opinions, recommend or interact with the books.

In these videos, the publisher's promotion relies on the receivers and peer dialogue, by trying to arouse empathy in the child audience. Similarly, some videos point out how to use a book, contain activities linked to the work or show its ludic and creative possibilities apart from reading it, even referring to other virtual epitexts, such as interactive games. These virtual epitexts emphasize the reception process of the book and offer replicable models for child readers; in other words, they provide ideas and proposals that extend the reading of the work and boost its potentiality. Again, there are two types among the publications showing this tendency.

TABLE 3 Audio-visual documents reviewed and documents selected.

Audio-visual documents reviewed		Audio-visual documents selected for analysis (virtual epitexts)	
Total publications in 2022	765	Publications selected from 2022	157
Total publications in 2021	1,078	Publications selected from 2021	252
Total publications in 2020	1,518	Publications selected from 2020	427
Total publications reviewed	3,361	Total publications analyzed	836

TABLE 4 Register sheet for audio-visual documents.

Name of publisher		Register date of audio-visual data		
Name of YouTube or Vimeo channel	Channel creation date	Number of channel subscribers	Total number of channel views	Channel link
Number of total publications by publisher in 2022		Number of publications selected from 2022 whose objective is the promotion of children’s picturebooks		
Number of total publications by publisher in 2021		Number of publications selected from 2021 whose objective is the promotion of children’s picturebooks		
Number of total publications by publisher in 2020		Number of publications selected from 2021 whose objective is the promotion of children’s picturebooks		
Total number of publications by publisher		Number of publications selected for analysis		
Register of each of the virtual epitexts selected (one row per video)				
Video title	Date of video upload	Video duration	Video views	Video link

On one hand, there are more spontaneous videos, which consist of everyday recordings intended for immediate dissemination, for instance, some publications by Fondo de Cultura Económica or the publisher Pintar-Pintar, or the initiative by Wonder Ponder designed in 2020, which encourages children to participate by explaining how they felt during lockdown. On the other hand, although less frequently, there are more elaborate videos in which promotion is the underlying purpose. Their construction involves design, editing, animation and attention to the esthetics of the virtual epitext, as in some significant examples by the publishers Kalandraka or Zahorí Books (Table 10).

3.2. Strategies for promoting critical reading

After exploring the current tendencies that define the virtual epitexts selected, the analysis then focused on identifying the discourse markers and strategies likely to encourage critical reading, while also considering the theoretical framework established previously. To this effect, a selective analysis of a qualitative and interpretive nature was conducted with the purpose of fulfilling study objectives 2 and 3; in other words, in this case, the interest of the study lay not so much in the frequency with which a specific discourse marker appeared in the virtual epitexts, but rather in identifying, in the wide range of publications previously described, the discourse markers present in the virtual epitexts most likely to promote critical reading. To this effect, we separated and analyzed the discourse markers that identified strategies that promoted a type of reading focused on cracks in the discourse, on questioning the sources and developing empathy as a method to raise awareness resulting from access to knowledge (Ennis, 2011, 2018; Kuhn, 2018, 2019; Bezanilla et al., 2019). The analysis was complemented by a selection of virtual epitexts with comments, which are suggested, ultimately, as resources that may be of interest to mediators and readers

in promoting reading and training critical readers in socio-educational contexts. The virtual epitexts in the final selection promote books that should also be recommended for classroom reading in order to favor the development of biliterate critical reading by merging analog and digital formats (Wolf, 2018). Table 11 shows the discourse markers and strategies identified by this second analysis that are subsequently commented and exemplified.

3.2.1. The visible author as a metadiscursive strategy

As confirmed by the tendency analysis, one of the recurring markers in virtual epitexts used for the promotion of children's books is the presence of the author. Documents containing examples of this discourse marker can be found in promotions such as those by the publisher A buen paso for *Semillas. Un Pequeño gran viaje* by Alonso and Paschetta (2018) and *Zum zum. El viaje de la semilla* by Ferrada and Paschetta (2021). In this case, it involves a conversation between the editor and the

TABLE 5 Emergent macrocategories.

Macrocategories	Description
Emphasis on the materiality of the book	Videos focusing on displaying the book as an object, on page-turning and paratexts: illustrations, book covers, foldouts, pop-ups.
Emphasis on the video editing and the esthetic component	Animated and cinematic-style book trailers
Emphasis on the interpretation of the work and/or its construction process	Participation by authors, editors, booksellers and other adult mediators to present a book, provide comments or reflect on the creative process
Emphasis on the reading of the book	Reading aloud, songs, storytellers, oral narration, performed by adult mediators
Emphasis on the reception of the book and its receiver	Participation by child readers, activities suggested in the book, extension of the work.

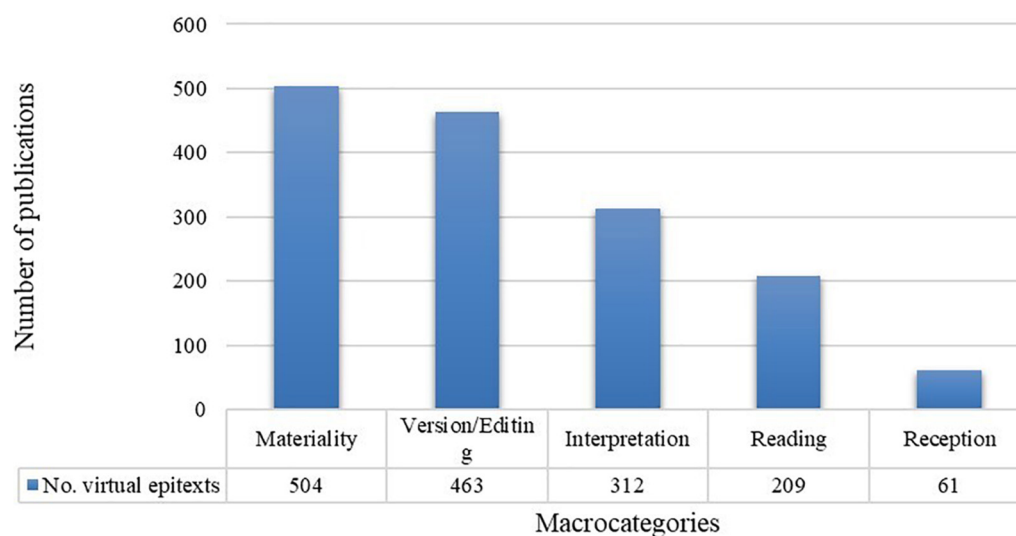


FIGURE 3
Macrocategories detected in the virtual epitexts.

illustrator of the two works regarding the creative processes.¹ This provides an explanation about the universe in which the illustrator works according to the publisher's proposal. Similarly, Jeffers (2019) appears in the book trailer for *El destino de Fausto*, a classic-style book trailer with sophisticated animation.² Along the same lines, Pacheco (2020) explains his interpretation process regarding the work by Sister Juana Inés de la Cruz in *Caperucita Roja*, *Primero sueño*,³ Goes (2016) develops his creative activity⁴ or Dautremere (2021) demonstrates her creative process in her workspace, without saying a word, in *Tan solo un instante*.⁵ In these cases, the visibility of the people behind the books increases, which leads to the personalization of a discourse that does not come from the anonymity typical of children's books, but rather from a discourse with its own ideology that corresponds to the author's 'world view.' Therefore, whether fictional or non-fictional, the discourse is presented as a subjective proposal that responds to the bias of the person who has created it and that may, in this way, be questioned by the reader, who is required to form an opinion. More specifically, certain epitexts emerge as paradigmatic, such as those related to the recent promotion of two particularly relevant picturebooks: *Dos ositos* by Ylla (2021) and *Ernesto el elefante* by Browne (2021). The virtual promotion document of *Dos ositos*⁶ follows the style of a traditional book trailer in the first part—using a narrative voiceover with moving images of the illustrations in the foreground, which are photographs of the protagonists—and, the second, shorter part presents the author, her particular features, her life plan and her image contextualized in the time the work was created. Thus, the

reader and mediator are presented with the directions the dialogue may subsequently take in the work, awareness of sustainable development and love of animals and the use of documentary images, which ensures the hybridization of reality and fiction. To a lesser extent, the promotional book trailer of *Ernesto el elefante*,⁷ incorporates the figure of the author to contextualize the proposal and suggest contemplation of the work within the framework of an entire authorial universe with an ideological bias.

On these lines, we recover a study by Sander's (2018) on the connection of the non-fiction book with critical reading, since we understand that his research may apply, to some extent, to the analysis of children's books, whether fictional or not, in a critical sense. In his study *A literature of questions. Non-fiction for the Critical Child*, Sander's (2018) defines the genre as a literature of questions rather than answers requiring a critical reader to construct its meaning.

Sanders analyses the non-fiction book in terms of the presentation of information, of the bias authors place on the book simply by selecting some pieces of information to the detriment of others; he also investigates the esthetic strategies that invite readers to become involved in the texts, to interact with them, to engage with the information in a reflective, dialogic way. The intention is not to present information for it to be merely absorbed, but rather to encourage readers to investigate for themselves by arousing their curiosity and wonder in relation to what is presented.

These works propose critical readers who call into question what is presented to them, have an opinion on what they are shown, notice cracks in the discourse and question the credibility of the information. Based on ideas Barthe's (1994) on the 'death of the author' and Freir's (2002), Freire (2010) premises regarding critical pedagogy, along with Bakhtin's (1981) dialogic concept of the novel, Sander's (2018) identifies the critical reader in works with cracks

1 https://youtu.be/UBU0B9GKy_g

2 <https://www.youtube.com/watch?v=TDSPN4JK6dc>

3 <https://youtu.be/mB9YTBW-N2s>; <https://youtu.be/WyRzjtFfa-Y>; <https://youtu.be/lx0H8n0rqYI>

4 <https://www.youtube.com/watch?v=2lg4JPGWLWw>

5 <https://www.youtube.com/watch?v=nRWtaTr3lpk>

6 <https://www.youtube.com/watch?v=tOMsWpkXEI>

7 <https://www.youtube.com/watch?v=i860KJY0h9Y>

in the authority of the text and questions that guide the process of intellectual inquiry. Therefore, centralization of meaning in a single, authorized source of true knowledge is questioned. Using the markers that Sanders identifies as the ‘visible author,’ the discourse presents the story using the voice, attitude and viewpoint of the author, who becomes part of the story through their ideas and opinions. Although, rhetorically, the interpretation may seem to be conditioned, Sanders bases his assertions on the research by Paxton and Zarnowski (Sander’s, 2018, p. 58), who showed, at the time, that readers’ interactive responses were more likely to occur when the author was present.

3.2.2. Empathic projection of the author on the discourse

Along the same lines, with an emphasis on metadiscursive visibility, book presentation from an emphatic perspective acquires particular relevance. In this way, readers identify themselves through exposure to the authors’ own experiences, whether they are writers or illustrators. This is the case, for example, in the promotional epitext for *El patito feo* (Andersen and Abramović, 2021), in which the illustrator, Marina Abramović, describes her experiential process of reading Andersen’s (2021) work⁸ or the promotional epitext for *Muñeco de barro*, in which the illustrator, Carme Solé-Vendrell (Reyes and Solé-Vendrell, 2020), talks about her personal interpretation of the characters and their surroundings.⁹ Also relating to awareness-raising, the editor of Libros del Zorro Rojo discusses *Hablo como el río* by Scott and Smith (2021), suggesting the lines of dialogue the story opens.¹⁰

3.2.3. Experiential reading: Between reality and fiction

In recent years, promotional virtual epitexts that defend removing the boundaries between fiction and non-fiction have taken the same direction. The intention is, ultimately, to relate reading to the reader’s personal experience and environment. Therefore, they show documents that encourage an emotional and physical connection between the books and the contents proposed. Thus, for example, in cultural and ideological terms, book trailers contain very clearly identified spaces and music. The book trailer for *¡Gracias, madre Tierra!*, illustrated by Starkoff (2022), based on the Haudenosaunee Thanksgiving Address is a clear example.¹¹ It is the translation, in foldout format, of a remarkable prayer offering a greeting and gratitude to all living beings, which has been recited every morning for many centuries by the whole community in a territory located between the United States and Canada. This defines the book as an inventory of the natural world and the wonders that nature offers. The natural backdrop against which the work unfolds establishes a relationship between the book and its physical surroundings by reflecting on fictional and artistic discourses and their natural extension along the lines proposed by the Sustainable Development Goals contained in the 2030 Agenda and the report *Reimagining our Futures Together: A New Social Contract for Education*, prepared by the Comisión Internacional

sobre los Futuros de la Educación, 2022. The book trailer for *Tiempo de haikus* by Santaaulàlia and Lozano (2022)¹² contains the same theme and promotional concept.

3.2.4. The author, the reader, the book and the environment

In the register of experiential reading, there are promotional proposals that address the reader through questions raised directly by the works and generate interpretations of the setting by concealing the boundaries between the book and reality through the use of real scenarios where the act of reading occurs. An example of this proposal is the book trailer for the collection *¿Quién soy?* by Seceda and García (2021)¹³ where narrators’ and characters’ voices combine in open spaces in which fiction intertwines with readers’ realities and the contents are revealed as being part of receivers’ own lives. In this sense, the promotional epitext of *Brujas, guerreras y diosas* by Hodges and Lee-Merrion (2020) is paradigmatic, as it is a document structured around testimonies from female readers who have chosen stories about fairies, vampires, sorceresses and goddesses with whom they identify and present them as contemporary role-models. Therefore, the reader’s reality and the fictional discourse are connected, thus removing the boundaries between reality and fiction. The book trailer for the picturebook *¿Debo argumentar el sinsentido de la esclavitud?* by Frederick Douglass, with commentary by Squilloni and Fosch (2021) and translated by Jordi Pigem and Catarina Sacramento is of particular interest.¹⁴ It concerns the speech given by the ex-slave Frederick Douglass on Independence Day in the United States. The appeal to the reader, inherent in its ironic content, is made more powerful through the voices of the young black people who read aloud the first sentences of this rhetorical work of art while looking straight at the camera.

3.2.5. From the materiality of the discourse to in-depth reading

Another of the discourse markers present in promotional virtual epitexts in recent years is the relevance that the physicality of the act of reading acquires through the conception of the book as an object. While in previous studies (Tabernero-Sala, 2016b, 2019) we insisted on the tendency to emphasize the materiality of the book as one of the categories used to define the book trailer, in the last three years this category has become so usual that it has been identified as a dominant tendency, which may be explained by the digital environment with which people engage. Thus, book trailers such as those relating to the promotion of *Bienvenida* by Comín (2021a),¹⁵ *Mi arbolito de Navidad* by Comín (2021b),¹⁶ *Los volcanes* by Geis (2022)¹⁷ or *¡No chupes este libro!* by Ben-Barak and Frost (2020)¹⁸ combine the dimension of the work as an object with specific references to its physical

⁸ <https://www.youtube.com/watch?v=fK4dVVAL45g>

⁹ https://www.youtube.com/watch?v=sDp_ZioYU-4

¹⁰ https://www.youtube.com/watch?v=OPQpuzlv_Ho

¹¹ <https://vimeo.com/779672069>

¹² <https://vimeo.com/738779606>

¹³ <https://www.youtube.com/watch?v=ykvn6oebpLY>

¹⁴ <https://vimeo.com/657981925>

¹⁵ <https://www.youtube.com/watch?v=RpwKjIPie68>

¹⁶ <https://www.youtube.com/watch?v=7VS36HTYjEg>

¹⁷ <https://www.youtube.com/watch?v=Hia7Xel1FOW>

¹⁸ <https://www.youtube.com/watch?v=Pcyf1FKLz7A>

TABLE 6 Examples of virtual epitexts based on the book's materiality.

Book title	Publisher	Date of upload to channel	Duration	Views (to date indicated)
<i>¡Gracias, Madre Tierra!</i>	Akiara Books	09/12/2022	00:01:36	191 (11/01/2023)
Link	https://vimeo.com/779672069			
<i>Bajo el mar</i>	Penguin España	24/10/2022	00:00:29	62 (09/01/2023)
Link	https://youtu.be/prDtVgbuJU			
<i>Reinos minúsculos</i>	Edelvives	15/11/2022	00:01:00	571 (06/01/2023)
Link	https://youtu.be/EoePirPE1A4			
<i>Bienvenida</i>	A buen paso	22/02/2021	00:01:03	777 (29/12/2022)
Link	https://youtu.be/RpwKjIPie68			
<i>¡Ves lo que yo veo?</i>	Tecolote	30/09/2020	00:00:47	67 (03/01/2023)
Link	https://youtu.be/P4K9ujte2eM			
<i>Pompas de jabón</i>	Kókinos	23/12/2020	00:01:16	391 (11/01/2023)
Link	https://vimeo.com/494118535			

TABLE 7 Examples of animated and cinematic-style book trailers.

Book title	Publisher	Date of upload to channel	Duration	Views (to date indicated)
<i>La Reina de las Nieves</i>	Edelvives	20/12/2022	00:00:34	212 (06/01/2023)
Link	https://youtu.be/JztN1YIMLeg			
<i>Bibliotecarias a caballo</i>	A fin de cuentos	09/12/2022	00:00:36	13 (30/12/2022)
Link	https://youtu.be/gcb8BkdXsrg			
<i>Con esta línea</i>	Combel	04/11/2022	00:01:03	298 (30/12/2022)
Link	https://youtu.be/ib969TsDK8I			
<i>Circo</i>	Kalandraka	04/04/2022	00:02:22	197 (11/01/2023)
Link	https://youtu.be/S2vwWiHNb6g			
<i>Bajo las piedras</i>	Akiara Books	01/07/2020	00:01:58	721 (11/01/2023)
Link	https://vimeo.com/434288011			
<i>Travesía</i>	Libros del Zorro Rojo	09/06/2020	00:01:28	1,897 (29/12/2022)
Link	https://youtu.be/CZGkTGfPqb0			

handling by the reader, at a leisurely pace that is a delight to the senses. This material dimension is part of the identification of in-depth reading to the development of the critical reader, involving the defense, as recent research recommends (Wolf, 2018; Zafra, 2019; Gil-Pelluch et al., 2020; Støle, 2020) of the time to think and memory retention that paper affords in contrast to the immediacy of the characteristic 'skim reading' of screens. The study conducted by Zafra (2017, p. 19), among others, refers to reading online every day as an experience based more 'on impressions than on concentration,' more on screenshots than on deliberate reflection. Zafra (2019) consequently defends the notion of 'time to think' as the only way to achieve autonomous, critical thinking.

3.2.6. The sensorial experience. The narrator's voice

This register also includes the importance attributed to the narrator's voice by the different promotional epitexts, whether it is the actual author, writer/illustrator, or oral narrators who read

aloud in spaces physically suited to the content of the book. There is an increased presence of documents in which the narrator's voice becomes one of the key elements of the promotional proposal. In this manner, book trailers such as the one for the picturebook *¡Artista!*, by Hernández-Sevillano and Cerro-Rico (2019),¹⁹ which presents a dramatized recitation of the text over the images, or the reading aloud of *Nunca dejes de brillar* by Alonso and Muñoz (2020)²⁰ may serve as examples of pronounced discourse markers in the promotion of reading for children since the year 2020, closely linked to the necessity of defending reading as a sensorial experience.

3.2.7. The multimodal discourse

Cinematographic book trailers that develop a multimodal discourse in which music, image, animation and voice produce artistic documents with an entity of their own follow the same

¹⁹ <https://www.youtube.com/watch?v=BW8lyRWH7O0>

²⁰ <https://www.youtube.com/watch?v=kPttxOHEMWE>

TABLE 8 Examples of virtual epitexts focusing on understanding and the creative process.

Book title	Publisher	Date of upload to channel	Duration	Views (to date indicated)
<i>Martin y la nuez inolvidable</i>	Fondo de Cultura Económica	16/11/2022	00:04:05	143 (08/01/2023)
Link	https://youtu.be/Lp1jEybJpXc			
<i>El Elefante</i>	Nórdica	29/04/2022	00:01:28	88 (03/01/2023)
Link	https://youtu.be/KEtTpOEd_IE			
<i>Tan solo un instante</i>	Edelvives	15/11/2021	00:01:17	5,838 (06/01/2023)
Link	https://youtu.be/nRWtaTr3Ipk			
<i>iNoche, toca los platillos!</i>	Pequeño Editor	02/08/2021	00:03:04	7,634 (03/01/2023)
Link	https://youtu.be/MuY46jHtE44			
<i>El destino de Fausto</i>	Andana	17/09/2020	00:02:13	1,338 (30/12/2020)
Link	https://youtu.be/TDSPN4JK6dc			
<i>La bañera</i>	Takatuka	09/07/2020	00:01:02	92 (03/01/2023)
Link	https://youtu.be/cXULdqSXRw			

TABLE 9 Examples of virtual epitexts focusing on the reading of the book.

Book title	Publisher	Date of upload to channel	Duration	Views (to date indicated)
<i>La pirata Guiomar</i>	Maeva	24/05/2022	00:08:46	42 (02/01/2023)
Link	https://youtu.be/fLOgNGC3jxA			
<i>Dos ositos</i>	Kalandraka	03/12/2021	00:01:35	243 (11/01/2023/)
Link	https://youtu.be/tOMsWpkXELI			
<i>Sol solito</i>	Libre Albedrío	19/03/2021	00:00:37	1,186 (09/01/2023)
Link	https://youtu.be/7sZw5Q3f2-c			
<i>Palmir</i>	Ediciones Castillo	05/02/2021	00:01:07	89 (02/01/2023)
Link	https://youtu.be/DPjgKAMeSe8			
<i>La oficina de objetos perdidos</i>	Juventud	01/05/2020	00:06:04	1,905 (30/12/2022)
Link	https://youtu.be/4CcIyS9cp24			
<i>iQué lío cósmico!</i>	Carambuco	07/04/2020	00:10:09	1,593 (04/01/2023)
Link	https://youtu.be/zQCR7vJSTII			

direction (Unsworth, 2015). Thereby, book trailers such as those for *Bambi* by Benjamin Lacombe (Saltén and Lacombe, 2020),²¹ *El bolero de Ravel* by Abad and Delicado (2020),²² *Shh... Tenemos un plan* by Haughton (2019)²³ or *Abecedario* by Kaufman and Franco (2017), Letra Ñ²⁴ use ellipsis, metaphors and appeals to the reader in their composition to move readers and interact with them.

4. Discussion

In terms of specialized publishers' promotion of children's books, the analysis presented reveals that, in recent years,

virtual epitexts have evolved in proportion to the new challenges posed by the digital society regarding reader training (Cordón-García, 2018, 2020; Lluich, 2018). The data examined from the sample selected reveal an increase in the number of virtual documents created by publishers relating to the dissemination of children's books in recent years, with the peak incidence rate in 2020, possibly due to the changes produced by the COVID-19 health crisis. This fact is consistent with the results of the reports on post-pandemic reading habits among young people and adults (Cencerrado Malmierca and Yuste Tuero, 2020; ERI-Lectura, 2020; Sánchez-Muñoz, 2022). Similarly, the presence of the book trailer as a document to promote reading has decreased in comparison to less elaborate, more immediate virtual materials, as required by social media, where the ephemeral prevails over anything requiring a longer production time with the consequent economic costs involved. In response to the speed that characterizes the digital

²¹ <https://www.youtube.com/watch?v=tT1KHo2ZJoQ>

²² https://www.youtube.com/watch?v=GuDC_Zgtfik

²³ <https://www.youtube.com/watch?v=7jn2dGrku5k>

²⁴ <https://www.youtube.com/watch?v=VFy7hGS48hE>

TABLE 10 Examples of virtual epitexts that emphasize the role of the child reader and extend the reception process.

Book title	Publisher	Date of upload to channel	Duration	Views (to date indicated)
<i>¿Quién soy? Crías de animales</i>	Kalandraka	29/07/2022	00:02:20	69 (11/01/2022)
Link	https://youtu.be/ykvn6oebpLY			
<i>Agua y Tierra</i>	Amanuta	16/05/2022	00:03:45	33 (30/12/2022)
Link	https://youtu.be/ZjbkKPIfox0			
<i>Ppprrrrriit</i>	Zahorí Books	15/02/2022	00:01:05	125 (04/01/2023)
Link	https://youtu.be/3XGCXHWi2CA			
<i>Mi gran granja</i>	Lóquez	02/07/2021	00:02:40	56 (02/01/2023)
Link	https://youtu.be/gifZb79vr_M			
<i>Secreto de familia</i>	Fondo de Cultura Económica	09/04/2020	00:05:05	153 (08/01/2023)
Link	https://youtu.be/KxFy80fStLI			
<i>Etenko y los patines maravillosos</i>	Pintar-Pintar	07/04/2020	00:02:29	64 (04/01/2023)
Link	https://youtu.be/rdUnRsLDHWY			

society, only a few publishers have opted to attend to the esthetic quality of virtual epitexts and offer what Unsworth (2015) terms ‘multimodal literary narratives’ in their digital extensions.

As far as the tendencies observed are concerned, the five macrocategories defined in the initial descriptive analysis identify significant differences compared to the virtual dissemination scenario of previous years (Lluch et al., 2015; Tabernero-Sala, 2016b). Materiality emerges as the tendency with the greatest presence in virtual promotional documents, in keeping with developments in the field of children’s book publishing (Kümmerling-Meibauer, 2015; Tabernero-Sala, 2016a, 2019), although this is unevenly reflected in the documents analyzed, as some of them are simply recordings of a reader handling a book. Since 2020, the voices of authors and readers fill the virtual epitexts used in book promotion, emphasizing the documentary nature of the act of reading and reinforcing the presence of creators, mediators and readers in line with the overuse of the testimonial literary practices typical of the digital culture (Zafra, 2019).

To answer the question concerning the connection between promotional digital epitexts and the proposal of critical reading, we understand that what we have termed the ‘visible author’ (Sander’s, 2018) can be defined as one of the most significant discourse markers in the virtual epitexts selected in the qualitative content analysis. In this sense, the presence of the author in all aspects establishes cracks in the concept of ‘truth,’ by offering the reader a personal and ideological relationship to what is explained or related, and shows the mediator, and consequently the reader, the possibility of establishing methods of argumentation and questioning to use as a basis for the development of critical thinking, through the integration of different viewpoints, as advocated by Ennis (2011, 2018) and Kuhn (2018, 2019). It is not surprising that this concept of the ‘visible author’ includes the figure of the editor, thereby establishing the concept of a joint authorship that constantly addresses the reader. In the same vein, the empathic projection of

the ‘visible author’ paves the way to raising receivers’ awareness, involving, once again, the connection between emotion and knowledge, one of the most significant aspects in contemporary theories on non-fiction picturebooks (Grilli, 2020; Tabernero-Sala, 2022). Likewise, promotional virtual epitexts should emphasize and encourage the presence of discourse markers linked to the hybridization of fictional and non-fictional discourses through the use of different types of language, leading to, as the case may be, what we have termed experiential reading. In this way, the promotion reinforces the relationship between the reader, the book and reality by suggesting the connection of the book with reading environments in which shared natural settings are the priority, in line with the proposals contained in the Sustainable Development Goals. The artistic and experiential component extends the relationship between the book and its environment, through feedback that transforms the individual’s relationship to their context, as studied in the ecocritical paradigm of research on children’s books (Goga et al., 2018).

The emphasis on the materiality of the book and the presence of voiceovers, oral narrators and musical discourses aimed to arouse emotion and wonder, as well as curiosity, recover the sensorial aspect of reading, since the act of reading engages the body (Littau, 2006). With materiality as a basis, in-depth reading is identifiable in the concept of multimodal reading, which requires the reader to pause (Wolf, 2018; Zafra, 2019; Gil-Pelluch et al., 2020; Støle, 2020) for a time, which results in deliberation and enjoyment typical of artistic discourse.

In summary, virtual epitexts for the promotion of children’s picturebooks may provide mediators and readers with means, such as those revealed in the selection above, with which to develop the basis of critical thinking and meet, in this way, the challenge posed by the digital society in training readers within the framework of a biliteracy that has already been established (Wolf, 2018). Following on from previous studies (Tabernero-Sala et al., 2022), the selected virtual epitexts present discourse markers typical of the promotion of the non-fiction picturebook, thereby confirming,

TABLE 11 Discourse markers and strategies identified.

Discourse markers	Strategies for encouraging critical reading
Presence of the author of the book: voice of authority, opinion, reflection, subjectivity, questions for the reader.	The author is made visible as a metadiscursive strategy: this encourages reflection and questioning of the voice of authority
	Transmission of the author's empathic projection on the discourse: this encourages empathy, understanding and awareness
Hybridization of fiction and non-fiction: scientific information, questions, fictional characters, plural testimonies, symbiosis between the natural world and the fictional world.	Intensification of experiential reading: links are created between the reader's own experience and reading
	Strengthening of the relationship between the author, the reader, the book and the surroundings: curiosity and searching is encouraged, the relationship between the book and the surroundings is extended, through feedback that transforms the relationship of the individual with their context.
Emphasis on the materiality of the book: format, texts, illustrations, book covers, pop-ups.	Encouragement of exploration and the physical handling of the book, of enjoyment and detailed, in-depth reading
Presence of a narrator: voice, reading the book aloud, encouragement of reading, interpretation, guidance.	Proposal of models for reading aloud and oral narration: reading promoted as a sensorial experience
Discursive multimodality: texts, images, music, voice.	Proposal of a poetic and artistic discourse that addresses, moves, engages and interacts with the reader

as reported by recent studies (Von-Merveldt, 2018; Grilli, 2020; Salisbury, 2020; Dindelli, 2021; Jan, 2021), the vitality of this genre that has changed the panorama of children's books by defining a new reading model based on the hybridization of fictional and non-fictional discourses according to the contexts of the reader and the mediator.

5. Conclusion

The analysis of the results defines this research based on the following findings:

- Firstly, new tendencies resulting from the challenges posed by the digital society in recent years and new literate practices are described from the perspective of digital promotion (Cordón-García and Muñoz-Rico, 2023).
- Secondly, the results distinguish discourse markers that favor critical reading, which was previously non-existent in promotional virtual epitexts. The visibility of the author, the removal of the boundaries between reality and fiction, the experiential connection between the author, the reader and their environment, and the enhancement of the material

aspects of the discourse all emphasize in-depth reading through the use of digital promotion (Wolf, 2018), along the lines established in the premises on critical reading advocated in international documents, such as *Reimagining our Futures Together: A New Social Contract for Education* (Comisión Internacional sobre los Futuros de la Educación, 2022).

- It is therefore understood that promotional virtual epitexts create meaning (Gray, 2010) and propose, from the digital paradigm, a way of reading that fosters the essential analogical aspect of in-depth reading, which is key to critical thinking. Consequently, a corpus such as the one selected for this research could be defined as a selection of recommended good practices in the context of training future mediators (Trigo-Ibáñez and Santos-Díaz, 2023; Álvarez-Ramos et al., 2023).

This opens future lines of research to delve into the analysis of the virtual epitexts used in the promotion of reading and into the reading paradigm that they present. Similarly, from this perspective, there is a need for research into how social media influence the contents and methods used in the promotion of reading in the digital society. Finally, the study of model virtual epitexts should continue to transfer the results to reading mediators.

6. Limitations

This research was essentially conducted in the context of Spanish publishing, although some Latin American publications were also selected, since the main criterion was for the works to be published in Spanish. Therefore, we consider that the sample of 836 audio-visual documents is representative for the research objectives but the results are not generalizable to other contexts. Furthermore, the audio-visual documents in the sample have been taken from two digital media, YouTube and Vimeo, given that other social media replicated content and were more unstable; this means that some publishers who had abandoned these platforms may have produced significant publications via other digital media on the dates indicated that have not been registered. Regarding the method used, the approach is quantitative as regards data registration and detection of the initial categories, by frequency and saturation, and qualitative-interpretive as regards content analysis and reporting of results. As a consequence, the limitations involved in research with these characteristics, concerning representativeness or generalization, should be assumed in the interests of the reflection and in-depth analysis required by the objectives of this study as a contribution to educational research.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

RT-S: conceptualization, analysis, writing of the manuscript, review, and supervision. MC-C: methodology, analysis, editing, review, and supervision. Both authors contributed to the article and approved the submitted version.

Funding

This research was framed within the R&D&I Project entitled “Non-fictional readings for the integration of critical citizens in the new cultural ecosystem. RENFICE” (PID2021-126392OB-I00), financed by the Spanish Ministry of Science and Innovation.

References

- Abad, J. A., and Delicado, F. (2020). *El bolero de Ravel*. Pontevedra: Kalandraka.
- Alonso, J. R., and Paschetta, M. (2018). *Semillas. Un Pequeño gran viaje*. Barcelona: A buen paso.
- Alonso, S., and Muñoz, P. (2020). *Nunca dejes de brillar. Vilanova i la Geltrú*: Carambuco Ediciones.
- Álvarez-Ramos, E., Mateos-Blanco, B., and Alejaldre-Biel, L. (2023). Prácticas letradas de la sociedad tecnológica: Consumos lectores en educación superior. *Texto Livre* 16:e42044.
- Andersen, H., and Abramović, M. (2021). *El patito feo*. Barcelona: Libros del Zorro Rojo.
- Asociación Álbum (n.d.). *i Álbum! 25 editoriales independientes [websites]*. Available online at: <https://webdelalbum.org/quienes-somos/> (accessed December 20, 2022).
- Bajtín, M. M. (1981). “Discourse in the novel. 1934-1935,” in *The dialogic imagination: Four essays*, ed. M. M. Bajtín (Austin: University of Texas Press), 259–422.
- Bakhtin, M. M. (1981). *The dialogic imagination*. Austin: University of Texas Press.
- Baron, N. S. (2015). *Words onscreen: The fate of reading in a digital world*. New York, NY: Oxford University Press.
- Baron, N. S. (2021). *How we read now: Strategic choices for print, screen and audio*. Oxford: Oxford University Press. doi: 10.1093/oso/9780190084097.001.0001
- Barthes, R. (1994). “La muerte del autor,” in *El susurro del lenguaje. Más allá de la palabra y de la escritura*, ed. R. Barthes (Barcelona: Paidós), 65–72.
- Basaraba, N. (2016). Creating persuasive book trailers as a new media marketing tool. *Logos* 27, 34–51. doi: 10.1163/1878-4712-11112110
- Ben-Barak, I., and Frost, J. (2020). *i No chupes este libro!*. Barcelona: Takatuka.
- Bezanilla, M. J., Fernández-Nogueira, D., and Poblete, M. (2019). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. *Think. Skills Creat.* 33, 1–10. doi: 10.1016/j.tsc.2019.100584
- Biblioteca Nacional de España (n.d.). *Literatura Infantil y Juvenil. Editoriales especializadas [websites]*. Available online at: https://www.bne.es/es/Micrositios/Guias/Literatura_Infantil/EditorialesLibrerias/EditorialesEspecializadas/ (accessed December 20, 2022).
- Bilushchak, O., Radkovets, O., and Syerov, Y. (2020). “Internet marketing strategy promotion of a book in social media,” in *Proceedings of the 2nd International workshop on control, optimisation and analytical processing of social networks (Coapsn 2020)*, 260–272. Available online at: <http://ceur-ws.org/Vol-2616/paper22.pdf> (accessed December 20, 2022).
- Browne, A. (2021). *Ernesto, el elefante*. Pontevedra: Kalandraka.
- Caro-Valverde, M. T. (2018). La comunicación argumentativa en la Sociedad del Conocimiento, clave del liderazgo distribuido para un cambio educativo desde el desarrollo profesional. *Rev. Educ. Distancia (RED)* 8, 1–30. doi: 10.6018/red/56/8
- Cencerrado Malmierca, L. M., and Yuste Tuero, E. (2020). Tiempos confinados, cambios de hábitos en el ocio y la lectura, un valor en alza. *Rev. Electrón. Leer Escribir Descubrir* 1, 82–88.
- Comín, M. (2021a). *Bienvenida*. Barcelona: A buen paso.
- Comín, M. (2021b). *Mi arbolito de Navidad*. Barcelona: Combel Editorial.
- Comisión Internacional sobre los Futuros de la Educación (2022). *Reimaginar juntos nuestros futuros, un nuevo contrato social para la educación*. UNESCO y Fundación SM. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000381560> (accessed December 20, 2022).
- Cordón-García, J. A. (2018). Combates por el libro. Inconclusa dialéctica del modelo digital. *Prof. Inf.* 27, 467–481. doi: 10.3145/epi.2018.may.02
- Cordón-García, J. A. (2020). La lectura digital. Intelectión, apropiación y contextos. *Biblioteche Oggi Trends* 6, 28–40.
- Cordón-García, J. A., and Muñoz-Rico, M. (2023). La socialización de la lectura y la escritura: Fandom, redes y visibilidad editorial. *Texto Livre* 16:e37067.
- Cruces, F. (2017). *¿Cómo leemos en la sociedad digital? Lectores, booktubers y prosumidores*. Barcelona: Ariel y Fundación Telefónica.
- Dautremere, R. (2021). *Tan solo un instante*. Zaragoza: Edelvives.
- De Amo Sánchez-Fortún, J. M. (2021). *La lectura y la escritura como prácticas sociales en la cultura digital*. Valencia: Tirant lo Blanch.
- De Amo, J. M., and García-Roca, A. (2021). Mechanisms for interpretative cooperation: Fan theories in virtual communities. *Front. Psychol.* 12:699976. doi: 10.3389/fpsyg.2021.699976
- Delgado, P., Vargas, C., Ackerman, R., and Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on Reading comprehension. *Educ. Res. Rev.* 25, 23–38. doi: 10.1016/j.edurev.2018.09.003
- Dimova, M., Slavova-Petkova, S., and Luchev, D. (2018). Models of digital book trailers and applications of digital storytelling approach for educational purposes. *Digit. Present. Preserv. Cult. Sci. Heritage* 8, 197–206. doi: 10.55630/dipp.2018.8.19
- Dindelli, I. (2021). *I nuovi non-fiction picturebook. Uno studio sugli albi illustrati di divulgazione per l'infanzia*. Ph.D. thesis. Bologna: Università di Bologna.
- Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *Univ. Illinois* 2, 1–8.
- Ennis, R. H. (2018). Critical thinking across the curriculum: A visión. *Topoi* 37, 165–184. doi: 10.1007/s11245-016-9401-4
- ERI-Lectura (2020). *Hábitos lectores durante el confinamiento por la covid-19. Informe ejecutivo*. RODERIC, Universitat de València. Available online at: <https://roderic.uv.es/handle/10550/74797> (accessed December 20, 2022).
- España, Cortes Generales (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación (BOE-A-2020-17264). *Boletín Oficial Estado* 340, 122868–122953.
- España, Ministerio de Educación y Formación Profesional (2022). Real Decreto 157/2022, de 1 de marzo, por el que se establecen la ordenación y las enseñanzas mínimas de la Educación Primaria (BOE-A-2022-3296). *Boletín Oficial Estado* 52, 1–109.
- Federación de Gremios de Editores de España (n.d.). *Directorio de editoriales. Infantil y juvenil. Publishing houses directory. Infant and youth. Ministerio de industria, comercio y turismo*. Available online at: https://www.federacioneditores.org/img/documentos/INFANTIL_JUVENIL.pdf (accessed December 20, 2022).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ferrada, M. J., and Paschetta, M. (2021). *Zum zum. El viaje de la semilla*. Barcelona: A buen paso.
- Freire's, P. (2002). *Pedagogía del oprimido*. Madrid: Siglo XXI.
- Freire, P. (2010). *Cartas a quien pretende enseñar*. Buenos Aires: Siglo XXI.
- García-Roca, A., and De-Amo, J. M. (2019). Jóvenes escritores en la red: Un estudio exploratorio sobre perfiles de Wattpad. *Ocnos Rev. Estudios Sobre Lectura* 18, 18–28. doi: 10.18239/ocnos_2019.18.3.1968
- Geis, P. (2022). *Los volcanes*. Barcelona: Combel Editorial.
- Genette, G. (2004). *Metalepsis. De la figura a la ficción*. Buenos Aires: Fondo de Cultura Económica.
- Gil-Pelluch, L., Delgado-Herrera, P., Vargas-Pecino, C., Vergara-Martínez, M., and Salmerón-González, L. (2020). La lectura en pantalla en las aulas. *Textos Didáctica Lengua Lit.* 89, 41–47.
- Goes, P. (2016). *La línea del tiempo*. Madrid: Maeva Ediciones.
- Goga, N. (2020). “Verbal and visual informational strategies in non-fiction books awarded and mentioned by the Bologna Ragazzi Award. 2009-2019,” in *Non-fiction picturebooks. Sharing knowledge as an esthetic experience*, ed. G. Grilli (Pisa: Edizioni ETS), 51–68.
- Goga, N., Guanío-Uluru, L., Hallás, B. O., and Nyrnes, A. (Eds.) (2018). *Ecocritical perspectives on children's texts and cultures. Nordic dialogues*. London: Palgrave Macmillan. doi: 10.1007/978-3-319-90497-9
- Goga, N., Iversen, S. H., and Teigland, A. S. (eds.) (2021). “Verbal and visual strategies in nonfiction picturebooks,” in *Theoretical and analytical approaches* (Oslo: Scandinavian University press). doi: 10.18261/9788215042459-2021
- Gray, J. (2010). *Show sold separately. Promos, spoilers, and other media paratexts*. New York, NY: New York University press.
- Grilli, G. (2020). “Beauty and the World. Some questions to five illustrators (often also authors) of non-fiction picturebooks for children on their work, poetics, inspiration(s),” in *Non-fiction picturebooks. Sharing knowledge as an esthetic experience*, ed. G. Grilli (Pisa: Edizioni ETS), 267–293.
- Grilli, G. (2021). “The artistic nonfiction picturebook,” in *Verbal and visual strategies in nonfiction picturebooks. Theoretical and analytical approaches*, eds N. Goga, S. H. Iversen, and A. S. Teigland (Oslo: Scandinavian University Press), 22–36.
- Gron, R. (2014). Literary experience and the book trailer as intermedial paratext. *SoundEffects* 4, 90–107. doi: 10.7146/se.v4i1.20330
- Guzmán-Cedillo, Y. I., and Flores-Macías, R. C. (2020). La competencia argumentativa como meta en contextos educativos. *Rev. Lit. Educ.* 56, 15–34. doi: 10.5565/rev/educar.1009
- Haughton, C. (2019). *i Shhh! Tenemos un plan*. Madrid: NubeOcho.
- Hernández-Sevillano, D., and Cerro-Rico, M. (2019). *i Artista!*. Madrid: Bookolia.
- Hodges, K., and Lee-Merriam, H. (2020). *Brujas, guerras, diosas*. Barcelona: Libros del Zorro Rojo.
- Ibarra-Rius, N., and Ballester-Roca, J. (2017). De literatura, cine, publicidad e internet: Relatos digitales y book trailers en la formación de docentes. *@tic. Rev. D'innov. Educ.* 19, 47–54. doi: 10.7203/attic.19.11006
- Jan, C. (2021). Literatura infantil y juvenil: Erase una vez un éxito. *El País Diciembre* 569, 2–3.
- Jeffers, O. (2019). *El destino de Fausto*. Picassent: Andana Editorial.
- Kaufman, R., and Franco, R. (2017). *Abecedario. Abrir, bailar, comer y otras palabras importantes*. Madrid: Bruño.
- Kovač, M., and Van Der Weel, A. (eds.) (2020). *Lectura en papel vs. Lectura en pantalla*. Bogotá: CERLALC-UNESCO.
- Krippendorff, K. (2019). *Content analysis. An introduction to its methodology*. London: Sage. doi: 10.4135/9781071878781
- Kuhn, D. (2018). A role for reasoning in a dialogic approach to critical thinking. *Topoi* 37, 121–128. doi: 10.1007/s11245-016-9373-4
- Kuhn, D. (2019). Critical thinking as discourse. *Hum. Dev.* 62, 146–164. doi: 10.1159/000500171
- Kümmerling-Meibauer, B. (2015). From baby books to picturebooks for adults: European picturebooks in the new millennium. *Word Image* 31, 249–264. doi: 10.1080/02666286.2015.1032519
- Kümmerling-Meibauer, B., and Meibauer, J. (2021). “How descriptive picturebooks engaged children in knowledge about coal, oil, and gas,” in *Verbal and visual strategies in nonfiction picturebooks. Theoretical and analytical approaches*, eds N. Goga, S. H. Iversen, and A. S. Teigland (Oslo: Scandinavian University Press), 189–200. doi: 10.18261/9788215042459-2021
- Lebrun, M., Lacelle, N., and Boutin, J. F. (eds.) (2018). Dispositifs numériques pour l'enseignement de la littérature. *Rev. Rech. Lit. Méd. Multimodale* 8. doi: 10.7202/1050932ar
- Littau, K. (2006). *Theories of reading. Books, bodies and bibliomania*. Cambridge: Polity Press.
- Lluch, G. (2017). “Jóvenes y adolescentes comparten la lectura,” in *¿Cómo leemos en la sociedad digital? Lectores, booktubers y prosumidores*, ed. F. Cruces (Barcelona: Ariel y Fundación Telefónica), 81–102.
- Lluch, G. (2021). “El canon lector creado entre iguales. Estudio de caso: la recomendación virtual,” in *La lectura y la escritura como prácticas sociales en la cultura digital*, ed. J. M. De-Amo Sánchez-Fortún (Valencia: Tirant humanidades), 51–73.
- Lluch, G. (2018). *Claves para promocionar la lectura en la red*. Madrid: Síntesis.
- Lluch, G., Tabernero-Sala, R., and Calvo-Valios, V. (2015). Epitextos virtuales públicos como herramientas para la difusión del libro. *Prof. Inf.* 24, 797–804. doi: 10.3145/epi.2015.nov.11
- Mangen, A., Olivier, G., and Velay, J. L. (2019). Comparing comprehension of a long text read in print book and on kindle: Where in the text and when in the story? *Front. Psychol.* 10:38. doi: 10.3389/fpsyg.2019.00038
- Millán, J. A. (ed.) (2017). *La lectura en España. Informe 2017*. Madrid: Ministerio de Educación y Cultura and Federación de Gremios de Editores de España.
- Mizrachi, D., Salaz, A. M., Kurbanoglu, S., Boustany, J., and on behalf of the Arfis Research Group. (2018). Academic reading format preferences and behaviors among university students worldwide: A comparative survey analysis. *PLoS One* 13:e0197444. doi: 10.1371/journal.pone.0197444
- OECD (2021). *21st-century readers: Developing literacy skills in a digital world*. Paris: PISA, OECD Publishing. doi: 10.1787/a83d84cb-en
- Pacheco, G. (2020). *Caperucita roja (primero sueño)*. Santa Cruz de Tenerife: Diego Pun Ediciones.
- Reyes, E., and Solé-Vendrell, C. (2020). *Muñeco de barro*. Barcelona: Libros del Zorro Rojo.
- Romero-Oliva, M. F., Heredia-Ponce, H., and Romero-Claudio, C. (2023). Promoción de la lectura y transmedia: De la creación editorial al book-trailer como epitexto ficcional. *Texto Livre* 16.
- Romero-Oliva, M. F., Heredia-Ponce, H., and Sampérez-Hernández, M. (2019). El 'book-trailer' como herramienta digital en la formación lectora de los futuros docentes. Un estudio de caso. *Caracteres. Estudios Cult. Crít. Esfera Digit.* 8, 92–128.
- Romero-Oliva, M. F., Heredia-Ponce, H., and Trigo-Ibáñez, E. (2020). Las TIC en el fomento lector de los adolescentes. Un estudio de caso desde las creencias docentes. *Contextos Educ.* 25, 105–125. doi: 10.18172/con.4249
- Rovira-Collado, J. (2017). Booktrailer y booktuber como herramientas LIJ 2.0 para el desarrollo del hábito lector. *Invest. Sobre Lectura* 7, 55–72. doi: 10.24310/revistaisl.vi7.10981
- Ruiz-Olabuénaga, J. I. (2009). *Metodología de investigación cualitativa*. Bilbao: Universidad de Deusto.
- Salisbury, M. (2020). “True stories and big books: New creative opportunities in non-fiction picturebooks,” in *Non-fiction picturebooks. Sharing knowledge as an esthetic experience*, ed. G. Grilli (Pisa: Edizioni ETS), 91–108.
- Salten, F., and Lacombe, B. (2020). *Bambi, una vida en el bosque*. Zaragoza: Edelvives.
- Sánchez-Muñoz, E. (2022). Impacto del confinamiento por COVID-19 en las plataformas de préstamo digital: El caso de GaliciaLe. *Rev. Española Doc. Cien.* 45, 1–14. doi: 10.3989/redc.2022.1.1843
- Sander's, J. S. (2018). *A literature of questions. Nonfiction for the Critical Child*. Minneapolis, MN: University of Minnesota Press. doi: 10.5749/j.ctt1pwt67w
- Santaulàlia, J. N., and Lozano, L. (2022). *Tiempo de haikus*. Barcelona: Akiara Books.
- Schilhab, T., Balling, G., and Kuzmčová, A. (2020). “La disminución de la materialidad en el tránsito de la lectura impresa a la lectura en pantalla,” in *Lectura en papel vs. Lectura en pantalla*, eds M. Kovač and A. Van der Weel (Bogotá: CERLALC-UNESCO), 106–127.
- Scott, J., and Smith, S. (2021). *Hablo como el río*. Barcelona: Libros del Zorro Rojo.
- Seceda, T., and García, E. (2021). *¿Quién soy? Crías de animales*. Pontevedra: Kalandraka.
- Smith, J. M., and Robertson, M. K. (2019). Navigating award-winning nonfiction children's literature. *Read. Teach.* 73, 195–204. doi: 10.1002/trtr.1811
- Squilloni, A., and Fosch, C. (2021). *Frederick Douglass: ¿Debo argumentar el sinsentido de la esclavitud?*. Barcelona: Akiara Books.
- Starkoff, V. (2022). *i Gracias, madre Tierra!*. Barcelona: Akiara Books.
- Støle, H. (2020). “El mito del nativo digital: ¿por qué necesitan libros?” in *Lectura en papel vs. Lectura en pantalla*, eds M. Kovač and A. Van der Weel (Bogotá: CERLALC-UNESCO), 49–69.
- Tabernero-Sala, R. (2016b). Los epitextos virtuales en la difusión del libro infantil: Hacia una poética del book-trailer. Un modelo de análisis. *Ocnos Rev. Estudios Sobre Lectura* 15, 21–36. doi: 10.18239/ocnos_2016.15.2.1125
- Tabernero-Sala, R. (2016a). El lector en el espacio del libro. *Fuera Margen* 19, 22–25.

- Tabernero-Sala, R. (2018). "El bibliotrailer en la promoción del libro y de la lectura," in *Claves para promocionar la lectura en la red*, ed. G. Lluch (Madrid: Síntesis), 64–78.
- Tabernero-Sala, R. (2021). "Entre paradigmas analógicos y digitales. El book-trailer en la formación de lectores y mediadores," in *La lectura y la escritura como prácticas sociales en la cultura digital*, ed. J. M. De Amo Sánchez-Fortún (Valencia: Tirant lo Blanch), 111–136.
- Tabernero-Sala, R. (ed) (2022). *Leer por curiosidad. Los libros de no ficción en la formación de lectores*. Barcelona: Graó.
- Tabernero-Sala, R. (2019). *El objeto libro en el universo infantil. La materialidad en la construcción del discurso*. Zaragoza: Prensas de la Universidad de Zaragoza.
- Tabernero-Sala, R., and Calvo-Valios, V. (2016). Book-trailers as tools to promote reading in the framework of the Web 2.0. *New Rev. Child. Lit. Librariansh.* 22, 53–69. doi: 10.1080/13614541.2016.1120071
- Tabernero-Sala, R., and Laliena, D. (2023). El libro ilustrado de no ficción en la formación de lectores: Análisis de las claves discursivas y culturales para leer en la sociedad digital. *Texto Livre* 16, 1–13. doi: 10.1590/1983-3652.2023.41926
- Tabernero-Sala, R., Colón-Castillo, M. J., Sampériz-Hernández, M., and Campos-Bandrés, I. O. (2022). Promoción de la lectura en la sociedad digital. El book-trailer del libro ilustrado de no ficción como epitexto virtual en la definición de un nuevo discurso. *Prof. Inf.* 31, 1–17. doi: 10.3145/epi.2022.mar.13
- Trigo-Ibáñez, E., and Santos-Díaz, I. C. (2023). Ferramentas virtuais na concepção de leitura de futuros professores: Uma exploração lexical. *Texto Livre* 16, 1–19. doi: 10.1590/1983-3652.2023.41798
- Trigo-Ibáñez, E., Saiz-Pantoja, R., Sánchez-Arjona, E., and Romero-Oliva, M. F. (2022). Desarrollar el pensamiento crítico con el libro ilustrado de no ficción en el marco del tercer espacio educativo. *Rev. Interuniv. Formación Prof. Continuación Antigua Rev. Escuelas Norm.* 36:97. doi: 10.47553/rifop.v97i36.3.96690
- Unión Europea (2016). *Promoting Reading in the digital environment. Publicaciones de la Unión Europea*. Available online at: <https://op.europa.eu/en/publication-detail/-/publication/9052931a-2ece-11e6-b497-01aa75ed71a1> (accessed December 20, 2022).
- Unsworth, L. (2015). Persuasive narratives: Evaluative images in picture books and animated movies. *Vis. Commun.* 14, 74–96. doi: 10.1177/1470357214541762
- Von-Merveldt, N. (2018). "Informational picturebooks," in *The routledge companion to picturebooks*, ed. B. Kümmerling-Meibauer (London: Routledge), 231–245. doi: 10.4324/9781315722986
- Witarsa, and Muhammad, S. (2023). Critical thinking as a necessity for social science students capacity development: How it can be strengthened through project based learning at university. *Front. Educ.* 7:983292. doi: 10.3389/feduc.2022.983292
- Wolf, M. (2018). *Reader, come home. The reading brain in a digital world*. New York, NY: Harper.
- Ylla. (2021). *Dos ositos*. Pontevedra: Kalandraka.
- Zafra, R. (2017). "Itinerarios del yo en un cuarto propio conectado," in *¿Cómo leemos en la sociedad digital? Lectores, booktubers y prosumidores*, ed. F. Cruces (Barcelona: Ariel y Fundación Telefónica), 81–102.
- Zafra, R. (2019). Conferencias Aranguren ¿Fin de la intimidad? La (im)posibilidad de un mundo sin párpados. *Ensayo Sobre Intimidad Conectada Isegoria* 60, 51–68. doi: 10.3989/isegoria.2019.060.03



OPEN ACCESS

EDITED BY
Pedro García Guirao,
WSB Universities,
Poland

REVIEWED BY
Jean-Paul Fischer,
University of Lorraine,
France
Xingshan Li,
Institute of Psychology (CAS),
China
Meina Luo,
Zhejiang Normal University,
China

*CORRESPONDENCE
Li Yin
✉ yinl@tsinghua.edu.cn

SPECIALTY SECTION
This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

RECEIVED 23 January 2023
ACCEPTED 13 March 2023
PUBLISHED 06 April 2023

CITATION
Yin L and McBride C (2023) A longitudinal study
on sensitivity to symmetry in writing and
associations with early literacy abilities.
Front. Educ. 8:1150075.
doi: 10.3389/feduc.2023.1150075

COPYRIGHT
© 2023 Yin and McBride. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

A longitudinal study on sensitivity to symmetry in writing and associations with early literacy abilities

Li Yin^{1*} and Catherine McBride²

¹Department of Foreign Languages and Literatures, Center for the Study of Language and Psychology, Tsinghua University, Beijing, China, ²Department of Human Development and Family Science, Purdue University, West Lafayette, IN, United States

Introduction: Young children reportedly find symmetrical prints, such as H, easier to copy and produce from memory than asymmetrical ones. Little is known about how sensitivity to symmetry in print relates to early word reading and writing development. We examined this in Chinese, a visually complex writing system featuring symmetrical patterns, such as 金 or 美.

Methods: Sixty preschoolers in Beijing completed a character decision task (Experiment 1, $M_{\text{age}}=62.16$ months) and a character learning task (Experiment 2, $M_{\text{age}}=63.96$ months), using stimuli matched on visual complexity and frequency but differing in symmetry, and were assessed on Chinese word reading and word writing abilities twice with a 2-year interval.

Results: Children were significantly more likely to endorse symmetrical complex stimuli as possible Chinese than asymmetrical complex ones, and they learned to read symmetrical characters significantly better than asymmetrical ones, reflecting their sensitivity to symmetry in Chinese characters. However, with age and nonverbal IQ statistically controlled, sensitivity to symmetry was not uniquely associated with Chinese word reading or word writing abilities at any time; rather, asymmetrical character learning, which necessitates reduced sensitivity to symmetry, was uniquely associated with Chinese word writing concurrently and longitudinally.

Discussion: Findings highlight the close relationship between analytic perception of written words and early writing ability.

KEYWORDS

symmetry, writing, perception, early literacy, Chinese

Introduction

Symmetry in writing refers to that if a word or character were folded along an axis, the geometrical patterns of the two sides would largely overlap. Symmetrical patterns in writing can be around a vertical axis (such as M in English or 非 in Chinese), a horizontal axis (such as B or 臣), or both (such as H or 甲). There are increasing observations showing the relative facility of symmetrical shape perception and production. For example, 4-month-old infants processed vertical symmetrical patterns more efficiently than asymmetrical or horizontally symmetrical ones (Bornstein et al., 1981). Young learners of the Latin alphabet found symmetrical letters easier to copy and write from memory than asymmetrical letters (Treiman and Kessler, 2011). Chinese preschoolers wrote their names better with symmetrical characters than with

asymmetrical ones (Yin and Treiman, 2013). French 5 to 6-year-olds wrote symmetrical digits 0 and 8 with a higher percentage of correctness than the eight other asymmetrical digits (Fischer, 2018).

How is sensitivity to symmetry in writing associated with early word reading and writing abilities? For learning to read and write, unlearning symmetrization or mirror invariance is needed because children evolutionarily tend to treat symmetrical patterns as mirror images of one another and consequently take them as the same thing (Dehaene et al., 2010). For example, to correctly identify a word, children need to develop print-specific understanding that symmetrical images of *d* and *b* are two different letters. Sensitivity to spatial orientation has been found to significantly relate to reading skills in English-reading first graders (Fernandes et al., 2016) and Chinese-reading 6-year-olds (Zhang et al., 2021). Compared with early word reading, there is less report on the association between sensitivity to symmetry in print and early word writing. Writing a word requires a more detailed and accurate representation of the graphic unit in words than reading, thus writing may reveal more individual differences in the analytic perception of sub-word units than reading. This study examined young children's sensitivity to symmetry in Chinese and its associations with early writing as well as reading abilities.

Symmetry effect in reading and writing Chinese

Chinese writing features abundance of symmetrical patterns. According to the study by Xu (2012), among the 3,500 characters in the List of Common Characters in Modern Chinese (China National Working Committee on Languages and Characters, 1988), 414 characters are visually symmetrical. Of these, the majority (93.48%, 378 in total) are vertically symmetrical, which include 188 perfectly-symmetrical characters (45.41%), showing an absolute overlap of the two sides when folded along the vertical axis (such as 凹 or 草), and 199 relatively-symmetrical characters (48.07%) (such as 容 or 再). There are 3 characters that are only horizontally symmetrical (巨, 旧, and 臣) and 24 characters that are both vertically and horizontally symmetrical (such as 王 or 中). It should be noted that these symmetrical whole characters can also appear as sub-character components in a large number of compound characters in Chinese (such as 榕, 溶, 忠, 钟, 枉, 忤), as detailed below.

Chinese character is visually complex. The majority of the 2,570 characters in school Chinese in mainland China have 7 to 12 strokes; some have more than 20 strokes (Shu et al., 2003). Over 80% of Chinese characters are compound characters consisting of a semantic radical that gives a clue to the character's meaning and a phonetic radical that gives a clue to the character's pronunciation. Expert readers can identify the radical or reoccurring stroke pattern in a character, such as “艹” in characters “花” and “草,” whereas novice readers may process the character as a whole rather than by sub-character components (Hsiao and Cottrell, 2009).

A symmetry effect in Chinese character recognition was reported in adults (Yu and Cao, 1992; Chen and Huang, 1999; Huang et al., 2002). Yu and Cao (1992) found that expert readers processed balanced characters, where strokes are distributed equally across the two halves of the character, faster than unbalanced ones. This effect was found in low-frequency characters but not in high-frequency ones. According to the researchers, symmetrical patterns enhance perceptual wholeness. In expert character recognition, whole

character processing and component processing compete for attentional resources. For low-frequency characters, whole-character processing has a cognitive advantage over component processing in symmetrical characters due to enhanced perceptual wholeness, whereas for high-frequency (i.e., familiar) characters, whole-character processing has a cognitive advantage over component processing for both symmetrical and asymmetrical characters and thus no symmetry effect was found. According to Chen and Huang (1999), information redundancy also accounts for the symmetry effect in Chinese. In their study of adult readers, symmetrical characters were processed faster than asymmetrical ones in high-complexity characters but not in low-complexity ones. The researchers interpreted that in symmetrical complex characters, recognizing half of the character leads to recognition of the other, while in asymmetrical complex characters, two or more parts need to be recognized. This effect was not found in low-complexity characters because the processing time is short for both symmetrical and asymmetrical characters.

Few studies explicitly reported symmetry effect in Chinese acquisition in young children. Li (1964) examined Chinese preschoolers' perceptual sensitivity to graphemes in characters. Ten 6-year-olds were asked to look at a character and then write it. The researcher recorded the stroke order and stroke pause children made. It was found that children perceived similar and roughly symmetrical strokes as one group, such that in writing the character, they would first write these parts out and then filled in the remaining parts, showing a tendency to perceive similar or roughly symmetrical parts as one grapheme. For example, in writing the character “蜀,” children first wrote the two “弓” parts on the two sides and then filled in the middle part “百.” However, the sample size of this study was small and the relation between sensitivity to perceptual symmetry in characters and children's reading or writing abilities was not examined.

Visual-perceptual sensitivity and early literacy development

Children begin to learn about the visual form of writing as early as 2 and 3 years old before they learn about how the visual form represents language (Zhang et al., 2017). By age 5, children have developed quite a bit of knowledge about the visual form of writing they are exposed to and such knowledge is significantly related to concurrent and subsequent reading and writing abilities (Kessler et al., 2013; Yin and McBride, 2015; Treiman, 2017). For example, in the study of Yin and McBride (2015), Chinese 5-year-olds demonstrated sensitivity to the positional regularity of radicals in Chinese and such sensitivity explained unique variance in their word reading and writing abilities 1 year later. To date, most studies investigated differentiation between writing and drawing (e.g., Levin and Bus, 2003), discrimination of native writing symbols from foreign symbols (e.g., Lavine, 1977; Zhang et al., 2017), and detection of graphotactic patterns in written words (e.g., Pollo et al., 2009). Less research has examined sensitivity to symmetry in writing, a visual feature that transcends orthographic units in a word and carries no phonological or semantic information, and its association with literacy abilities.

Learning to read or write involves more than visual perception of the written word. A meta-analysis of 34 studies on the relationship between visual skills and Chinese reading acquisition (Yang et al., 2013) showed that various visual skills (including visual perception,

speed of processing visual information, and pure visual memory) had low-to-moderate correlations with Chinese reading for children in the lower grades (preschool to second grade), but not in higher grades (second through sixth grades). Contrastingly, visual-verbal association skill accounted for 34% and 41% of the variance in Chinese reading acquisition for children in both lower and higher grades, respectively. Learning to read asymmetrical characters may reveal more visual-verbal association skill than learning to read symmetrical characters and thus might be more closely associated with word reading or writing abilities because it involves more analytic perception of sub-components in a character. We tested this idea in this study.

The present study

The present study examined Chinese preschoolers' sensitivity to symmetry in characters and its concurrent and longitudinal associations with Chinese word reading and word writing abilities. We used a character decision task to tap pure perceptual sensitivity to symmetry in characters and a character learning task to tap analytic perception of sub-character components as well as sensitivity to symmetry across sub-character components. We assessed children's Chinese word reading and word writing abilities twice, with a 2-year interval.

In the character decision task, children were asked to decide whether a rare character was a possible Chinese. There were 25 symmetrical characters and 25 asymmetrical characters matched on frequency and visual complexity. If children were sensitive to symmetry in Chinese, they would more likely endorse symmetrical stimuli as possible Chinese than asymmetrical ones.

In the character learning task, children were taught to pronounce two sets of five rare characters matched on frequency and visual complexity but differing in symmetry. Given that learning to read involves mapping a visual form onto a spoken form, if children found symmetrical patterns easier to perceive than asymmetrical ones, they would learn to pair symmetrical characters with a pronunciation better than asymmetrical ones.

In view of the critical role of orthographic skills in early Chinese literacy development (Wang et al., 2015; Yin and McBride, 2015) and the limited role of pure visual perceptual skills compared with visual-verbal association skill in Chinese reading acquisition (Yang et al., 2013), we hypothesized that analytic perception of sub-character components, rather than pure perceptual sensitivity to symmetry in characters, may uniquely predict Chinese literacy abilities longitudinally.

Method

Participants

Sixty preschoolers from a public preschool in a middle-class neighborhood in Beijing participated in the study. Thirty children participated in Experiment 1 ($M_{\text{age}} = 62.16$ months, 10 girls,) and thirty children participated in Experiment 2 ($M_{\text{age}} = 63.96$ months, 12 girls). At Time 1, the majority (78.3%) of children were in their last (third) year of kindergarten and 21.7% were in their second year of kindergarten (7 and 6 children in Experiment 1 and 2, respectively). Written consent was obtained from the guardians of all children.

Procedure and tasks

At Time 1, we administered a character decision task in Experiment 1, and a character learning task in Experiment 2, and assessed Chinese word reading, Chinese word writing, and non-verbal IQ for all children. At Time 2, 2 years later, when the majority of children became second graders and 21.7% became first graders in the primary school, we assessed Chinese word reading and Chinese word writing for all children again.

Character decision

Children were asked to decide whether a stimulus was a possible Chinese character. The stimuli included 50 rare characters selected based on symmetry and visual complexity. Asymmetrical and symmetrical characters were matched on frequency (time of occurrence per million) and visual complexity (number of strokes) (frequency: *Mean* (SD) = 47 (93) and 56 (260) for asymmetrical and symmetrical characters, respectively, $p = 0.361$; visual complexity: *Mean* (SD) = 8.36 (1.89) and 8.24 (2.19) for asymmetrical and symmetrical characters, respectively, $p = 0.704$). High- and low-complexity characters were matched on the frequency and differed in complexity (frequency: *Mean* (SD) = 43 (92) and 60 (264) for high-complexity and low-complexity characters, respectively, $p = 0.07$; visual complexity: *Mean* (SD) = 9.96 (1.16) and 6.56 (1.12) for high-complexity and low-complexity characters, respectively, $p < 0.001$). Table 1 shows the sample stimuli. The Cronbach's α for this task was 0.95.

Character learning

This task was modeled on the word-learning paradigm used with young children in previous studies (McBride-Chang and Treiman, 2003). Children were taught to pronounce two sets of five rare characters differing in symmetry but matched on frequency (time of occurrence per million) and visual complexity (number of strokes) (frequency: *Mean* (SD) = 3 (6) and 72 (145) in Set A and 0 (0) and 1 (0) in Set B for asymmetrical characters and symmetrical characters, respectively, $ps > 0.278$; visual complexity: *Mean* (SD) = 9 (3.53) and 9 (3.46) in Set A and 9 (2.00) and 9.8 (2.59) in Set B for asymmetrical characters and symmetrical characters, respectively, $ps = 0.767$). Table 1 shows the sample stimuli. The task contained a demonstration phase and up to 8 learning trials. The experimenters began by saying that they would play a learning game with the child, that the characters to be learned were unfamiliar, and that the child should try his or her best. In the demonstration phase, the experimenter showed the child a card and pronounced the character, running her fingers under the character. The experimenter then asked the child to follow by saying "This character is read ____." We used a pseudo pronunciation, which was not the correct pronunciation for either the symmetrical or asymmetrical character. The procedure was repeated for all five stimuli in each set and the order of demonstration was randomized for each child. Once all five stimuli had been introduced in the first learning trial, the experimenter showed the child one of the cards and asked the child to give the pronunciation of the character to which it corresponded. The experimenter praised the child if s/he pronounced the character correctly and provided the correct answer if s/he pronounced it incorrectly. Learning trials continued until the child achieved the

TABLE 1 Sample stimuli in character decision in Experiment 1 (upper panel) and character learning in Experiment 2 (lower panel).

Condition	Sample stimulus in character decision	Complexity	Number of strokes
Symmetrical	𠂇	Low-complexity	6
	𠂈	High-complexity	11
Asymmetrical	𠂉	Low-complexity	6
	𠂊	High-complexity	11

Condition	Sample stimulus in character learning	Pseudo pronunciation used	Number of strokes
Symmetrical	𠂇	/bù/	7
Asymmetrical	𠂈		8
Symmetrical	𠂉	/fāng/	9
Asymmetrical	𠂊		9
Symmetrical	𠂋	/gǔ/	7
Asymmetrical	𠂌		7

criterion of responding correctly to all five items on two consecutive trials (McBride-Chang and Treiman, 2003). The Cronbach's α for this task was 0.97.

Chinese word reading

Children read aloud 50 Chinese single-character words presented in order of increasing difficulty (Wang et al., 2015). The Cronbach's α was 0.97 and 0.96 in Experiment 1 and 2, respectively at Time 1, and was 0.94 and 0.92 for Experiment 1 and 2, respectively, at Time 2.

Chinese word writing

Children were asked to write 10 orally-familiar words including 3 single-character words and 7 two-character words, totaling 17 characters (Wang et al., 2015). Each word was orally presented twice in a meaningful sentence to ensure the child understood the intended word. The Cronbach's α was 0.83 and 0.85 in Experiment 1 and Experiment 2, respectively, at Time 1, and was 0.89 and 0.87 for Experiment 1 and Experiment 2, respectively, at Time 2.

Non-verbal IQ

Sets A and B of Raven's Standard Progressive Matrices (Raven et al., 1996) were administered to measure children's nonverbal IQ at Time 1. The Cronbach's α was 0.85.

Results

Experiment 1

We analyzed the data in this study using IBM SPSS Version 27. Table 2 shows the descriptive statistics and correlation matrix in Experiment 1. We first analyzed the data of character decision using generalized linear mixed modeling (GLMM). As an extension of generalized linear modeling, GLMM allows for the linear predictors containing random effects as well as fixed effects and the response variables from different distributions, such as binary responses. In the model we built, the dependent variable was whether the rare

character was endorsed as a possible Chinese or not (coded as 1 and 0, respectively), the random effects were child and item, and the fixed effects were symmetry (symmetrical vs. asymmetrical), visual complexity (high vs. low), and their interaction. We selected binary logistic regression for target distribution and logit link transformation. The final model that offered the best fit to the observed performance of children had an overall correct-classification rate of 73.7%, $F(3, 1,496) = 12.08$, $p < 0.001$. As shown in Figure 1, there was a main effect of visual complexity, $F(1, 1,496) = 23.31$, $p < 0.001$, such that children were more likely to endorse high-complexity stimuli as possible Chinese than low-complexity ones ($\beta = 1.01$, $SE = 0.20$, 95% CI [0.62, 1.41], $p < 0.001$). There was also a significant interaction between symmetry and visual complexity, $F(1, 1,496) = 5.87$, $p = 0.016$. Children were more likely to endorse symmetrical high-complexity stimuli as possible Chinese than asymmetrical high-complexity ones ($\beta = 0.68$, $SE = 0.28$, 95% CI [0.13, 1.22], $p = 0.016$). No such difference was found for low-complexity stimuli. The random effect of child was significant (EST = 3.58, SEE = 1.12, 95% CI [1.94, 6.61], $p = 0.001$). The random effect of item was not significant, $p = 0.502$. These results indicate that Chinese children averaging 5 years of age are sensitive to symmetry in visually-complex Chinese characters.

Before running association tests to examine the concurrent and longitudinal associations between sensitivity to symmetry tapped in the character decision task and children's literacy abilities, we examined the normality of the distribution of the data. Shapiro-Wilk tests showed that data of character decision in all conditions and literacy abilities at both time points were not normally distributed (except for symmetrical high-complexity character decision, $p = 0.074$), $ps < 0.05$. Considering that parametric analysis of transformed data is a potentially better strategy than non-parametric analysis as it appears to be more powerful than the latter (Rasmussen and Dunlap, 1991), we normalized the data using rank-case transformation, which was considered to work better for small sample size in association tests than logarithm and Box-Cox transformations (Goh and Yap, 2009) and used normal scores obtained using Blom's formula in subsequent analyses.

TABLE 2 Descriptive statistics and correlation matrix of variables in Experiment 1 (upper panel) and Experiment 2 (lower panel).

		Mean (SD)	1	2	3	4	5	6	7	8	9	10
1	Age (month)	62.16 (9.72)										
2	Non-verbal IQ (max: 24)	13.73 (4.73)	0.67**									
3	Symmetrical, low-complexity CD	0.52 (0.32)	−0.15	−0.21								
4	Symmetrical, high-complexity CD	0.69 (0.25)	0.17	0.14	0.72**							
5	Asymmetrical, low-complexity CD	0.54 (0.30)	−0.03	−0.14	0.82**	0.77**						
6	Asymmetrical, high-complexity CD	0.60 (0.33)	0.10	0.08	0.71**	0.85**	0.76**					
7	Time 1 Chinese word reading (max:50)	13.57 (14.30)	0.48**	0.40*	−0.46*	−0.32	−0.34	−0.38*				
8	Time 1 Chinese word writing (max: 17)	2.80 (2.89)	0.67**	0.51**	−0.30	−0.11	−0.20	−0.21	0.74**			
9	Time 2 Chinese word reading (max: 50)	32.33 (11.55)	0.61**	0.72**	−0.44*	−0.13	−0.30	−0.21	0.81**	0.71**		
10	Time 2 Chinese word writing (max:17)	13.01 (5.59)	0.83**	0.77**	−0.16	0.22	−0.07	0.14	0.37*	0.58**	0.68**	

		Mean (SD)	1	2	3	4	5	6	7	8
1	Age (month)	63.96 (8.70)								
2	Non-verbal IQ (max: 24)	14.53 (4.84)	0.72**							
3	Symmetrical character learning	0.62 (0.28)	0.60**	0.51**						
4	Asymmetrical character learning	0.68 (0.27)	0.45*	0.37*	0.80**					
5	Time 1 Chinese word reading (max: 50)	13.93 (13.90)	0.56**	0.43*	0.50**	0.50**				
6	Time 1 Chinese word writing (max: 17)	3.10 (2.88)	0.64**	0.43*	0.61**	0.63**	0.79**			
7	Time 2 Chinese word reading (max: 50)	32.77 (12.80)	0.61**	0.65**	0.53**	0.54**	0.85**	0.73**		
8	Time 2 Chinese word writing (max: 17)	14.03 (5.09)	0.80**	0.71**	0.62**	0.60**	0.45*	0.52**	0.68**	

N=30. CD = character decision, * $p < 0.05$, ** $p < 0.01$.

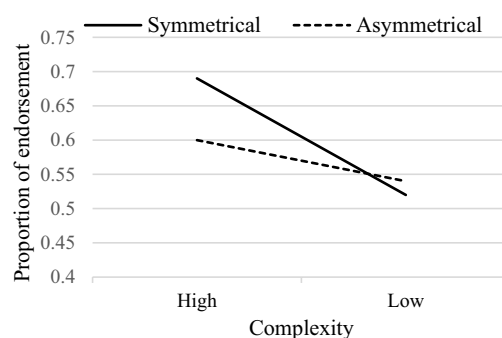
We conducted hierarchical regression analyses only on Chinese word reading because, as shown in Table 2, Chinese word writing was not significantly correlated with character decision in any condition concurrently or longitudinally ($ps > 0.113$). For Time-1 word reading, age and IQ were entered in the first step and the significant correlates of Time-1 word reading were entered in the second step. For Time-2 word reading, age and IQ were entered in the first step, Time-1 word reading was entered in the second step to control the autoregressive effect, and the significant correlate of Time-2 word reading was entered in the third step. Results show that character decision was not significantly associated with Chinese word reading at any time ($ps > 0.293$), suggesting that pure perception of symmetrical patterns in characters is not uniquely related to early word reading or writing abilities in Chinese.

Experiment 2

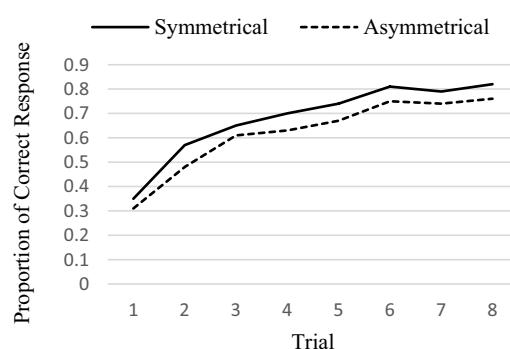
Table 2 also shows the descriptive statistics and correlation matrix of Experiment 2. As in Experiment 1, we first conducted GLMM analysis on the character learning data. The dependent variable was character learning accuracy (correct vs. incorrect, coded as 1 and 0, respectively), the random effects were child, set, and item, and the fixed effects were symmetry (symmetrical vs. asymmetrical), trial (1 to 8), and their interaction. The final best-fitting model had an overall correct-classification rate of 79.7%, $F(15, 2,384) = 18.59$, $p < 0.001$. As shown in Figure 1, there was a significant main effect of symmetry

($\beta = 0.74$, $SE = 0.36$, 95% CI [0.04, 1.44], $p = 0.039$), such that children learned to read symmetrical characters significantly better than asymmetrical ones. There was a significant main effect of trial [$F(7, 2,384) = 38.23$, $p < 0.001$], children improved across trials in both conditions. There was no significant interaction between symmetry and trial ($p = 0.921$). For the random effects, child was significant ($EST = 3.40$, $SEE = 0.97$, 95% CI [1.94, 5.95], $p < 0.001$), set and item were not significant, $ps > 0.510$. These results indicate that Chinese young children are sensitive to symmetry in characters when learning to read them.

Normal scores obtained from rank case transformation with Blom's formula were used in association analyses ($ps < 0.05$ in Shapiro–Wilk tests for all variables). Hierarchical regression analyses were conducted for Chinese word reading and word writing at Time 1 and Time 2, respectively, since character learning was significantly correlated with them in both conditions at both time points. In all models, age and IQ were entered in the first step. For Time-1 word reading or writing, significant correlates were entered in the second step. For Time-2 word reading or writing, Time-1 reading or writing (depending on the analysis) was entered in the second step, and the corresponding significant correlates of Time 2 were entered in the third step. Table 3 presents the results of the final models. Results show that for Chinese word reading, neither symmetrical nor asymmetrical character learning was uniquely associated with it in both time points ($ps > 0.24$). For Chinese word writing, asymmetrical character learning was a significant unique predictor of Time-2 word writing, accounting for 9% additional variance after age, IQ, and Time-1 word writing were statistically



Proportion of endorsement as possible Chinese in character decision as a function of visual complexity and symmetry



Proportion of correct response in character learning as a function of symmetry and trial

FIGURE 1

Illustration of symmetry effect in character decision in Experiment 1 (upper panel) and character learning in Experiment 2 (lower panel).

controlled ($\beta = 0.46$, $SE = 0.15$, $p = 0.017$). Asymmetrical character learning was also marginally significantly associated with Chinese word writing at Time 1 ($p = 0.052$). These results indicate that analytic perception of written characters, rather than pure sensitivity to symmetry in characters, is crucial for Chinese word writing ability, particularly longitudinally.

Discussion

This study examined Chinese preschoolers' sensitivity to symmetry in Chinese characters and its associations with Chinese word writing as well as word reading abilities. Corroborating the relative facility of symmetrical pattern recognition and production (Bornstein et al., 1981; Treiman and Kessler, 2011; Yin and Treiman, 2013; Fischer, 2018), we found that Chinese children averaging 5 years of age are sensitive to symmetry in characters: they endorsed symmetrical stimuli as possible Chinese more often than asymmetrical ones and they learned to read symmetrical characters better than asymmetrical ones. Extending existing literature, we found that pure sensitivity to symmetry in characters did not uniquely predict Chinese word reading or writing; rather, analytic perception of sub-character units, tapped in learning to read

asymmetrical characters, uniquely predict Chinese word writing concurrently and longitudinally.

Perceptual wholeness and information redundancy in symmetrical patterns (Chen and Huang, 1999) may explain children's sensitivity to symmetry in Chinese in this study. Li (1964) found that Chinese 6-year-olds perceived symmetrical parts in a character as one grapheme: when writing a character, they showed a tendency of first writing down the symmetrical parts in a character and then writing the remaining parts. Such perceptual wholeness of symmetrical parts may be especially conspicuous in high-complexity Chinese characters, which may partially explain why children were more likely to endorse high-complexity symmetrical stimuli as possible Chinese than low-complexity symmetrical stimuli in this study. Also, symmetrical parts contain similar or repetitive visual information. As reading involves mapping a visual form onto a spoken form, redundant information may facilitate perception of a symmetrical visual form and make linking it to a spoken form easier. This may explain why children in this study learned to read symmetrical characters better than asymmetrical ones.

Notably, learning to read asymmetrical characters, which necessitates analytic perception of sub-character components, uniquely predicted the 5-year-olds' Chinese word writing ability concurrently and longitudinally 2 years later. Learning to read asymmetrical characters may better tap the engagement of local

TABLE 3 Final model predicting Chinese word reading and Chinese word writing at Time 1 and Time 2 from character learning at Time 1.

Outcome and predictor	R^2	ΔR^2	ΔF	b	SE b	β	p value
Time 1 word reading							
Age	0.31	0.31	6.18*	0.51	0.31	0.40	0.118
IQ	–	–	–	0.01	0.04	0.03	0.899
Time 1 symmetrical character learning	0.39	0.08	1.60	0.00	0.29	0.00	0.999
Time 1 asymmetrical character learning	–	–	–	0.30	0.25	0.31	0.238
Time 1 word writing							
Age	0.42	0.42	9.61**	0.67	0.26	0.54*	0.015
IQ	–	–	–	–0.02	0.04	–0.13	0.519
Time 1 symmetrical character learning	0.57	0.16	4.54*	–0.00	0.23	–0.00	0.994
Time 1 asymmetrical character learning	–	–	–	0.42	0.21	0.44	0.052
Time 2 word reading							
Age	0.46	0.46	11.49***	–0.11	0.18	–0.09	0.536
IQ	–	–	–	0.08	0.02	0.40**	0.003
Time 1 word reading	0.83	0.37	56.19***	0.71	0.11	0.70***	0.000
Time 1 symmetrical character learning	0.84	0.01	0.74	–0.10	0.15	–0.11	0.501
Time 1 asymmetrical character learning	–	–	–	0.17	0.14	0.17	0.244
Time 2 word writing							
Age	0.67	0.67	27.71***	0.70	0.20	0.63**	0.002
IQ	–	–	–	0.04	0.03	0.24	0.125
Time 1 word writing	0.67	0.00	0.05	–0.18	0.14	–0.20	0.213
Time 1 symmetrical character learning	0.76	0.09	4.40*	–0.11	0.16	–0.13	0.498
Time 1 asymmetrical character learning	–	–	–	0.39	0.15	0.46*	0.017

$N=30$. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

attention to components for greater precision than learning to read symmetrical characters and such precision is called for more in production (writing) than in recognition (reading). Thus, although symmetrical character learning was significantly correlated with Chinese word reading and word writing at both time points, it was not uniquely associated with them after age, IQ, and autoregressive effects were statistically controlled. From a novel perspective of symmetry sensitivity, this study contributes to current literature on the associations between breaking mirror invariance in print (such as shape orientation sensitivity) and literacy acquisition by showing that such link manifests not only in reading (Fernandes et al., 2016; Zhang et al., 2021) but more particularly in writing. This study also adds to the recent demonstrations by Fischer and Luxembourg (2022) of the difference between a recognition task (e.g., distinguishing between the reversible letters of b and d) and a writing task (e.g., reversing letters or digits in writing), the two tasks often considered as equivalent in the literature on mirror invariance but actually found to be at least partly independent processes (having a negative relationship between the two rates) in French 6 year olds. Future studies may further examine this issue with Chinese young children.

A limitation of the present study is that we were unable to disentangle visual-perceptual skill from orthographic knowledge that children might be using when making character decision or learning to read characters, despite the very low frequency of stimuli used.

Future studies should use pseudocharacters and take identifiable subcomponents and configurations into consideration. Another limitation is the small sample size in each experiment despite the statistical adjustments we made before running association analyses. Future work should involve a larger sample with diverse language backgrounds to examine the generalizability of the symmetry effect across writing systems.

Despite the limitations, the present research provides novel evidence that young children are sensitive to symmetry in writing, but it is the analytic perception of sub-word units rather than pure sensitivity to perceptual symmetry that is significantly related to early literacy, particularly writing.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Tsinghua University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

LY and CM conceived of the presented idea and collaborated in analyzing the data and drafting the article. LY conducted the data collection. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from National Natural Science Foundation of China (Grant no. 81371497) and Tsinghua University Initiative Scientific Research Program (Grant no. 2022THZWC11) awarded to LY.

References

- Bornstein, M. H., Ferdinandsen, K., and Gross, C. G. (1981). Perception of symmetry in infancy. *Dev. Psychol.* 17, 82–86. doi: 10.1037/0012-1649.17.1.82
- Chen, C., and Huang, X. (1999). Research on characteristics of visual recognition to symmetrical structural Chinese characters. *Acta Psychol. Sin.* 31, 154–161. (in Chinese)
- China National Working Committee on Languages and Characters (1988). *List of common characters in modern Chinese*. Beijing, China: People's Republic of China.
- Dehaene, S., Nakamura, K., Jobert, A., Kuroki, C., Ogawa, S., and Cohen, L. (2010). Why do children make mirror errors in reading? Neural correlates of mirror invariance in the visual word form area. *NeuroImage* 49, 1837–1848. doi: 10.1016/j.neuroimage.2009.09.024
- Fernandes, T., Leite, I., and Kolinsky, R. (2016). Into the looking glass: literacy acquisition and mirror invariance in preschool and first-grade children. *Child Dev.* 87, 2008–2025. doi: 10.1111/cdev.12550
- Fischer, J. P. (2018). Studies on the written characters orientation and its influence on digit reversal by children. *Educ. Psychol.* 38, 1–16. doi: 10.1080/01443410.2017.1359239
- Fischer, J. P., and Luxembourg, C. (2022). Typical 6-year-old children's confusion between “b” and “d” in reading cannot be assimilated to reversal. *Read. Writ.* 35, 2433–2451. doi: 10.1007/s11145-022-10290-6
- Goh, L., and Yap, V. B. (2009). Effects of normalization on quantitative traits in association test. *BMC Bioinformatics* 10:415. doi: 10.1186/1471-2105-10-415
- Hsiao, J. H., and Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist. *Psychol. Sci.* 20, 455–463. doi: 10.1111/j.1467-9280.2009.02315.x
- Huang, C., Chen, C., and Yu, H. (2002). The processing of structurally symmetrical Chinese character recognition (I): an experimental research on character-level process. *Psychol. Sci.* 25, 4–6. doi: 10.16719/j.cnki.1671-6981.2002.01.002 (in Chinese)
- Kessler, B., Pollo, T. C., Treiman, R., and Cardoso-Martins, C. (2013). Frequency analyses of prephonological spellings as predictors of success in conventional spelling. *J. Learn. Disabil.* 46, 252–259. doi: 10.1177/0022219412449440
- Lavine, L. O. (1977). Differentiation of letterlike forms in prereading children. *Dev. Psychol.* 13, 89–94. doi: 10.1037//0012-1649.13.2.89
- Levin, I., and Bus, A. G. (2003). How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. *Dev. Psychol.* 39, 891–905. doi: 10.1037/0012-1649.39.5.891
- Li, Z. (1964). A study of Chinese grapheme perception unit in 6-year-old children. *Acta Psychol. Sin.* 2, 178–184. (in Chinese)
- McBride-Chang, C., and Treiman, R. (2003). Hong Kong kindergarteners learn to read English analytically. *Psychol. Sci.* 14, 138–143. doi: 10.1111/1467-9280.t01-1-01432
- Pollo, T. C., Kessler, B., and Treiman, R. (2009). Statistical patterns in children's early writing. *J. Exp. Child Psychol.* 104, 410–426. doi: 10.1016/j.jecp.2009.07.003
- Rasmussen, J. L., and Dunlap, W. P. (1991). Dealing with nonnormal data: parametric analysis of transformed data vs nonparametric analysis. *Educ. Psychol. Meas.* 51, 809–820. doi: 10.1177/001316449105100402
- Raven, J. C., Court, J. H., and Raven, J. (1996). *Standard progressive matrices*. Oxford, UK: Oxford Psychologists Press
- Shu, H., Chen, X., Anderson, R. C., Wu, N., and Xuan, Y. (2003). Properties of school Chinese: implications for learning to read. *Child Dev.* 74, 27–47. doi: 10.1111/1467-8624.00519
- Treiman, R. (2017). Learning to spell words: findings, theories, and issues. *Sci. Stud. Read.* 21, 265–276. doi: 10.1080/10888438.2017.1296449
- Treiman, R., and Kessler, B. (2011). Similarities among the shapes of writing and their effects on learning. *Writ. Lang. Lit.* 14, 39–57. doi: 10.1075/wll.14.1.03tre
- Wang, Y., Yin, L., and McBride, C. (2015). Unique predictors of early reading and writing: a one-year longitudinal study of Chinese kindergarteners. *Early Child. Res. Q.* 32, 51–59. doi: 10.1016/j.ecresq.2015.02.004
- Xu, N. (2012). Symmetrical structures of modern Chinese characters. *J. Longyan Univ.* 30, 39–43. doi: 10.16813/j.cnki.cn35-1286/g4.2012.04.007 (in Chinese)
- Yang, L. Y., Guo, J. P., Richman, L. C., Schmidt, F. L., Gerken, K. C., and Ding, Y. (2013). Visual skills and Chinese reading acquisition: a meta-analysis of correlation evidence. *Educ. Psychol. Rev.* 25, 115–143. doi: 10.1007/s10648-013-9217-3
- Yin, L., and McBride, C. (2015). Chinese kindergartners learn to read characters analytically. *Psychol. Sci.* 26, 424–432. doi: 10.1177/0956797614567203
- Yin, L., and Treiman, R. (2013). Name writing in mandarin-speaking children. *J. Exp. Child Psychol.* 116, 199–215. doi: 10.1016/j.jecp.2013.05.010
- Yu, B., and Cao, H. (1992). How distributions of numbers of strokes affect the cognition of Chinese characters. *Psychol. Sci.* 4, 5–10. (in Chinese)
- Zhang, W., Ni, A., and Li, S. (2021). Development of orientation sensitivity to Chinese compound characters in 4-to-6-year-old children. *Psych J.* 10, 33–46. doi: 10.1002/pchj.405
- Zhang, L., Yin, L., and Treiman, R. (2017). Chinese children's early knowledge about writing. *Br. J. Dev. Psychol.* 35, 349–358. doi: 10.1111/bjdp.12171

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Lénia Carvalhais,
Infante D. Henrique Portugalense University,
Portugal
Vassilios Papadimitriou,
Hellenic Open University, Greece
George Manolitsis,
University of Crete, Greece

*CORRESPONDENCE

Georgia Z. Niolaki
✉ g.niolaki@bathspa.ac.uk

[†]These authors have contributed equally to this work and share first authorship

[†]These authors share last authorship

RECEIVED 02 March 2023

ACCEPTED 20 April 2023

PUBLISHED 11 May 2023

CITATION

Niolaki GZ, Negoita A, Vousden JI, Terzopoulos AR, Taylor L and Masterson J (2023) What spelling errors can tell us about the development of processes involved in children's spelling. *Front. Psychol.* 14:1178427. doi: 10.3389/fpsyg.2023.1178427

COPYRIGHT

© 2023 Niolaki, Negoita, Vousden, Terzopoulos, Taylor and Masterson. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

What spelling errors can tell us about the development of processes involved in children's spelling

Georgia Z. Niolaki^{1*†}, Alexandra Negoita^{2†}, Janet I. Vousden^{3†}, Aris R. Terzopoulos^{4†}, Laura Taylor^{5†} and Jackie Masterson^{6†}

¹School of Education, Bath Spa University, Bath, United Kingdom, ²School of Psychological, Social and Behavioral Sciences, Coventry University, Coventry, United Kingdom, ³Department of Psychology, Nottingham Trent University, Nottingham, United Kingdom, ⁴School of Psychology, Birmingham City University, Birmingham, United Kingdom, ⁵School of Psychology, Northampton University, Northampton, United Kingdom, ⁶Department of Psychology, UCL, Institute of Education, London, United Kingdom

Introduction: Spelling is an essential foundation for reading and writing. However, many children leave school with spelling difficulties. By understanding the processes children use when they spell, we can intervene with appropriate instruction tailored to their needs.

Methods: Our study aimed to identify key processes (lexical-semantic and phonological) by using a spelling assessment that distinguishes different printed letter strings/word types (regular and irregular words, and pseudowords). Misspellings in the test from 641 pupils in Reception Year to Year 6 were scored using alternatives to binary correct versus incorrect scoring systems. The measures looked at phonological plausibility, phoneme representations and letter distance. These have been used successfully in the past but not with a spelling test that distinguishes irregularly spelled words from regular words and pseudowords.

Results: The findings suggest that children in primary school rely on both lexical-semantic and phonological processes to spell all types of letter string, but this varies depending on the level of spelling experience (younger Foundation/Key stage 1 and older Key stage 2). Although children in younger year groups seemed to rely more on phonics, based on the strongest correlation coefficients for all word types, with further spelling experience, lexical processes seemed to be more evident, depending on the type of word examined.

Discussion: The findings have implications for the way we teach and assess spelling and could prove to be valuable for educators.

KEYWORDS

spelling, phonological plausibility, letter distance, automated measures of phonemes representations, primary age students

Introduction

Learning to spell is a lengthy and challenging process, and it is not surprising that some children struggle and need extra support (Nag et al., 2018). Therefore, it is important to find accurate and sensitive methods for assessing children's spellings (Treiman and Kessler, 2004). One way to achieve this is to move away from the traditional binary correct vs. incorrect scoring method. Research has suggested that looking at children's misspellings beyond the dichotomous scoring system might reveal patterns of development in linguistic knowledge (Masterson and Apel, 2010; de Bree and van

den Boer, 2019; Treiman et al., 2019; for non-English orthographies see also Niolaki and Masterson, 2012; Carvalhais et al., 2020, 2021) and underlying cognitive processes that children use when they spell (Caravolas et al., 2001; Kessler, 2009; Treiman et al., 2016). These processes involve perception, attention, memory and expertise, all of which past research has explored in considerable depth (Georgiou et al., 2012; Rønneberg and Torrance, 2019; Zoccolotti et al., 2020; Niolaki et al., 2022). As such, to date, spelling has been primarily studied more in relation to cognitive abilities than to the actual spelling errors made. Recent research has suggested that it is worth looking in more depth at the characteristics of spelling attempts with different types of words (i.e., regular and irregular words and pseudowords; de Bree and van den Boer, 2019; Niolaki et al., 2020).

Dual route (DR) models of spelling have been proposed to explain how information is processed when spelling words (Barry, 1994). Similarly, models such as the integration of multiple patterns (IMP) model that support statistical learning (which is considered a powerful strategy for generalizing learning from untaught items) agree that both sublexical and lexical/semantic processes exist. The IMP model also recognizes that items can be divided into regularly and irregularly spelled words (Treiman, 2018, p. 648). These models suggest that people use sound-to-spelling rules, or *sublexical processing*, as well as stored knowledge of whole-word spellings, or *lexical processing* (Barry, 1994; Martin and Barry, 2012). Children are believed to possess the basis for lexical processing before they start to spell, as a result of spoken language and early exposure to printed text (Barry, 1994). Children are also believed to memorize the orthographic rules and learn to spell correctly as they gain more experience with print. As a result, children perform better on words that occur frequently in books (Kessler et al., 2012). A more pronounced frequency effect has been observed for irregularly spelled words (such as *DEBT*, *YACHT*, and *MORTGAGE*), indicating that children need to hear, read and write these words more times than regularly spelled words (such as *BARGE*, *SLATE*, and *TARGET*) in order to create a correct representation in memory (Kroese et al., 2000).

Empirical evidence suggests that novice spellers rely heavily or even exclusively on the sub-lexical route (Caravolas et al., 2001), and they progressively develop the lexical-semantic route. In the UK, phonics training is the starting fuel for children to understand how phoneme-grapheme correspondences operate. This is important as phonemes in English can have several spelling options (Barry and Seymour, 1988), for example, /k/ can be spelled <c>, <k>, or <ck>. The options vary in probability of occurrence, and this depends in part on whether context-free or context-sensitive probabilities are considered (context sensitive as in <magic—>magician>; Spencer, 2009). Overall, spelling encompasses different tiers of linguistic awareness (phonology, orthography, morphology and semantics), so to establish solid lexical representations, the spellers need to be aware of all these and their interrelationships.

To gain insight into the processes children use in spelling, sensitive, non-binary scoring systems have been developed to examine spelling errors (e.g., Bruck and Waters, 1988; Treiman and Kessler, 2004; Kessler, 2009). These systems have also been found to have good discriminatory power for students struggling with spelling (Treiman et al., 2016; for non-English languages see Niolaki et al., 2014; Joye et al., 2020). The non-binary measures rely, for example, on how close the misspelled word is to the correct spelling and can be used as a fine-grained measure to monitor the progress a child is making in developing spelling skills (Masterson and Apel, 2010). Many children, especially in the Foundation Year, produce more errors than correct

responses (Georgiou et al., 2020; Carvalhais et al., 2021), and the measures can provide us with a clear picture of the use of lexical-semantic and sublexical processes, and how children's reliance on these processes changes as they progress in spelling skill. Also, error analysis is not affected by floor effects which is an issue when we solely look at accuracy especially for young spellers (Treiman et al., 2019). We next present the measures we included in the current investigation.

Phonologically plausible errors (e.g., *elephant* spelled <elefant>) are considered to reflect use of sublexical spelling processes and are particularly apparent in novice spellers. Caravolas et al. (2001) analysed spellings from 153 children in the three first years in school in the UK. They emphasized the critical predictive role of phonological spelling ability for later reading and spelling accuracy. The researchers argued that children need phonics training (structured instruction that helps children to spell unfamiliar words) to build a solid sound-letter mapping system and then to become skilled spellers who competently use orthographic rules (Caravolas et al., 2001).

In a longitudinal and cohort study with 95 Portuguese students from two age groups—Grades 4–7 and 6–9, Carvalhais et al. (2021) conducted spelling error analyses and looked at phonological plausibility as one of the critical variables. They found that younger children made more misspellings than the older children and phonologically inappropriate errors were less in the older group. These findings are consistent with similar observations made in the English orthography (Treiman and Bourassa, 2000; Caravolas et al., 2001). The research highlights the importance of phoneme-grapheme associations in the earliest stages of spelling in English.

In the current study we employed two separate measures of phonological plausibility—a binary phonological plausibility score (PhP, e.g., spelling *elephant* as <elefant>), and an automatized continuous measure, the Automated Measure of Phoneme Representation (AMPR). The main difference between PhP and AMPR is the first is a binary measure (phonologically plausible error or not) and is hand scored (by the research team), whereas the second provides a score computed across the word. For AMPR, values closer to zero indicate a lower quality of error, as the target is distant from the response (Treiman and Kessler, 2004), meaning that as children's spelling skill develops the AMPR score should increase. We also used letter distance (LD)¹ in our analyses to capture the number of letter additions, deletions and substitutions needed to create the correct phonological and orthographic spelling from an error.

Treiman et al. (2016) investigated a range of scoring measures in a longitudinal study with children from Kindergarten to Grade 2. The findings indicated that children possess some phonological knowledge early in spelling development, however, LD, the lexically-related measure, proved a better predictor of spelling accuracy in beginner spellers than the phoneme distance measure, the sublexically-related one (Treiman et al., 2016). We considered that it would be informative to see, in our study with UK children, and almost 21 years after the implementation of synthetic phonics in schools, whether PhP or LD would be more strongly associated with emergent spelling, and whether the association would differ for regular words, irregular words and pseudowords, as these are assumed to draw on different

¹ LD: calculated using Ponto <http://spell.psychology.wustl.edu/ponto/> (Kessler, 2009).

processes, that is, whole-word and sublexical. For LD, values that are closer to zero indicate less distance from the correct spelling. Thus, as children develop their spelling knowledge LD values should decrease.

Several studies have reported that more sensitive scoring methods can effectively capture developmental changes in spelling and strong associations with reading and phonological ability (Ritchey et al., 2009; Clemens et al., 2014; Frisby, 2016). As noted above, Treiman et al. (2016) investigated a range of measures in a longitudinal study. Participants were 374 children from kindergarten age to Grade 2 in the USA and Australia. The researchers employed letter-based measures (LD, correctness, letter sequence) and phoneme-based measures (AMPR, phoneme distance, sound-spelling) at two-time points. The findings revealed that letter-based measures accounted for more variance in spelling accuracy than phoneme-based measures. Treiman et al. (2019) reported a replication study with British English spellers where correctness (a binary correct/incorrect measure) was more predictive of single word spelling at Time 4 assessment (when the children attended the spring term of Year 2) than non-binary measures. Binary measures might be useful for predicting who is likely to struggle with spelling but would not tell us why. In this case non-binary measures would give us more insight.

Carvalho et al. (2020) in their study with Portuguese students (also reported earlier) found that orthographically related variables (such as stress mark errors and orthographic misspellings) were the most common errors in older learners (Grades 4, 6, and 7). In the English-based studies reported above children were tested in the earlier school grades, so there is not much evidence for English on how the phonological- and orthographic-related measures perform in older spellers.

In summary, it has not been established if developmental changes in spelling errors apply equally to regular and irregular words and pseudowords. Moreover, several scoring methods have been found to relate to one another, for example, PhP and AMPR (Treiman et al., 2019), yet they have not been investigated concerning spelling errors with different letter string/item types. The current study addressed these gaps in the literature by exploring patterns of spelling errors in primary school children split into three age groups, Foundation Year/Key Stage 1 (F/KS1; Kindergarten to grade 2), Early Key Stage 2 (EKS2; Grade 3 to 4) and Advanced Key Stage 2 (AKS2; Grade 5–6). Thus, the goal of the present study was to investigate the strategies children use when they spell different types of words, through analysis of their errors, and how this may change from reception year to year 6 (5- to 12-years).

The current study

We aimed to investigate whether children rely on different processes for different types of words in their spelling, and explored whether this reliance might change from Reception year to Year 6 (equivalent to Kindergarten to Grade 6 in the USA). Research has indicated that both lexical-semantic and sub-lexical processes are employed for spelling by beginner spellers, while the former processes seem to become more prevalent when children gain more experience with reading and spelling (Bruck and Waters, 1988; Caravolas et al., 2001; Kessler et al., 2012; Share, 1995). Treiman et al. (2016, 2019) found that lexically related variables were strong predictors of more advanced spelling (however, the older children only went up to Grade/Year 2). When investigating the spelling of different types of letter strings, it is likely that children rely more on sub-lexical processes for pseudowords, and lexical processes for irregular

words, while the spelling of regular words will tap both processes (Niolaki et al., 2022). Based on the research reviewed above, we aimed to test the following hypotheses. We expected that the sub-lexically related (phoneme-based) measure AMPR would be more important for spelling pseudowords than irregular words. We also anticipated that the sublexically- and lexically-related (letter based) measure LD would be important for real words; but less strongly associated with pseudoword spelling.

Specifically, we made the following predictions:

- There will be a significant interaction between key stages and word types for spelling accuracy, letter-based and phoneme-based measures. Lower key stages should evidence lower scores on accuracy and phoneme-based scores and higher scores on letter-based measures for each word type than higher key stages.
- There will be significant associations between phoneme-based and letter-based measures and accuracy for all word types in all key stages. The strength of associations will vary depending on key stage and letter string type.

Methods

Participants

Participants were 641 UK primary school children attending Reception Year to Year 6, from a mix of urban and rural schools (seven different state schools). School years were divided into three levels as follows: Group 1 comprised children in Foundation /Key Stage 1 (F/KS1), i.e., Reception Year to Year 2, Group 2 were children in the first half of Early Key Stage 2 (EKS2), i.e., Year 3 and 4, and Group 3 were children in the second half of Advanced Key Stage 2 (AKS2), i.e., Year 5 and 6. Pupils were grouped this way to allow for the identification of strategies used for spelling in developmental stages (beginning, early and late stages/advanced spellers). The number of children in each age group, together with their mean chronological age, is shown in Table 1. Data were collected after parents/carers of participating students returned consent forms and after children assented to participate in the study. The University Ethics' Committee granted ethical approval for the study.

Materials

Data from a new interpretive spelling test for primary school children were collected. The test consists of three sections comprising 36 irregular words (e.g., <yacht>), 36 regular words (e.g., <cat>), and 34 pseudowords (e.g., <trelishly>). Spelling regularity was calculated based on the frequency of occurrence of sound-letter correspondences

TABLE 1 Number and mean chronological age of participants per age group (standard deviations are in parentheses).

	F/KS1	EKS2	AKS2
Number of participants	309	165	167
Age (years)	6.34 (0.95)	8.77 (0.94)	10.78 (0.56)

in the word (Spencer, 2009; Vousden et al., 2011) but also spelling instruction in UK schools. Pseudowords were formed by combining the first half of a regular word and the last half of another regular word. When spelled, these items follow regular words' structure, but they were unfamiliar to the children. For pseudoword accuracy (using binary scoring) we categorized any plausible spelling as accurate. The first and third authors agreed on the phonologically plausible acceptable responses for the 36 pseudowords (for example the item <clep> was phonologically plausible if spelled as <clepp> and <klep>).

Items were matched across the three subtests on word length, and across regular and irregular words on zipfrequency and zipf contextual diversity (see Niolaki et al., 2022,² for details). The reliability of each subtest is high based on the accuracy scores of the binary assessment: irregular words $\alpha=0.97$, regular words $\alpha=0.96$, and pseudowords $\alpha=0.94$. Each spelling response was scored using binary, non-binary and categorical measures. Table 2 provides examples of scoring with the different measures.

Accuracy

Children's spellings were given a score of zero for incorrect and one for correct spellings. The maximum possible accuracy score was 106.

Orthographic measure

Letter distance

Letter distance was calculated using Ponto [available online at <http://spell.psychology.wustl.edu/ponto/> (Kessler, 2009)]. This online tool allocates points for each deletion, addition, transposition or substitution needed for the child's written response to be transformed into the conventional spelling. Mean distance scores were generated for each child for regular words, irregular words and pseudowords separately.

Phonological measures

Phonological plausibility

An error was given a score of one if it was phonologically plausible and zero if there were incorrect phoneme-grapheme correspondences (phonemic error), or if additional elements were present or absent. Errors were calculated as PhP or not by the second author and agreed by the first and third authors.

Automated measures for phoneme representation

The AMPR scoring metric comprises the number of phonologically plausible phonemes in a word divided by the total number of phonemes. The AMPR calculates a lower score when phonologically implausible errors are made, for example, *life* spelled

TABLE 2 Examples of scoring using all scoring measures.

Target word	Child's spelling	Accuracy	PhP	AMPR	LD
<i>Life</i>	<i>Live</i>	0	0	0.66	1.4
<i>Nature</i>	<i>Nocher</i>	0	0	0.66	6.2
<i>Flavor</i>	<i>Flaver</i>	0	1	1	2.4
<i>Aspire</i>	<i>Aspier</i>	0	1	1	2

PhP, phonological plausibility; AMPR, automated measure for phoneme representation; LD, letter distance.

as *live* would receive a score of 0.66 (2/3 phonemes correctly represented) while *life* spelled *live* would receive a score of 1 (3/3 phonemes correctly represented, the highest score). The measure was generated using the online software (available at <http://spell.psychology.wustl.edu/AMPR>), yielding scores between 0 and 1, where scores nearer 1 represent a phonologically plausible response and scores closer to zero represent a non-phonologically plausible response (Treiman and Kessler, 2004; see Table 3). Mean scores for each word type per child were calculated.

Procedure

The data were collected in Spring/Summer term by researchers and trained research assistants testing children in small groups in urban and rural primary schools (at least three research assistants supervised the children, and the class teacher also helped during the procedure). For details of the procedure please see Niolaki et al. (2022). Error analysis was computer-scored or hand-scored, as described above, by the authors. In the case of scoring by hand, categorisations were discussed and agreed upon between the first, second and third authors. Agreement between authors had to reach 100% for hand-scored variables.

Results

Total correct scores were computed for all participants for regular words, irregular words and pseudowords. Average scores for AMPR and LD were computed. Means and standard deviations per measure and word type are presented in Table 3.

A two-way Analysis of Variance (ANOVA) was conducted to explore Key stage group differences in accuracy. The between groups variable was Key stage group (F/KS1 (Reception Year to Year 2), EKS2 (Year 3 and Year 4), AKS2 (Year 5 and Year 6)), and the within groups variable was word type (regular word, irregular word, pseudoword). There was a significant interaction between word type and Key stage groups [$F(4, 1,240) = 76.44, p < 0.001, \eta_p^2 = 0.198$]. While pseudowords were spelled more accurately than irregular words in F/KS1 and EKS2, this reversed in AKS2 (see Figure 1A). Regular words were the most accurate in all key stages. There was a significant main effect of word type [$F(2, 1,240) = 402.7, p < 0.001, \eta_p^2 = 0.394$]—the children were more accurate in spelling regular words ($M = 23.8$) than irregular words ($M = 19.9$) and pseudowords ($M = 17.9; p < 0.001$). There was also a significant main effect of group [$F(2, 620) = 423.6, p < 0.001, \eta_p^2 = 0.577$], with AKS2 children spelling significantly more items accurately overall ($M = 28.0$) than

² A list of the items is in the appendix of the published paper Niolaki et al. (2022).

TABLE 3 Means for all scoring measures (standard deviations are in parentheses).

		Accuracy%	PhP%	AMPR	LD
Year_ groups	Type of word				
	Irregular	17.1 (19.9)	26 (18)	0.78 (0.19)	3.28 (1.55)
F/KS1	Regular	38.5 (25.9)	16 (13.5)	0.77 (0.18)	1.86 (1.57)
	Pseudoword	34.4 (22.7)	-	0.72 (0.17)	2.12 (1.66)
	Irregular	54.8 (23.1)	42.3 (20.8)	0.89 (0.07)	1.31(0.88)
EKS2	Regular	74.2 (18.8)	23.7 (21)	0.86 (0.09)	0.57 (0.58)
	Pseudoword	62.7 (17.1)	-	0.86 (0.26)	0.83 (0.65)
	Irregular	76.9 (17.2)	61.8 (24.9)	0.95 (0.04)	0.54 (0.47)
AKS2	Regular	86.4 (11.4)	34.1 (29.7)	0.91 (0.06)	0.24 (0.22)
	Pseudoword	70.1 (12.8)	-	0.85 (0.07)	0.61 (0.37)

PhP, phonological plausibility; LD, letter distance; AMPR, automated measure for phoneme representation.

EKS2 ($M = 23.0$) and F/KS1 ($M = 10.6$; all $p_s < 0.001$). The difference between EKS2 and F/KS1 was also significant.

The same analysis was carried out for the PhP, AMPR and LD scores (see **Figures 1B–D** for plots of the mean scores). For PhP the scores were for regular and irregular words only as all phonologically plausible responses for pseudowords were counted as correct. The results of the two-way ANOVA revealed a significant interaction of word type and Key stage group [$F(2, 601) = 23.3$, $p < 0.001$, $\eta_p^2 = 0.07$]. The main effect of word type was significant [$F(1, 601) = 341.9$, $p < 0.001$, $\eta_p^2 = 0.36$] as children made more PhP errors for irregular ($M = 42.1$) than regular words ($M = 24.1$). The main effect of group was significant [$F(2, 601) = 125.8$, $p < 0.001$, $\eta_p^2 = 0.29$]. The difference between F/KS1 ($M = 46$) and EKS2 ($M = 23$) and AKS2 ($M = 11$) was significant, as was that between EKS2 and AKS2.

For the AMPR scores, the interaction of word type and key stage group was significant [$F(4, 1,162) = 5.37$, $p < 0.005$, $\eta_p^2 = 0.018$]. AMPR scores increased from F/KS1 to AKS2 for each word type except pseudowords. In the case of pseudowords, scores increased from F/KS1 to EKS2 only. There was a significant main effect of word type [$F(2, 1,162) = 64.8$, $p < 0.001$, $\eta_p^2 = 0.100$] with significantly lower scores for pseudowords ($M = 0.81$) than regular words ($M = 0.85$) and irregular words ($M = 0.88$; $p_s < 0.01$ for all differences). The difference between regular and irregular words was also significant ($p < 0.001$). There was also a main effect of group [$F(2, 581) = 65.06$, $p < 0.001$, $\eta_p^2 = 0.183$]. F/KS1 had significantly lower scores ($M = 0.766$) than ESK ($M = 0.875$) and ASK2 ($M = 0.905$) but the difference between ESK2 and ASK2 did not reach significance.

Finally, for LD, there was a significant interaction between word type and key stage group [$F(4, 1,220) = 125.9$, $p < 0.001$, $\eta_p^2 = 0.29$]. LD scores decreased from F/KS1 to EKS2 for each word type. Irregular words had the highest score but in AKS2 pseudowords scored higher. There was a significant main effect of word type [$F(2, 1,220) = 525.7$, $p < 0.001$, $\eta_p^2 = 0.46$]. There were significantly higher scores for irregular words ($M = 1.71$) than regular words ($M = 0.91$) and pseudowords ($M = 1.21$; all $p_s < 0.001$). The difference was also significant between regular words and pseudowords ($p < 0.001$). There was a significant main effect of group [$F(2, 610) = 197.4$, $p < 0.001$, $\eta_p^2 = 0.39$]. F/KS1 had significantly higher scores ($M = 2.45$) than EKS2 ($M = 0.91$) and AKS2 ($M = 0.47$; all $p < 0.001$). The difference between EKS2 and AKS2 was also significant ($p < 0.001$).

Correlation analyses

Partial correlations were conducted for each KS group, controlling for grade, in order to explore the associations of the phoneme and letter-based measures with spelling accuracy for regular words, irregular words and pseudowords for AMPR and LD, and for regular and irregular words for PhP. The results are presented in **Table 4**.

For the F/KS1 group, significant associations were observed between accuracy for all word types and all measures. For EKS2, all associations were significant, with the exception that AMPR pseudoword scores were not associated with pseudoword accuracy. Notably, for AKS2, associations with PhP and AMPR for regular words are lower than for the youngest age group. High levels of accuracy could explain these weaker associations in the older age groups. However, LD was consistently strongly correlated with accuracy across the age groups and letter string types.

In order to determine differences in the associations Eid et al.'s (2011) comparison of correlations (online calculator³) was used. LD and AMPR were compared, as PhP was less strongly associated with accuracy according to the results of the correlations (see **Table 4**).

Comparison of correlations for F/KS1

The relationship between irregular word accuracy and AMPR irregular word score ($r = 0.290$) was significantly less strong than that for regular word accuracy and AMPR regular word score ($r = 0.529$; $z = 5.1$, $p < 0.001$). The association between regular word accuracy and AMPR regular word score ($r = 0.529$) was significantly lower than that for pseudoword accuracy and AMPR pseudoword score ($r = 0.606$; $z = 1.9$, $p < 0.05$).

There were no significant differences in the associations between LD scores and accuracy for regular vs. irregular and pseudoword vs. irregular. Unexpectedly, the coefficient for LD regular words and regular word accuracy ($r = -0.682$) was significantly less strong than the coefficient for LD pseudowords and pseudoword accuracy ($r = -0.751$; $z = 2.4$, $p < 0.01$).

³ Eid et al. (2011): <https://www.psychometrica.de/correlation.html#dependent>.

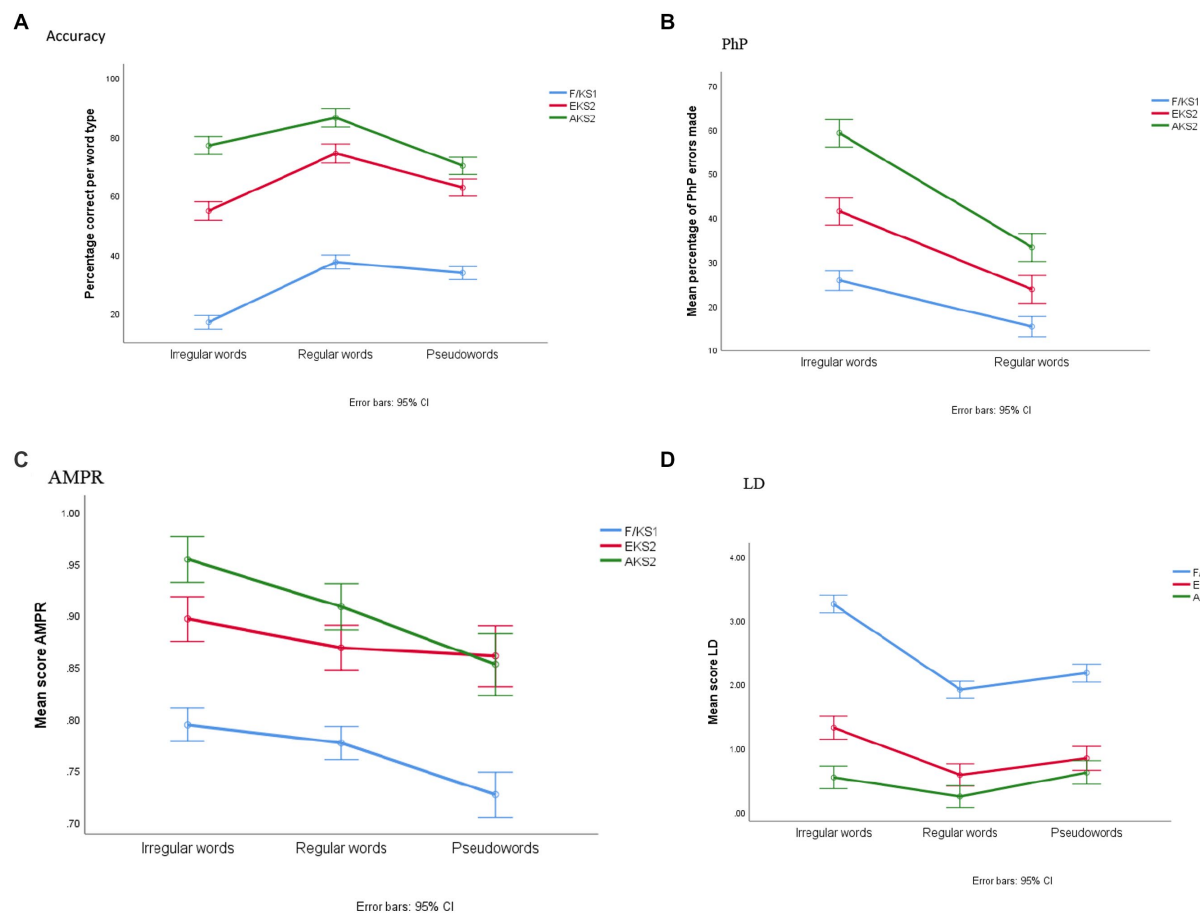


FIGURE 1

Mean accuracy, phonological plausibility (PhP), automated measure for phoneme representation (AMPR), and letter distance (LD) score for each word category and group.

TABLE 4 Correlations between spelling accuracy and scoring metrics controlling for year_group.

Scoring metric	F/KS1 accuracy			EKS2 accuracy			AKS2 accuracy		
	Ir	Rg	Pw	Ir	Rg	Pw	Ir	Rg	Pw
PhP Ir	0.391***	-		0.435***	-		0.446***	-	
PhP Rg	-	0.420***		-	0.334***		-	0.202*	
AMPR Ir	0.290***	-	-	0.456***	-	-	0.411***	-	-
AMPR Rg	-	0.529***	-	-	0.237**	-	-	0.187*	-
AMPR Pw	-	-	0.606***	-	-	-0.073	-	-	0.170*
LD Ir	-0.719***	-	-	-0.951***	-	-	-0.956***	-	-
LDRg	-	-0.682***	-	-	-0.943***	-	-	-0.930***	-
LD Pw	-	-	-0.751***	-	-	-0.882***	-	-	-0.927***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Ir, irregular; Rg, regular; Pw, pseudowords; PhP, phonological plausibility; AMPR, automated measure for phoneme representation; LD, letter distance.

Comparison of correlations for EKS2

The coefficient for irregular word accuracy and AMPR irregular words ($r = 0.456$) was significantly higher than that for regular word

accuracy and AMPR regular words ($r = 0.237$; $z = 3.2$, $p < 0.001$). This is different to what we found for F/KS1 children. The difference between the coefficient for LD irregular words/irregular word accuracy and the coefficient for LD regular word/regular word

accuracy was not significant, indicating reliance on lexical processes for both word types. The coefficient of the association between LD and accuracy for pseudowords was significantly lower than for irregular words ($z=5.86$, $p<0.001$) and regular words ($z=4.87$, $p<0.001$).

Comparison of correlations for AKS2

The difference between the coefficient for irregular word accuracy and AMPR irregular words ($r=0.411$) and regular word accuracy and AMPR regular words ($r=0.187$) was significant ($z=10.1$, $p<0.001$). This is consistent with the findings for EKS2 but not for F/KS and this is to be expected due to the consistent sound–letter associations that the regular words have.

The coefficient for LD irregular words and irregular word accuracy ($r=-0.956$) was significantly higher than that for LD regular words and regular word accuracy ($r=-0.930$), ($z=3.3$, $p<0.001$). This suggests that for irregular words there is more reliance on lexical processes than for regular words. The coefficient for LD irregular words and irregular word accuracy was also significantly higher than the coefficient for LD pseudowords and pseudoword accuracy (-0.927 ; $z=3.3$, $p<0.001$). This indicates that pseudowords in comparison to irregular words rely less on lexical processes.

Discussion

This study aimed to examine the interactions between key stages, spelling accuracy and measures of sublexical (phonological plausibility: PhP, Automated Measures for Phoneme Representation: AMPR) and lexical spelling processes (Letter Distance: LD). Spelling was examined in relation to the type of word, regular and irregular words and pseudowords, as they are differentially affected by lexicality. Beginning spellers seem to rely more on sublexical processes for spelling, and as children gain experience, lexical processes become more important. It was expected that spelling of irregular words would be primarily associated with lexically-related variables, while spelling of pseudowords would be associated with sublexically-related variables. For regular words, there should be reliance on both lexical and sublexical processing.

In terms of accuracy, similarly to past findings (de Bree and van den Boer, 2019), it was observed that children were the most accurate in spelling regular words, for all key stages. While pseudowords were spelled more accurately than irregular words up to the second half of KS2, the AKS2 children were more accurate with irregular words than pseudowords. This suggests that as they gain more experience with spelling and reading, children rely more on orthographic (lexical) processes. This is supported by the significant interaction between KS and LD, with improving scores for LD for all letter string types from early KS to AKS. The improvement for irregular words was such that they outperformed pseudowords at AKS2. The results of the present study are in line with the findings of former studies which indicated that spelling development is continuous rather than stage driven, reflecting gradual improvements in children's phonological and orthographic knowledge (de Bree and van den Boer, 2019; McMurray, 2020; Carvalhais et al., 2021).

AMPR also significantly improved from early to advanced KS for regular and irregular words. Pseudowords improved between F/KS1 and EKS2 but remained similar between EKS2 and AKS2, highlighting that as children put more effort into applying lexical processes at AKS2, phonologically implausible errors decrease more for irregular and regular words than pseudowords. This implies that for real words, phonological plausibility becomes strongly linked to lexical processing and automatic retrieval, which is also indicated by stronger associations between AMPR and accuracy as KS advances, when for pseudowords these associations become weaker. Another marker of sublexical processing, PhP, confirmed the regular word superiority over irregular words, as irregular words had consistently more phonologically appropriate errors than regular words. However, this result does not capture the subtle changes and improvements as shown by the AMPR and KS interaction effects as AMPR can be used for pseudowords. The results endorsed past research suggesting that non-binary measures are good metrics with which to monitor spelling development (Masterson and Apel, 2010; Werfel and Krimm, 2015).

To further confirm this, the strength of the relationships between spelling accuracy and phoneme- and letter-based measures differed as a function of KS and the type of letter string. Across KS groups, the strongest associations were found with LD, and less so for PhP and AMPR. The discrepancy in the strength of associations between LD and phoneme-based measures might be also explained by the inclusion of multisyllabic words in the spelling test, that need more effort at the whole-word level (Heggie and Wade-Woolley, 2017) and the different types of letter strings that have differential reliance on lexical and sublexical processing. This may explain the difference to Treiman et al. (2016, 2019), who found that the strongest predictors of spelling were accuracy followed by PhP and then LD.

For the early F/KS1 spellers the association between sublexical processes (i.e., AMPR) and accuracy for regular words and pseudowords was stronger than that for irregular words. This is expected as irregular words have less predictable letter-sound associations. At EKS2 the association between AMPR pseudoword and pseudoword accuracy was non-significant and smaller in comparison to the coefficients for regular and irregular words. This may suggest that at this stage knowledge of how the word is spelled is more important than sublexical phoneme-grapheme knowledge, which, by this KS level reaches a plateau.

Similarly, at AKS2, the association of irregular word accuracy with AMPR was stronger than the association of regular word accuracy and AMPR. While the influence of phonics in Key Stage 2 becomes less pronounced, for irregular words sublexical processes can be still important due to the inconsistent phoneme-grapheme correspondences. At KS2 there is more reliance on lexical processes (reflected in LD scores) for irregular words than for pseudowords, which is consistent with the findings for F/KS1. It is also noteworthy that AMPR and PhP show consistently significant and fairly similar in strength associations for regular and irregular words in all KS. This finding confirms that the two variables tap the same construct.

At F/KS1 LD was more strongly associated with accuracy for pseudowords than regular words. This is likely because regular words benefit from being able to draw on both lexical and sublexical processes so will be the closest to the target spellings. In the case of pseudowords, they need to be spelled using phoneme-grapheme correspondences (PGCs) and we counted any legal alternative PGC as

correct (unlike in the case of the words where it must be only the target one for that word in order to be correct). This lenient criterion for the pseudowords means they can be accurate phonologically while having many letters different from the target. In the case of the irregular words, the children will be trying to use PGCs but these will be disadvantageous and will lead to a big difference in letters compared to the target.

At EKS2 the difference in the associations between LD and accuracy for regular and irregular words did not reach significance indicating reliance on lexical processes for both word types, demonstrated also by the significantly stronger associations in comparison to the LD-accuracy association for pseudowords. However, at AKS2 LD-accuracy irregular word associations were stronger than those for both regular words and pseudowords, showing a shift to relying more on lexical processes when spelling irregular words.

AMPR and LD are two non-binary measures that relate to spelling accuracy with all letter string types (except AMPR and pseudoword accuracy at EKS2). The strength of associations varied depending on the group and letter string category. AMPR was consistently less strongly associated with accuracy in comparison to LD. For pseudoword and regular word accuracy this is not an odd finding as for children with more competence in spelling, sublexical processing will be less critical for their spelling due to the straightforward PGCs. For irregular word spelling, the reliance on lexical processes is strong at later stages, as reflected in the LD scores outperforming LD for pseudowords at AKS2. For FKS1 children there is also less influence of phoneme-based measures (AMPR) than LD for irregular word spelling even if the influence of phonics teaching is strong in this age group. The influence of phonics, although strong, is not the optimal strategy to spell irregular words. Regular words, on the other hand, at all KS, are the most accurate word type as they benefit from input from both lexical and sublexical processes, indicated also by the strong associations between LD scores and accuracy. The findings suggest that there is a gradual unification of spelling processes (orthography, phonology and semantics) from KS1 to AKS2 as suggested by the lexical quality hypothesis (Perfetti and Hart, 2002) and the 'linguistic trilogy' that suggests spelling effectively requires all three processes (Wolter and Dilworth, 2014).

Finally, for pseudoword spelling and for F/KS1, there was influence of LD and for EKS2, LD was more strongly associated with pseudoword accuracy than with regular and irregular word accuracy. This may be partly explained to the way pseudowords were devised for the spelling test, which implicates lexical processing. In an attempt to spell pseudowords that somewhat resemble real words, children employ lexical processes to a greater extent than for regular and irregular words, as the latter two are more easily recognized. Lexical processing becomes a strategy that can be transferred to other types of letter strings. Another potential interpretation could be that pseudowords need to be spelled using PGCs and we counted any alternative PGC as correct. This lenient criterion for the pseudowords means they can be accurate phonologically while having many letters different from the target.

This important finding agrees with results from the longitudinal study of Carvalhais et al. (2021). Those researchers found for the older cohort of Portuguese learners that orthographic and stress errors were more prevalent than phonological errors, indicating that the influence

of lexical semantic processes are stronger as phonological skills are by now very well mastered. That means from a school perspective that spelling instruction should foster more orthographic strategies in these later stages of education.

Use of non-binary ways to score spelling errors can provide the teachers with a powerful tool to unpick the type of errors children make and if they need more support in phonological and lexical semantic (orthographic elements) processes. The spelling test with irregular and regular words and pseudowords can further uphold the teacher's work in finding the linguistic gaps that children might have and tailor an appropriate bespoke intervention. It is also evident that in addition to phonological processes, orthographic knowledge plays a significant role in learning to spell. Therefore, an educational system that emphasizes solely alphabetic strategies may put children at a disadvantage and it might also be of disservice.

As with any study, this one is not free of limitations. Several variables were generated by computer programs; however, some errors can still be difficult to categorize. Also, the data are cross-sectional while a longitudinal design could be more informative. However, these limitations are minimized by strengths, such as the reliable measures that were used to score children's spellings. PhP has been employed in many earlier studies examining the type of errors children make. The large sample size ($N=641$) and items in the spelling test ($N=106$) produced a large number of spellings—the total number of words including correct answers and no responses exceeded 65,000. To further our understanding of spelling development, researchers could apply a similar method in a cross-cultural study where spellings of native speakers of other orthographies are examined. This will allow us to ascertain whether these findings are universal or not.

Many years of schooling need to take place in order for spellings to become established entries in the mental lexicon. This study is one of the first to demonstrate when the lexical and sublexical processes based on the DR model (Martin and Barry, 2012) start to function in a more integrated and interdependent way. It will be informative for further studies to be conducted with more transparent and opaque orthographies as these can indicate if there is a universal age at which this occurs.

Conclusion

The scoring measures included previously showed primary school children's reliance on lexical and sublexical spelling processes (Caravolas et al., 2001; Treiman et al., 2016, 2019). The measures assign credit for partially correct spellings and can allow researchers and educators to capture changes in the development of spelling ability, and to analyse spelling performance (Treiman and Kessler, 2004; Kessler, 2009; Masterson and Apel, 2010). The fine-tuned measures employed in calculating the spelling errors, in addition to use of an assessment that differentiates between different word types, means we were able to capture changes as the children become more advanced in spelling. It was observed that as spelling skill increased, and the influence of systematic synthetic phonics presumably became less strong, children appeared to become more reliant on lexical processes. We believe that our study will inspire more research in spelling which will help unlock the mystery of learning to spell conventionally, and highlight that every letter counts for understanding spelling processes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Bath Spa University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This article was supported by Research England HEQR seed funding from Bath Spa University and British Academy (BA)/

References

- Barry, C. (1994). "Spelling routes (or roots or rutes)" in *Handbook of Spelling: Theory, Process and Intervention*. (London: Wiley), 27–49.
- Barry, C., and Seymour, P. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *Quart. J. Exp. Psychol. Section A* 40, 5–40. doi: 10.1080/14640748808402280
- Bruck, M., and Waters, G. (1988). An analysis of the spelling errors of children who differ in their reading and spelling skills. *Appl. Psycholinguist.* 9, 77–92. doi: 10.1017/S0142716400000461
- Caravolas, M., Hulme, C., and Snowling, M. (2001). The foundations of spelling ability: evidence from a 3-year longitudinal study. *J. Mem. Lang.* 45, 751–774. doi: 10.1006/jmla.2000.2785
- Carvalho, L., Limpo, T., and Pereira, L. Á. (2021). The contribution of word-, sentence-, and discourse-level abilities on writing performance: a 3-year longitudinal study. *Front. Psychol.* 12:668139. doi: 10.3389/fpsyg.2021.668139
- Carvalho, L., Limpo, T., Richardson, U., and Castro, S. L. (2020). Effects of the Portuguese GraphoGame on Reading, spelling, and phonological awareness in second graders struggling to read. *J. Writing Res.* 12, 9–34. doi: 10.17239/jowr-2020.12.01.02
- Clemens, N., Oslund, E., Simmons, L., and Simmons, D. (2014). Assessing spelling in kindergarten: further comparison of scoring metrics and their relation to reading skills. *J. Sch. Psychol.* 52, 49–61. doi: 10.1016/j.jsp.2013.12.005
- de Bree, E., and van den Boer, M. (2019). Knowing what we Don't know: cognitive correlates of early spelling of different target types. *Read. Writ.* 32, 2125–2148. doi: 10.1007/s11145-019-09936-9
- Eid, M., Gollwitzer, M., and Schmitt, M. (2011). Statistik und Forschungsmethoden Lehrbuch. Weinheim: Beltz.
- Frisby, C. (2016). An empirical comparison of the words spelled correctly and correct letter sequence spelling scoring methods in third- and fourth-grade classrooms. *J. Appl. Sch. Psychol.* 32, 101–121. doi: 10.1080/15377903.2016.1151847
- Georgiou, G. K., Torppa, M., Landerl, K., Desrochers, A., Manolitsis, G., de Jong, P. F., et al. (2020). Reading and spelling development across languages varying in orthographic consistency: do their paths cross? *Child Dev.* 91, e266–e279. doi: 10.1111/cdev.13218
- Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., and Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Read. Writ.* 25, 321–346. doi: 10.1007/s11145-010-9271-x
- Heggie, L., and Wade-Woolley, L. (2017). Reading longer words: insights into multisyllabic word reading. *Perspectives ASHA Special Interest Groups* 2, 86–94. doi: 10.1044/persp2.SIG1.86
- Leverhulme Small Research Grant [SG170586] awarded to the GN and a British Psychological Society Grant awarded to the AN.
- Joye, N., Dockrell, J. E., and Marshall, C. R. (2020). The spelling errors of French and English children with developmental language disorder at the end of primary school. *Front. Psychol.* 11:1789. doi: 10.3389/fpsyg.2020.01789
- Kessler, B. (2009). Statistical learning of conditional orthographic correspondences. *Writing Syst. Res.* 1, 19–34. doi: 10.1093/wsr/wsp004
- Kessler, B., Pollo, T., Treiman, R., and Cardoso-Martins, C. (2012). Frequency analyses of prephonological spellings as predictors of success in conventional spelling. *J. Learn. Disabil.* 46, 252–259. doi: 10.1177/0022219412449440
- Kroese, J. M., Hynd, G. W., Knight, D. F., Hiemenz, J. R., and Hall, J. (2000). Clinical appraisal of spelling ability and its relationship to phonemic awareness (blending, segmenting, elision, and reversal), phonological memory, and reading in reading disabled, ADHD, and normal children. *Reading and Writing*, 13, 105–131.
- Martin, D., and Barry, C. (2012). Writing nonsense: the interaction between lexical and sublexical knowledge in the priming of nonword spelling. *Psychon. Bull. Rev.* 19, 691–698. doi: 10.3758/s13423-012-0261-7
- Masterson, J., and Apel, K. (2010). The spelling sensitivity score: noting developmental changes in spelling knowledge. *Assess. Eff. Interv.* 36, 35–45. doi: 10.1177/1534508410380039
- McMurray, S. (2020). Learning to spell for children 5–8 years of age: the importance of an integrated approach to ensure the development of phonetic, orthographic and morphemic knowledge at compatible levels. *Dyslexia* 26, 442–458. doi: 10.1002/dys.1663
- Nag, S., Vagh, S., Dulay, K., and Snowling, M. (2018). Home language, school language and children's literacy attainments: a systematic review of evidence from low- and middle-income countries. *Rev. Educ.* doi: 10.1002/rev3.3130
- Niolaki, G. Z., and Masterson, J. (2012). Transfer effects in spelling from transparent Greek to opaque English in seven-to-ten-year-old children. *Biling. Lang. Cogn.* 15, 757–770. doi: 10.1017/S1366728911000721
- Niolaki, G. Z., Terzopoulos, A. R., and Masterson, J. (2014). Varieties of developmental dyslexia in Greek children. *Writing Syst. Res.* 6, 230–256. doi: 10.1080/17586801.2014.893862
- Niolaki, G., Vousden, J., Terzopoulos, A., Shepherd, D. L., Debney, L., and Masterson, J. (2022). Spelling predictors: investigating the role of phonological ability and rapid naming in a large cross-sectional British study. *Learn. Instr.* 80:101635. doi: 10.1016/j.learninstruc.2022.101635
- Niolaki, G. Z., Vousden, J., Terzopoulos, A. R., Taylor, L. M., Sephton, S., and Masterson, J. (2020). Predictors of single word spelling in English speaking children: a cross sectional study. *J. Res. Read.* 43, 577–596. doi: 10.1111/1467-9817.12330

- Perfetti, C. A., and Hart, L. (2002). The lexical quality hypothesis. *Precursors Functional Literacy* 11, 67–86. doi: 10.1075/swll.11.14per
- Ritchey, K., Coker, D., and McCraw, S. (2009). A comparison of metrics for scoring beginning spelling. *Assess. Eff. Interv.* 35, 78–88. doi: 10.1177/1534508409336087
- Rønneberg, V., and Torrance, M. (2019). Cognitive predictors of shallow-orthography spelling speed and accuracy in 6th grade children. *Read. Writ.* 32, 197–216. doi: 10.1007/s11145-017-9751-3
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition* 55, 151–218.
- Spencer, K. A. (2009). Feedforward, backward, and neutral transparency measures for British English. *Behav. Res. Methods* 41, 220–227. doi: 10.3758/BRM.41.1.220
- Treiman, R. (2018). Statistical learning and spelling. *Lang. Speech Hear. Serv. Sch.* 49, 644–652. doi: 10.1044/2018_LSHSS-STLT1-17-0122
- Treiman, R., and Bourassa, D. C. (2000). The development of spelling skill. *Top. Lang. Disord.* 20, 1–18. doi: 10.1097/00011363-200020030-00004
- Treiman, R., and Kessler, B. (2004). The case of case: children's knowledge and use of upper- and lowercase letters. *Appl. Psycholinguist.* 25, 413–428. doi: 10.1017/S0142716404001195
- Treiman, R., Kessler, B., and Caravolas, M. (2019). What methods of scoring young children's spelling best predict later spelling performance? *J. Res. Read.* 42, 80–96. doi: 10.1111/1467-9817.12241
- Treiman, R., Kessler, B., Pollo, T., Byrne, B., and Olson, R. (2016). Measures of kindergarten spelling and their relations to later spelling performance. *Sci. Stud. Read.* 20, 349–362. doi: 10.1080/10888438.2016.1186168
- Vousden, J. I., Ellefson, M. R., Solity, J. E., and Chater, N. (2011). Simplifying Reading: applying the simplicity principle to Reading. *Cogn. Sci.* 35, 34–78. doi: 10.1111/j.1551-6709.2010.01134.x
- Werfel, K., and Krimm, H. (2015). Utility of the spelling sensitivity score to analyse spellings of children with specific language impairment. *Aust. J. Learn. Difficult.* 20, 39–53. doi: 10.1080/19404158.2015.1047871
- Wolter, J. A., and Dilworth, V. (2014). The effects of a multilinguistic morphological awareness approach for improving language and literacy. *J. Learn. Disabil.* 47, 76–85. doi: 10.1177/0022219413509972
- Zoccolotti, P., De Luca, M., Marinelli, C. V., and Spinelli, D. (2020). Predicting individual differences in reading, spelling and maths in a sample of typically developing children: a study in the perspective of comorbidity. *PLoS One* 15:e0231937. doi: 10.1371/journal.pone.0231937



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Carolina Cordeiro,
University of Porto, Portugal
Fareeha Javed,
Higher Education Department,
Government of Punjab, Pakistan
Macarena Becerra,
University of Malaga, Spain
Sara Kashefian-Naeeni,
Shiraz University of Medical Sciences, Iran

*CORRESPONDENCE

Lawrence Jun Zhang
✉ lj.zhang@auckland.ac.nz
Ye Zhu
✉ zhuye@shisu.edu.cn

RECEIVED 07 March 2023

ACCEPTED 18 May 2023

PUBLISHED 07 June 2023

CITATION

Zhang J, Zhang LJ and Zhu Y (2023)
Development and validation of a genre-based
second language (L2) writing self-efficacy
scale.
Front. Psychol. 14:1181196.
doi: 10.3389/fpsyg.2023.1181196

COPYRIGHT

© 2023 Zhang, Zhang and Zhu. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Development and validation of a genre-based second language (L2) writing self-efficacy scale

Jianhua Zhang^{1,2}, Lawrence Jun Zhang^{1,2*} and Ye Zhu^{3*}

¹School of Foreign Languages, Sichuan University of Arts and Science, Dazhou, China, ²Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand, ³State Language Commission of China, Institute of Linguistics, Shanghai International Studies University and China Center for Language Planning and Policy Studies, Shanghai, China

Writing self-efficacy serves as one of the essential motivational factors in L1 and L2 writing, which has been measured by a series of scales in L1 and L2 contexts. However, the issue of task specificity was not resolved appropriately. This study aims to tackle this issue by entailing the genre characteristics of L2 writing tasks through developing a genre-based L2 writing self-efficacy scale with pertinent items. The new scale was designed with reference to the available research into writing self-efficacy. Its factorial structure was examined by structural equation modeling. Convergent validity and discriminant validity of the scale were examined by taking into consideration the average variance extracted and composite reliability for each individual factor involved in the scale, whereas the predictive validity of the scale was computed through regression analysis. Results show that the genre-based L2 writing self-efficacy scale demonstrated sound psychometric qualities. Theoretical and pedagogical implications of these research findings are discussed.

KEYWORDS

development of a genre-based L2 writing self-efficacy scale, genre characteristics, second language (L2), writing self-efficacy, psychometric quality, foreign language

1. Introduction

Motivational factors play an essential role in the models of writing (e.g., [Boscolo and Hidi, 2007](#); [Hayes, 2012](#); [Graham, 2018](#); [Zhang, 2022](#)). Writing self-efficacy has been widely acknowledged as one of the motivational factors in the first language (L1) and second or foreign language (henceforth referred to as L2) writing ([Chen and Zhang, 2019](#); [Chen et al., 2021](#)) and is conceptualized as a multidimensional construct. For instance, [Pajares and Valiante \(1999\)](#) proposed that writing self-efficacy entailed ideation self-efficacy, writing convention self-efficacy, and self-regulation self-efficacy, whereas [Teng et al. \(2018\)](#) conceptualization included linguistic self-efficacy, self-regulation self-efficacy, and performance self-efficacy. Scholars have developed and validated scales or questionnaires for gaging writing self-efficacy in the L1 and L2 contexts due to the significance of self-efficacy in the learning process (e.g., [Shell et al., 1989](#); [Pajares and Valiante, 1999](#); [Bruning et al., 2013](#); [Teng et al., 2018](#); [Sun and Wang, 2020](#)). [Bandura \(2006\)](#) pointed out that items in self-efficacy scales or questionnaires are supposed to represent specific task demands. Notwithstanding, the available scales for measuring L1 and L2 writing self-efficacy have not dealt with the issue of task specificity appropriately. The employment of these scales might not lead us to gain a complete understanding of learners' writing self-efficacy. Informed by the multidimensional perspective of writing self-efficacy, the current study is intended to tackle this issue by incorporating one of the vital writing task features (i.e., genre

characteristics) in the development of items in the writing self-efficacy scale.

2. Literature review

2.1. Conceptualization and dimensions of writing self-efficacy

2.1.1. Self-efficacy

Self-efficacy has been widely conceived as a vital construct in psychology and sociology, where it acted as a basic mechanism in psychosocial functioning. Scholars interpreted self-efficacy from different perspectives (e.g., Kirsch, 1985; Bandura, 1986). From the perspective of reinforcement theory, Kirsch (1985) viewed it as “expectancy for success at a task on which success is perceived to be dependent on ability” (p. 5). Bandura, however, recast it within the framework of social cognitive theory. He defined it as “people’s judgment of their capabilities to organize and execute courses of action required to attain designated types of performance” (Bandura, 1986, p. 391). Bandura commented on Kirsch’s conceptualization of self-efficacy and criticized Kirsch’s misinterpretation of self-efficacy by stating, “Kirsch further misrepresents self-efficacy theory when he alleges that the theory postulates low perceived self-efficacy as the cause of fear, irrespective of the domain of activity” (1986, p. 370). Thus, Bandura’s conceptualization of self-efficacy has gained increasing popularity and has been widely utilized in various studies (e.g., Schunk, 1991; Zimmerman, 2000; Matthews, 2010; Zumbunn et al., 2020).

Self-efficacy, as Zimmerman (2000) pinpointed, served as an essential motive for learning, which can effectively predict students’ motivation and learning achievement. It acts on students by influencing their choices, effort, persistence and perseverance, thought patterns, and emotional reactions (Pajares, 2003). Students could develop their self-efficacy beliefs on the basis of four kinds of information: mastery experience, vicarious experience, social persuasion, and physiological and emotional states (Bandura, 1997). Mastery experience can be viewed as an interpretation of one’s previous performance, which served as a base on which to form their beliefs about their performance capabilities in subsequent tasks. Vicarious experience is interpreted as observing others’ performance or modeling. For instance, instructors can model planning skills to learners in the educational context. Social persuasion means evaluations of one’s capability that others give, which can be either positive or negative. Positive evaluations can enhance students’ self-efficacy beliefs, whereas negative ones weaken them. Physiological and emotional states such as pause, anxiety, stress, and arousal indirectly inform one’s self-efficacy beliefs.

Another strand of exploring self-efficacy was rooted in self-regulated learning theory. According to Bandura (1986), self-regulatory factors act as a central player in human functioning. Self-regulation was conceived as “a metacognitive process that requires students to explore their own thought processes so as to evaluate the results of their actions and plan alternative pathways to success” (Usher and Pajares, 2008, p. 443). Self-regulatory strategies could facilitate students to succeed in their development. It was found that students’ utilization of self-regulatory strategies was mainly determined by their beliefs about their capabilities to do so (Zimmerman and Cleary, 2006).

2.1.2. Writing self-efficacy

Research has established that self-efficacy influences human motivation and action (e.g., Schunk, 1991; Zimmerman, 2000; Pajares, 2003; Schunk and Zimmerman, 2007; Caprara et al., 2008; Usher and Pajares, 2008). It is widely recognized as an essential player in acquiring writing competence and skills (e.g., McCarthy et al., 1985; Pajares and Johnson, 1994; Pajares and Valiante, 1997, 1999; Pajares et al., 1999). Writing self-efficacy refers to students’ perception of the writing capabilities that they possess to complete writing tasks (Pajares, 2003). It involves students’ judgement of “various composition, grammar, usage, and mechanical skills” (Pajares and Johnson, 1996, p.166). After summarizing numerous previous studies, Bruning et al. (2013) proposed a model of writing self-efficacy that included three focal dimensions: ideation, or idea generation; writing conventions, or translating ideas into words and sentences; and self-regulation, or the management, monitoring, and evaluation of writing processes. Ideation self-efficacy paid more attention to students’ confidence in their abilities to generate ideas. Idea generation served an essential role in the models of the writing process (e.g., Hayes and Flower, 1980; Hayes, 1996, 2012). It influenced other writing processes because of its cyclic nature. Writing convention self-efficacy is concerned with students’ confidence in following “a set of generally accepted standards for expressing ideas in writing in a given language” (Bruning et al., 2013, p. 28). For instance, the standards in English could cover spelling, punctuation, word, phrase, sentence, paragraph, discourse and their combinations for different populations. Self-regulation self-efficacy focused on students’ confidence in implementing self-regulatory skills successfully through the writing processes. It entails “a writer’s self-management and affective control but also involves judgments about cognitive and linguistic features as writing is being produced” (Bruning et al., 2013, p. 29). This model incorporated the thoughts accumulated about the writing process, writing and writing development. Thus, it was conceived as “consonant with writing process models emphasizing working memory’s centrality.....as well as with other portrayals of writing and writing development.....” (Bruning et al., 2013, p. 28). Additionally, Limpo et al. (2020) extended the conception of writing self-efficacy by adding two dimensions of handwriting and text genre, particularly for beginning writers in the L1 context.

Obviously, writing self-efficacy is a multidimensional construct. Given that there are huge differences between L1 and L2 writers, Teng et al. (2018) reconceptualized writing self-efficacy from three aspects of writing: linguistic, self-regulation, and performance. Their conceptualization was grounded in social cognitive theory and self-regulated learning theory and also informed by the difficulties that L2 writers encounter and the characteristics of L2 writing (Teng et al., 2018). Linguistic self-efficacy is concerned with students’ perception of their capability to retrieve words, translate ideas into sentences, and meet discourse requirements. Performance self-efficacy focuses on learners’ confidence in their abilities to complete a writing task in the instructional context. Self-regulation or self-regulatory efficacy was defined as “students’ perceived capability to execute metacognitive control in the learning-to-write process” (Teng et al., 2018, p. 23).

Although scholars held various conceptualizations of writing self-efficacy, they might have neglected some of the characteristics of writing tasks in their conceptualizations, thus not addressing the issue of task-specificity proposed by Bandura (2006). Furthermore, research has found that task characteristics (i.e., genre) could elicit learners’

distinctive writing performance (Yang et al., 2015, 2023; Yoon, 2021; Zhang and Cheng, 2021; Zhang and Zhang, 2021; Li and Zhang, 2022; Yang and Zhang, 2023). Therefore, it is suggested that the conceptualization of writing self-efficacy incorporate the aforementioned advances in L2 writing to demonstrate students' perception of writing competence in different genres.

2.2. Measuring writing self-efficacy

Writing self-efficacy is defined as "students' judgments of their competence in writing, specifically their judgments that they possess various composition, grammar, usage, and mechanical skills" (Pajares and Johnson, 1996, p.166). Three ways have been proposed to measure writing self-efficacy (Pajares, 2003). The first approach concentrates on the writer's confidence in the specific skills they have mastered, such as grammar, usage, punctuation and storytelling. The second involves evaluating the confidence in accomplishing concrete writing tasks, for example, writing a term paper or a letter to a teacher or professor. The third approach focuses on assessing writers' confidence in reaching specific performance criteria set by the course. The aforementioned ways of measuring writing self-efficacy reveal that it is a complex construct with different dimensions. Therefore, a multidimensional approach to writing self-efficacy would be necessary and vital for providing fine-grained information about writers' self-perceived competence.

It is well-acknowledged that self-efficacy can be effectively measured by questionnaires/scales. Clear guidelines on the operationalization and measurement of self-efficacy beliefs were given by Bandura (2006), who emphasized: "self-efficacy assessment tailored to domains of functioning and task demands identify patterns of strengths and limitations of perceived capability" (p. 319). Therefore, developing self-efficacy scales entails fitting scale items with particular domains and specific demands of different tasks. In addition, Zimmerman (2000) stressed that: (1) self-efficacy measures entail performance capabilities rather than personal qualities; (2) self-efficacy measures should be administered to students prior to engaging in certain tasks. His emphasis might shed insights into the development of writing self-efficacy scales: the writing self-efficacy scale should incorporate items assessing learners' confidence in performing specific writing tasks. Furthermore, it might also warn us of the timing of evaluating writing self-efficacy.

Research has revealed that writing self-efficacy exerted significant predictive effects on writing performance not only in L1 but also in L2 contexts (e.g., Pajares and Johnson, 1996; Prat-Sala and Redford, 2012; Teng et al., 2018; Sun and Wang, 2020). Several scales for writing self-efficacy have been designed and/or validated in L1 and L2 contexts. Pajares (2003) summarized three ways of measuring writing self-efficacy. The first way focused on evaluating students' confidence in specific writing skills. They were operationalized as successful performance in using grammar, showing mastery of usage, writing a composition, and demonstrating mechanical writing skills (e.g., Shell et al., 1989; Pajares and Johnson, 1996), as specific story-writing skills (Graham and Harris, 1989), and as writing skills selected by teachers that were suitable for relevant tasks (e.g., Pajares and Valiante, 1997, 1999, 2001; Pajares et al., 1999). The second way entailed evaluating students' confidence to complete specific writing tasks, for instance, a term paper and a letter (e.g., Shell et al., 1989; Pajares and Johnson,

1994). The third way involved the integration of the first and the second ways, thus assessing students' confidence in demonstrating both specific writing skills and completing writing tasks.

Initially, Shell et al. (1989) created a scale for examining L1 students' writing self-efficacy at a tertiary school. The scale was composed of two subscales: the task subscale measuring students' confidence in performing writing tasks, and the skill subscale measuring their confidence in utilizing specific writing skills. They reported that scores of the skill subscale showed a sound predictive effect on writing achievement, but those of the task subscale did not. In other words, students' writing skill self-efficacy can predict their writing performance, but their writing task self-efficacy can not. The scale was validated with young subjects of three different grades from primary and secondary school by Shell et al. (1995). They also found that writing skills self-efficacy rather than writing task self-efficacy exerted a significant predictive effect on writing achievement. The same result was reported in Pajares and Johnson (1994).

Later, the Writing Self-Efficacy Scale, developed by Pajares and Valiante (1999), has been widely utilized in studies relevant to writing self-efficacy because of its acceptable stability and internal consistency. Items in this scale were designed to measure middle school students' perception of their confidence in how well they can utilize grammar, usage, composition, and mechanical writing skills in writing tasks. A 0–100 response format rather than a traditional Likert format was employed because the former showed better psychometric quality than the latter (Pajares et al., 2001). They found that writing skills self-efficacy served as a significant predictor of writing competence compared with variables such as writing self-concept, previous apprehension, and perceived value of writing. Subsequently, the underlying structure of this scale was examined through structural equation modeling by Pajares (2007), which uncovered that writing skills self-efficacy was composed of two factors: basic skills and composition skills.

While writing self-efficacy drew more and more attention from researchers, criticisms emerged regarding its evaluation. Bruning et al. (2013) criticized writing self-efficacy scales available at that time for the broad coverage of related writing skills and the lack of theoretical underpinnings. After reviewing a series of studies, Bruning et al. (2013) pointed out that "most writing self-efficacy measures, however, have broadly sampled writing-related skills and tasks, making them less than ideal for yielding information about writers' self-efficacy for specific dimensions of writing" (p. 27). Besides, these instruments were found to be difficult to be related "directly to models of writing or to potentially writing-relevant psychological and language-related processes" (Bruning et al., 2013, p. 26).

As mentioned in the above section, Bruning et al. (2013) put forward a model of writing self-efficacy where self-efficacy beliefs were tied to writing models. The constructs in this model were examined and validated with middle and high school students through the Self-Efficacy for Writing Scale. This scale was composed of 16 items: ideation (5 items), conventions (5 items), and self-regulation (6 items). The proposed three-factor model of writing self-efficacy was confirmed, and the results suggested its generalizability.

All the aforementioned scales were designed to measure L1 students' writing self-efficacy and adapted and modified to fit the specific research questions. To our knowledge, the Second Language Writer Self-Efficacy Scale was one of a few instruments available to evaluate students' writing self-efficacy in a second/foreign language

context. It was designed by [Teng et al. \(2018\)](#) on the basis of their conceptualization of writing self-efficacy. This scale consisted of 3 subscales with 20 items: linguistic self-efficacy (7 items), performance self-efficacy (7 items), and self-regulatory self-efficacy (6 items). Structural equation modeling was employed to validate the scale with Chinese-speaking EFL learners at the tertiary level. CFA revealed that the scale showed satisfactory psychometric qualities. Model comparisons demonstrated that the three-factor correlated model fits with the data collected better than the one-factor and three-factor uncorrelated models. However, items examining their conceptualizations were criticized for not having taken into consideration task specificity ([Sun and Wang, 2020](#)). Therefore, [Teng et al. \(2018\)](#) Second Language Writer Self-Efficacy Scale needs to be extended to incorporate features of writing tasks.

Recently, [Sun and Wang \(2020\)](#) developed a new scale named the Questionnaire of English Writing Self-Efficacy to assess writing self-efficacy in ESL or EFL contexts. Items in the questionnaire were drawn from the Self-Efficacy for Writing Scale and the Questionnaire of English Self-Efficacy, which were developed by [Wang and Bai \(2017\)](#) to evaluate EFL learners' general English self-efficacy. The new scale examined five writing-related dimensions: ideation, organization, grammar and spelling, use of English writing, and self-regulation, each of which acted as a subscale. It was designed in a 7-point Likert response format. Cronbach's alpha showed a good internal consistency of each subscale. CFA revealed that the five-factor model fitted the data. Notwithstanding, the Questionnaire of English Writing Self-Efficacy needs to be finetuned. This is because the issue of task specificity is left unsolved in that some items in the questionnaire are not directly relevant to the writing task included in their study, and features of a writing task, such as genre, are insufficiently covered. Moreover, the authors might not have given enough attention to the theories underpinning the design of the questionnaire.

As suggested by [Bandura \(2006\)](#), self-efficacy measurement should be tailored to cover task demands. In other words, the instruments of self-efficacy should be designed as task-specific. As known to us, writing tasks are defined in specific genres. Genres are characterized by different patterns of language use and rhetoric features ([Wingate and Tribble, 2012](#)). Genre features could impose constraints on written discourses, for example, employing certain linguistic patterns to achieve a persuasive purpose, thus constituting the high-level demands for writing essays, both in L1 and L2 contexts. Furthermore, it was found that learners demonstrated distinctive syntactic structures, phrases, and words across argumentative and narrative tasks ([Ong and Zhang, 2010](#); [Yang et al., 2015](#); [Yoon, 2021](#)). Thus, genre features of these tasks should be included as one of the essential components of writing self-efficacy instruments.

However, to our knowledge, few of the available instruments investigating writing self-efficacy took into consideration task characteristics, such as genre features. Employing these instruments to assess writing self-efficacy might result in, possibly, a partial understanding of learners' self-belief in completing certain writing tasks; consequently, inappropriate teaching intervention can be rectified to achieve the expected teaching effectiveness and efficiency. Consequently, the validity and reliability of these instruments would be left questionable. To properly address this issue, therefore, it is imperative and reasonable to develop a new writing self-efficacy scale in relation to L2 writing contexts by incorporating task features of L2 writing, thus facilitating teachers, students and

researchers to gain a more complete understanding of students' writing self-efficacy and simultaneously providing a tentative solution to the issue raised by [Bandura \(2006\)](#) about the lack of task specificity in self-efficacy research.

3. Present study

This study aimed to develop a new scale for evaluating writing self-efficacy in the L2 context by incorporating task-specificity in relation to genre features of writing tasks. Therefore, we tried to answer three research questions:

1. What were the factors of the newly developed writing self-efficacy scale?
2. How did the factorial structure of the newly developed writing self-efficacy scale fit with the target subjects?
3. How did multi-dimensional writing self-efficacy predict L2 writing performance?

4. Methods

4.1. Participants

A total of 664 EFL students as convenient samples from a population of 50,000 at two medium universities in Western China were recruited to participate voluntarily by employing Slovin's formula, namely, using the formula as shown here: $n = N / (1 + Ne^2)$; and the participants were divided into two groups: Sample A comprised 332 students, and Sample B consisting of 332 students. When participating in this study, the participants had studied English for at least 6 years since the majority of them started learning English while they were in junior high school, and their mother tongue is Mandarin Chinese. Their English writing instruction was embedded in the integrated English course. The participants ranged from the first year to the third year (55.42% freshmen, 15.663% sophomores, and 28.915% juniors), of which 59.04 per cent ($n = 392$) were females, and 40.96 per cent ($n = 272$) were males between the ages of 18 and 21 (Mean = 19.7). They were registered in the following majors: electronic engineering ($n = 100$, 15.06%), computer science ($n = 114$, 17.17%), education ($n = 114$, 17.17%), business ($n = 176$, 26.51%), administration ($n = 96$, 14.46%), and tourism ($n = 64$, 9.64%). The participants in both Sample A and Sample B were equivalent in grade, gender, and major/specialization distributions.

4.2. Measure development

Following the guide given by [Bandura \(2006\)](#) for constructing self-efficacy scales, we took the particular domain of functioning and task demands as the priority in this research. As implied by [Pajares \(2003\)](#), whether a self-efficacy scale is appropriate and adequate depends to a great extent on "the domain under investigation, its different features, the types of the capabilities it requires, and the range of the situations these capabilities might be applied" (p. 144). To establish the content validity of the scale appropriate for EFL writers

at the tertiary level, we consulted and examined in a nuanced fashion some established instruments such as the Writing Skills Self-Efficacy Scale (WSES, Pajares and Valiante, 1999), the Writing Self-Regulatory Efficacy Scale (WSRES, Zimmerman and Bandura, 1994), the Self-Efficacy for Writing Scale (SEWS, Bruning et al., 2013) and the Second Language Writer Self-Efficacy Scale (SLWSES, Teng et al., 2018).

Initially, a total of 25 items relevant to EFL writing self-efficacy were produced. The content and face validity of these items were examined by three scholars who are well-published in local and international journals and based in the country where the study was conducted. Specifically, they scrutinized the theoretical rationale adopted here, evaluated whether the generated initial items matched the construct being targeted and measured, and checked whether the diction of the scale was clear and readable. The examination and evaluation of the items of this scale were performed in two rounds. The first round ended with the elimination of unnecessary items, the rewording of double-barreled items and the addition of items to remedy the obvious omissions.

Although a scale with a 0–100 response format has been proven to show stronger psychometric qualities in comparison with a traditional five-/seven-point Likert one in gaging self-efficacy beliefs (Pajares et al., 2001), the 0–100 response format was recently found to cause potential confusion for EFL learners when being employed to measure self-efficacy beliefs (Chen and Zhang, 2019; Chen et al., 2022). As a result, a compromise must be made, and then a 7-point Likert response format was adopted in this research. The finalized instrument containing 24 items was arranged in a logical fashion on the basis of the clusters of subcategories. A 7-point Likert scale with a gradation rating from 1 (*not at all true of me*) to 7 (*very true of me*) was adopted to explore the trait and state features of writing self-efficacy.

All the items of the current scale were developed in English and translated into Chinese when presented to the participants, guaranteeing that they can fully understand the items and avoid potential misunderstandings. The accuracy and equivalence of the translation were verified and backed up by means of translating and back-translating. Finally, the GL2WSS were then subjected to statistical tests to examine its construct validity.

4.3. English writing tests

Two writing tasks with a given topic were employed to examine the participants' writing performance: a narrative writing task and an argumentative one. These tasks completely matched the genres covered in the newly developed GL2WSS and thus enabled the GL2WSS to be really task-specific. In addition, narrative writing was the first genre they learned to write, whereas argumentative writing was the one they practiced most because it was targeted in classroom assessments and national and international English tests.

Participants in this study were required to finish two English compositions of at least 150 words according to the given prompts (see Appendix A for details) within 40 min, respectively, in an online writing platform named Pigai in two rounds. The topic for the argumentative writing was chosen from the old item pool of the College English Test, Band 4 (CET 4), which showed high validity and reliability. In contrast, the topic for the narrative writing was designed as culturally inoffensive and closely related to participants' daily life.

Jacobs et al.'s (1981) ESL Composition Profile, one of the established analytic scoring rubrics, was employed as the scoring criteria to assess participants' writing performances. Jacobs et al.'s scoring rubric has been widely used in L2 writing studies to evaluate the writing proficiency levels of L2 students around the world by virtue of its relatively easy operationalization (e.g., Ong and Zhang, 2010; Huang and Zhang, 2020; Rahimi and Zhang, 2021). The rubric gages five aspects of written essays: content (i.e., including knowledge of the subject, development of a thesis, and relevance to topics), organization (i.e., including idea support, organization, and sequencing), language use (i.e., including constructions and grammatical errors), vocabulary (i.e., including word range, word choice and usage, form mastery, and word appropriacy) and mechanics (i.e., including mastery of conventions and error rate of spelling, punctuation, capitalization and paragraphing), which are given different weights in the scoring scheme: content (30%), organization (20%), language use (25%), vocabulary (20%) and mechanics (5%). These aforementioned aspects can be rated at four rating levels: excellent to very good, good to average, fair to poor, and very poor.

4.4. Procedures

After signing the consent form, the participants were asked to complete the newly developed scale through an online survey platform named *Wenjuanxing* (literally translated as Questionnaire Star): The Genre-based Second Language Writing Self-Efficacy Scale (GL2WSS). As mentioned in the above section, all the items of the current questionnaire were developed in English and translated into Chinese when presented to EFL students to guarantee that the items could be fully understood and potential misunderstandings avoided.

The scale was given to students to elicit authentic context-based information. Before answering the questions in the questionnaire, all the students were told that their answers would not be judged as right or wrong on specific criteria and that they would be highly appreciated if they could provide accurate reports of writing self-efficacy. They were also told that their responses to the survey would not have any impact on their course grade at all. Before distributing the survey links to the participants, the researcher reviewed and clarified the instructions. Any doubts and comments from the participants were recorded and addressed during and after responding to the questionnaire. On average, respondents spent approximately 5–8 min finishing the questions in the questionnaire. The responses of students in Sample A were utilized to explore the factorial structure of the GL2WSS scale through exploratory factor analysis, while those from Sample B were to identify the relationship between the measured variables and the constructs in the GL2WSS scale.

After completing the online survey, students in Sample B were required to finish two writing tasks mentioned in the above section. A total of 664 essays were collected to investigate EFL learners' writing achievement. Two EFL teachers who have taught English for at least 10 years in China and demonstrated excellent performance in scoring CET4 essays were invited to mark the collected essays independently under Jacobs et al. writing scoring rubrics. The inter-rater reliability between the raters was $r_{AB} = 0.939$, $p_{AB} = 0.000 < 0.05$ for argumentative writing, $r_{AB} = 0.885$, $p_{AB} = 0.000 < 0.05$ for narrative writing, indicating sound scoring reliability.

4.5. Data analysis

Data analysis included two phases: data preparation and instrument validation. Following the procedures proposed by Dörnyei and Taguchi (2010), data gathered from the questionnaire went through screening and cleaning first. In the data cleaning, the researcher performed corrections of as many errors and inaccuracies as possible, which included impossible answers, incorrectly entered answers, contradicting answers, and implausible data. The provided answers that indicated that participants who lacked effort, intentionally misbehaved, or responded in an inaccurate fashion were deleted from the database. The checking and cross-examination of missing data were done through a manual inspection first. Listwise deletion was adopted to remove all the cases of missing data using Microsoft Excel 2016. After data screening and cleaning, no participants were removed from participation in the pilot study. Therefore, 664 participants were retained for the final analysis in the pilot study.

The normality, linearity, and homogeneity of variance of the data were carefully checked prior to the actual multivariate analysis. As a common practice, Mardia's kurtosis and/or skewness were adopted as a reference to check the multi-normality of the collected data. Generally speaking, when the critical ratio for Mardia's skewness and kurtosis is less than 1.96, the data are assumed to be multi-normally distributed (Tavakoli, 2012); otherwise, the data would not show the property of multi-normality. The examination of Mardia's kurtosis and skewness was conducted with the help of Stata 8.4.

In the instrument validation phase, the researcher adopted statistically rigorous procedures to scrutinize the reliability and the construct validity of the questionnaire in this study by running such analyses as factor analysis on the collected data. The reliability (also named internal consistency) of the questionnaire was measured by a reliability coefficient, Cronbach's alpha. The construct validity was examined through two sub-constructs: convergent validity and discriminant validity.

Exploratory factor analysis (EFA) was utilized initially to explore the underlying factors or components of the newly designed scale, the Genre-based Second Language Writing Self-Efficacy Scale (GL2WSS). In EFA, the specific technique of a maximum-likelihood analysis with oblique rotation was employed (O'Connor, 2000). After that, confirmatory factor analysis (CFA) was applied to examine the relationship between the measured variables and the constructs or factors in the GL2WSS following the advice in Tavakoli (2012). IBM SPSS 25 was utilized to conduct EFAs on the aforementioned questionnaire, whereas MPlus 8.3, a latent variable modeling program offering various estimation methods for normal and non-normal data (Muthén and Muthén, 2018), was employed to perform CFAs on the questionnaire.

The convergent validity and discriminant validity of the questionnaire were examined by the combination of the average variance extracted and composite reliability for each individual factor involved in the questionnaire. The average variance extracted was employed to measure "convergence among a set of items representing a reflectively measured latent construct" (Hair et al., 2019, p.659), while composite reliability was to measure "reliability and internal consistency of the measured variables representing a latent construct" (Hair et al., 2019, p.659). The average variance extracted and the composite reliability were calculated with the help of an online tool, which is available at <https://mlln.cn/>. The critical value for the average variance extracted and the composite reliability are 0.5 and 0.7,

respectively (Hair et al., 2019). In contrast, discriminant validity was examined by comparing the squared root of the average variance extracted and the correlational coefficient of the factors involved in the questionnaire. If the squared root of the average variance extracted is large than the correlational coefficients, it might indicate the sound discriminant validity of the questionnaire. Otherwise, it might indicate the opposite. In addition, the predictive validity of the GL2WSS scale was examined by performing regression analyses on factors in the scale and writing quality of argumentative and narrative writing, respectively.

5. Results

5.1. Descriptive statistics

In order to show the trend in the collected data, we need to see the distribution of participants' responses to these items in the GL2WSS. As shown in Table 1, the mean scores of all 24 items involved in GL2WSS were in the range of 3.71 to 4.75, coupled with standard deviations ranging from 1.168 to 1.49. The values for the skewness and kurtosis for all the items ranged from -0.643 to 0.086 and from -0.581 to 0.331 , respectively. According to the critical/cut-off values of ± 3.0 and ± 8.0 for skewness and kurtosis,

TABLE 1 Descriptive statistics of GL2WSS.

	Mean	SD	Skewness	Kurtosis
Item 1	4.01	1.438	-0.029	-0.554
Item 2	4.01	1.441	-0.089	-0.581
Item 3	4.75	1.375	-0.643	-0.022
Item 4	4.28	1.359	-0.321	-0.241
Item 5	4.31	1.297	-0.361	-0.13
Item 6	4.49	1.3	-0.535	0.029
Item 7	4.44	1.221	-0.32	-0.061
Item 8	4.58	1.203	-0.345	-0.127
Item 9	4.38	1.207	-0.32	0.048
Item 10	4.42	1.233	-0.145	-0.345
Item 11	3.9	1.351	-0.049	-0.458
Item 12	3.71	1.49	0.086	-0.524
Item 13	3.77	1.422	-0.048	-0.401
Item 14	4.03	1.372	-0.053	-0.158
Item 15	4.2	1.28	-0.258	-0.108
Item 16	4.22	1.281	-0.244	-0.095
Item 17	4.55	1.189	-0.348	0.159
Item 18	4.43	1.233	-0.441	0.291
Item 19	4.39	1.225	-0.389	0.331
Item 20	4.37	1.231	-0.364	0.29
Item 21	4.4	1.27	-0.252	0.04
Item 22	4.25	1.205	-0.105	0.141
Item 23	4.17	1.246	-0.269	0.017
Item 24	4.17	1.168	-0.413	0.304

respectively (Kline, 2016), the responses to the items showed the property of normal distribution.

5.2. Factors extracted through exploratory factor analysis

The sampling adequacy was verified by the Kaiser-Meyer-Olkin (KMO) measure, the result of which ($KMO=0.944$) shows that the valid sample size of 332 was sufficient for factor analysis. The strength of the correlations between items in the GL2WSS scale was measured by Bartlett's test of sphericity, the result of which ($df=276, p<0.001$) indicates that these correlations were large enough to run factor analysis. Compared with other estimation methods, the maximum likelihood estimation is advantageous because it "allows for the computation of a wide range of indexes of the goodness of fit of the model, permits statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals" (Fabrigar et al., 1999, p. 277). Therefore, the maximum likelihood estimation was conducted on all 24 items via oblique rotation to extract factors. Following (O'Connor's 2000) SPSS commands for parallel analysis, four predominant factors were extracted from the maximum likelihood estimation, explaining 63.442% of the cumulative variance. The four-factor solution was further examined to eliminate unsatisfactory items, including hyperplane items and some irrelevant items with low loading and complex loading. Following the recommended benchmark (± 0.5) for interpretability (Comrey and Lee, 2016), we retained the items with loading larger than 0.5. Eight items (items 1, 10, 15, 16, 17, 18, 20, 24, on the initial list) were eliminated because of complex loading (items that load at 0.5 or higher on two factors). Therefore, the other 16 items with loading larger than the benchmark were retained as the final version of the GL2WSS scale.

The revised GL2WSS scale, including 16 retained items, was re-assessed by employing the maximum likelihood estimation with oblique rotation, and the four-factor solution was confirmed ($KMO=0.929, df=171, p<0.001$), explaining 64.852% of the total variance. No hyperplane items or items with complex and low loading were detected in the revised GL2WSS scale. Through the thematic analysis of items grouped around each factor, four categories of writing self-efficacy were identified and labeled: Factor 1 was labeled as *Linguistic Self-Efficacy* (47.318% variance); Factor 2 as *Classroom Performance Self-Efficacy* (7.617% variance); Factor 3 as *Genre-Based Performance Self-Efficacy* (5.663% variance); and Factor 4 as *Self-Regulatory Self-Efficacy* (4.254% variance). The final version of the 16-item GL2WSS scale and standardized factor loadings for those items, together with Cronbach's alpha coefficients, are shown in Table 2.

Table 2 shows that Cronbach's alpha coefficients for the four factors were 0.884 for linguistic self-efficacy, 0.901 for classroom performance self-efficacy, 0.903 for genre-based performance self-efficacy, and 0.858 for self-regulatory self-efficacy, which were larger than the critical value of no less than 0.70 for satisfactory reliability.

5.3. Four-factor correlated models through confirmatory factor analysis

Before conducting CFA on all 16 items in the GL2WSS scale, we examined the multivariate normality by using Mardia's kurtosis

TABLE 2 Factor loadings for three-factor model after EFA and internal reliability.

	Items	Factor loading				α
		1	2	3	4	
Linguistic self-efficacy (LS)	Item 3	0.765				0.884
	Item 4	0.757				
	Item 5	0.754				
	Item 6	0.654				
	Item 2	0.514				
Class performance self-efficacy (CPS)	Item 12		−0.942			0.901
	Item 13		−0.902			
	Item 14		−0.643			
	Item 11		−0.578			
Genre-based performance self-efficacy (GPS)	Item 21			0.844		0.903
	Item 23			0.822		
	Item 22			0.81		
	Item 19			0.591		
Self-regulatory self-efficacy (SRS)	Item 9				−0.84	0.858
	Item 8				−0.74	
	Item 7				−0.53	

α = Cronbach's alpha coefficient.

value, whose critical ratio is 1.96 (Raykov and Marcoulides, 2008). The critical ratio of Mardia's kurtosis value in the current study was 56.517, which is larger than the cut-off point, suggesting that the responses to the GL2WSS scale are multivariate non-normal. Therefore, the maximum likelihood estimation with robust standard errors proposed in Mplus (Muthén and Muthén, 2018) was employed here to examine the factorial structure of the GL2WSS.

The Four-Factor Models were constructed on the basis of the result of EFA in the above section. It specified 16 items into four distinct but correlated writing self-efficacy. Initially, Model 1 was generated where each indicator was constrained to load only the factor it was designed to measure; covariance for each factor pair was freely estimated, and measurement error for each indicator was also freely estimated and uncorrelated. Model 2 was then constructed by correlating the factors in the scale. Models 1 and 2 were subjected to omnibus fit statistical analyses, and fit indices for them were compared, as shown in Table 3.

Fit indices in Table 3 show that Model 2 demonstrated a more satisfactory model fit than Model 1 ($\chi^2=210.608; df=131; p<0.001$; $\chi^2/df=1.608$; CFI=0.969; TLI=0.96; RMSEA=0.043 [0.032, 0.053]). Model 2, the final Four-factor Correlated Model, is presented in Figure 1.

Table 4 shows that the parameter estimates for all 16 items were statistically significant at $p<0.001$, and standardized loadings of the items on the corresponding latent factors ranged from 0.687 to 0.908, which are higher than the recommended value of 0.50, indicating the large effect size (Raykov and Marcoulides, 2008). This might suggest that the latent factors showed sound representativeness of the corresponding items. The average variance extracted values for genre-based performance self-efficacy, linguistic self-efficacy, self-regulatory self-efficacy, and classroom performance self-efficacy were larger than 0.5, while their composite reliability estimates were higher than 0.8,

TABLE 3 Fit indices for models 1 and 2.

	χ^2	df	CFI	TLI	RMSEA	AIC	BIC
Model 1	327.837	146	0.930	0.918	0.061	17242.666	17482.389
Model 2	210.608	131	0.969	0.960	0.043	17083.864	1738.664

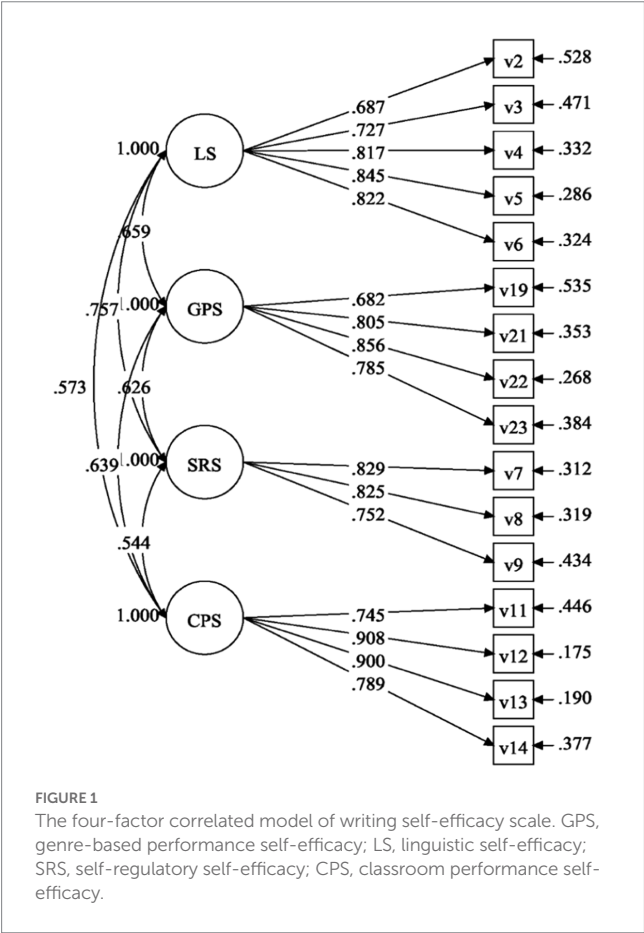


FIGURE 1 The four-factor correlated model of writing self-efficacy scale. GPS, genre-based performance self-efficacy; LS, linguistic self-efficacy; SRS, self-regulatory self-efficacy; CPS, classroom performance self-efficacy.

both of which might indicate the sound convergent validity of the GL2WSS scale.

Table 5 demonstrates that latent factors of genre-based performance self-efficacy, linguistic self-efficacy, self-regulatory self-efficacy, and classroom performance self-efficacy had a significant correlation at $p < 0.01$. Although the correlations across the above latent factors were more than 0.5, they were less than square roots of average variance extracted values for latent factors. All these estimates might suggest that latent factors had correlations to each other and a degree of differentiation, indicating a good discriminate validity of the GL2WSS scale.

5.4. Predictive value of GL2WSS

Stepwise regression analysis was employed to examine the predictive validity of the GL2WSS scale. We listed the scores of argumentative and narrative writing tasks, respectively, as dependent variables. Specifically, stepwise regression analyses were carried out where only genre-based performance self-efficacy entered into the regression model and other three factors (i.e., linguistic self-efficacy,

TABLE 4 CFA standardized regression weights for the four-factor correlated model of writing self-efficacy.

	Estimate	S.E.	p	AVE	CR
v2 \leftarrow LS	0.687	0.041	***	0.6086	0.8854
v3 \leftarrow LS	0.727	0.036	***		
v4 \leftarrow LS	0.817	0.034	***		
v5 \leftarrow LS	0.845	0.024	***		
v6 \leftarrow LS	0.822	0.027	***		
v7 \leftarrow SRS	0.829	0.03	***	0.5889	0.8505
v8 \leftarrow SRS	0.825	0.035	***		
v9 \leftarrow SRS	0.752	0.04	***		
v11 \leftarrow CPS	0.745	0.037	***	0.6671	0.889
v12 \leftarrow CPS	0.908	0.018	***		
v13 \leftarrow CPS	0.9	0.021	***		
v14 \leftarrow CPS	0.789	0.036	***		
v19 \leftarrow GPS	0.715	0.046	***	0.5639	0.8857
v21 \leftarrow GPS	0.785	0.041	***		
v22 \leftarrow GPS	0.766	0.044	***		
v23 \leftarrow GPS	0.727	0.045	***		

*** = $p < 0.001$; S.E., standard error; AVE, average variance extracted; CR, composite reliability; GPS, genre-based performance self-efficacy; LS, linguistic self-efficacy; SRS, self-regulatory self-efficacy; CPS, classroom performance self-efficacy.

self-regulatory self-efficacy, and classroom performance self-efficacy) did not enter into the model because of the low predictive values. Besides, another regression analysis was conducted to examine the predictive effects of overall writing self-efficacy on the quality of both argumentative and narrative essays. The results of these analyses are shown in Table 6.

It can be seen from the statistics in Table 6 that genre-based performance self-efficacy contributed to the writing quality of argumentative and narrative essays ($R^2_{\text{argumentation}} = 0.053$, $p_{\text{argumentation}} = 0.000$; $R^2_{\text{narration}} = 0.021$, $p_{\text{narration}} = 0.009$). Specifically, genre-based performance self-efficacy could explain 5.3% of the variability of the writing quality of argumentative essays but only 2.1% of that of narrative essays. In other words, students with higher levels of genre-based performance self-efficacy might write better argumentative and narrative essays than those with lower levels of genre-based performance self-efficacy. In contrast, other three factors generated insignificant predictive effects on the quality of the argumentative writing essays ($\beta_{\text{linguistic self-efficacy}} = 0.038$, $p = 0.564$; $\beta_{\text{self-regulatory self-efficacy}} = -0.1$, $p = 0.12$; $\beta_{\text{classroom performance self-efficacy}} = -0.014$, $p = 0.841$) and that of the narrative ones ($\beta_{\text{linguistic self-efficacy}} = -0.009$, $p = 0.89$; $\beta_{\text{self-regulatory self-efficacy}} = -0.053$, $p = 0.434$; $\beta_{\text{classroom performance self-efficacy}} = 0.030$, $p = 0.683$).

Additionally, Table 6 reveals that genre-based performance self-efficacy contributed to the writing quality of argumentative and narrative essays ($R^2_{\text{argumentation}} = 0.035$, $p_{\text{argumentation}} = 0.000$; $R^2_{\text{narration}} = 0.011$, $p_{\text{narration}} = 0.029$). It can be found from Table 6 that

overall writing self-efficacy had significant predictive effects on the quality of the argumentative essays ($\beta = 0.186, p = 0.001$) and that of the narrative ones ($\beta = 0.12, p = 0.029$).

6. Discussion

This study aims to tackle the issue of task-specificity by incorporating genre characteristics into the design and development of items of the GL2WSS. The results of EFA and CFA in the above section support the factorial structure of the newly developed GL2WSS, including four categories of self-efficacy: genre-based performance self-efficacy, linguistic self-efficacy, self-regulatory self-efficacy, and classroom performance self-efficacy. Scores of four sub-constructs were calculated separately to reveal the level of participants' perceived writing self-efficacy and collectively summed to show each student's overall level of each sub-construct in terms of linguistics, self-regulation, task, and situation. The results might offer preliminary evidence for including task characteristics (e.g., genre features) in the conceptualization of writing self-efficacy. Therefore, it might provide a tentative solution to the issue of task-specificity proposed by Bandura (2006), thus enabling the writing self-efficacy scale to be task-specific.

The findings of this study might also provide initial evidence for extending the conceptualization of writing self-efficacy in the L2 context. The models specified by SEM confirmed the conceptualization of writing self-efficacy as four distinctive but correlated sub-constructs subsumed under the construct of writing self-efficacy: linguistic self-efficacy, self-regulatory self-efficacy, classroom performance self-efficacy, and genre-based performance self-efficacy. The first three

sub-constructs were consistent with Teng et al. (2018) conceptualization of writing self-efficacy in the L2 context. Furthermore, as mentioned before, genre features could impose constraints on written discourses, for example, employing certain linguistic patterns to achieve the persuasive purpose, thus constituting the high-level demands for writing essays, both in L1 and L2 contexts. Therefore, incorporating genre features into the conceptualization of writing self-efficacy might cover the judgment of students' confidence in meeting higher writing requirements that were left untouched in the previous studies concerning the development of writing self-efficacy scales. The finding that the GL2WSS scale demonstrated sound convergent validity might suggest that linguistic self-efficacy, self-regulatory self-efficacy, classroom performance self-efficacy, and genre-based performance self-efficacy were correlated. Moreover, the findings of this study also revealed that the GL2WSS scale demonstrated sound discriminant validity, which might suggest that four categories of self-efficacy entailed in the GL2WSS scale were conceptually and empirically distinguished. Accordingly, it can be concluded that the inclusion of genre characteristics into the conceptualization of writing self-efficacy is empirically validated. Meanwhile, the findings of this study also corroborated the multidimensional nature of writing self-efficacy in the L2 context. Overall, compared with previous studies (e.g., Teng et al., 2018; Sun and Wang, 2020), the findings of this study might advance our understanding of the multidimensional nature of writing self-efficacy by entailing task characteristics (i.e., genre features). Thus, writing self-efficacy could be conceptualized in terms of linguistic skills, self-regulation, tasks, and situations.

Furthermore, the results of this study indicated that compared with the other sub-constructs subsumed in the GL2WSS scale, genre-based performance self-efficacy had a significant predictive effect on the quality of argumentative and narrative essays in the L2 context. That is, students who are efficacious in genre-based performance might show better performance in writing argumentative and narrative essays than those who are not. The better predictive effect of genre-based performance self-efficacy could be attributed to the fact that the items tapping into this category of writing self-efficacy were more closely relevant or matched to the specific genre features of writing tasks. The genre features of argumentative and narrative essays were entailed in the development of the GL2WSS scale, while students were required to write essays in two genres: argumentation and narration. The match between the genre features of writing tasks and those entailed in the GL2WSS scale could facilitate students to

TABLE 5 Square root of AVEs and correlation matrix of four factors.

	LS	GPS	SRS	CPS
LS	0.6086			
GPS	0.626	0.5639		
SRS	0.752	0.597	0.5889	
CPS	0.63	0.657	0.62	0.6671
Square root of AVE	0.78	0.751	0.767	0.817

Average variance extracted values were shown in diagonals. GPS, genre-based performance self-efficacy; LS, linguistic self-efficacy; SRS, self-regulatory self-efficacy; CPS, classroom performance self-efficacy.

TABLE 6 Predictive effects of GL2WSS on writing quality.

Regression models and predictors	Writing quality							
	Argumentation				Narration			
	B	β	t	p	B	β	t	p
Regression	$R^2 = 0.053 F = 18.539 p = 0.000$				$R^2 = 0.021 F = 6.917 p = 0.009$			
Genre-based performance self-efficacy	0.359	0.231	4.306	0.000	0.121	0.143	2.630	0.009
Linguistic self-efficacy		0.038	0.578	0.564		-0.009	-0.138	0.890
Self-regulatory self-efficacy		-0.100	-1.56	0.120		-0.053	-0.783	0.434
Classroom performance self-efficacy		-0.014	-0.200	0.841		0.030	0.409	0.683
Regression	$R^2 = 0.035 F = 11.865 p = 0.001$				$R^2 = 0.011 F = 4.838 p = 0.029$			
Overall writing self-efficacy	0.087	0.186	3.445	0.001	0.037	0.12	2.20	0.029

make a more accurate judgement about their performance in certain writing tasks. Therefore, the findings of this study might provide another independent evidence for the predictive effects of writing self-efficacy on writing performance as reported in the literature. On the whole, the findings provided substantial evidence for the utility of the GL2WSS scale as an effective measurement of writing self-efficacy.

Additionally, the differences in the predictive effects of genre-based performance self-efficacy and overall writing self-efficacy on the quality of the argumentative essays and that of the narrative ones might be attributed to the practice effect. As mentioned before, argumentative writing has been set in national tests (e.g., College English tests and Tests for English majors in China) and international ones (e.g., The International English Language Testing System and The Test of English as a Foreign Language). Therefore, instructors would focus on the teaching of argumentative writing due to the washback effects of these tests. Naturally, argumentative writing tasks might be frequently assigned to students and they may practice them accordingly, thus their skills in writing argumentative essays being sharpened.

7. Conclusion

This study is designed to develop and validate a new scale for assessing writing self-efficacy in L2 contexts by incorporating genre features of writing tasks. Statistical analyses demonstrated that the newly developed GL2WSS scale demonstrated sound psychometric qualities, including good reliability, sound factorial structure, convergent validity, and discriminate validity. The findings of this study that the GL2WSS scale entailing linguistic self-efficacy, self-regulatory self-efficacy, classroom performance self-efficacy, and genre-based performance self-efficacy offered further evidence to support the multidimensional conceptualization of writing self-efficacy in a specific L2 context. The results of this study might also suggest that writing self-efficacy could be conceptualized in terms of linguistic skills, self-regulation, tasks, and situations, thus providing a tentative solution to the issue of task specificity of writing self-efficacy and offering initial evidence to support that writing self-efficacy is context sensitive.

The GL2WSS scale could be employed as a pedagogical tool in the classroom to facilitate teachers and students of EFL writing in assessing different aspects of writing self-efficacy. The inclusion of task-specific features such as text genre in the newly developed writing self-efficacy scale might help students make more accurate judgments of writing self-efficacy. The GL2WSS scale might offer students an opportunity to understand their writing capabilities from linguistic, classroom performance, genre-based performance, and self-regulatory aspects, all of which could motivate them to enhance their writing proficiency. Besides, teachers are advised to use this scale to know the profile of their students' writing self-efficacy, which could facilitate teachers to adjust their writing instructions to engage students with achievement and enjoyment. The GL2WSS scale might be utilized (a) to elicit students' writing self-efficacy to provide the guidelines for designing curriculum and teaching activities in writing courses; (b) to examine the underlying factors of writing self-efficacy; (c) to gauge the effectiveness of teaching interventions; (d) to evaluate the preciseness of students' judgments of writing competence and help them to align them if necessary.

Despite the fact that optimistic findings were generated from this study, several limitations should be recognized due to the constraints of experimental conditions, experimental methods and available resources. To begin with, the target population in this study were a sample of Chinese EFL writers from two medium-ranking universities, and the four-factor structure of writing self-efficacy may not be valid for other L2 cohorts. Therefore, the GL2WSS scale might need further validation and refinements to suit different populations in different learning contexts. Besides, we entailed the genre features of argumentative and narrative writing without involving other genres. Therefore, it is recommended that the characteristics of other genres be incorporated, or the genre features of the GL2WSS scale be tailored to match the writing tasks employed to examine students' writing performance. In addition, there are three approaches available to examine criterion validity: retrospective validity, predictive validity, and concurrent validity (Horstmann et al., 2020). This study only focused on the predictive validity of the GL2WSS scale, thus leaving its retrospective validity and concurrent validity unexamined. Therefore, it might be interesting to investigate the extent to which writing self-efficacy and other motivational constructs, such as motive to write, might be conceptually and empirically distinguished or the relationship between writing self-efficacy and the criteria set previously.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by The University of Auckland Human Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JZ and LZ conceived and designed the study. JZ collected and analyzed the data and drafted the manuscript, and all the authors revised and approved the manuscript. YZ and LZ finalized it. All authors contributed to the article and approved the submitted version.

Funding

This work is funded by the Philosophy and Social Science Foundation of Shanghai, China (Grant no 2019BYY009).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations,

or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *J. Soc. Clin. Psychol.* 4, 359–373. doi: 10.1521/jscp.1986.4.3.359
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman/Macmillan Higher Education.
- Bandura, A. (2006). "Guide for constructing self-efficacy scales" in *Adolescence and education. Self-efficacy beliefs of adolescents*. eds. F. Pajares and T. Urdan (Charlotte, NC: Information Age Publishing), 307–338.
- Boscolo, P., and Hidi, S. (2007). "The multiple meanings of motivation to write" in *Studies in writing and motivation*. eds. S. Hidi and P. Boscolo (Netherlands: Elsevier), 1.
- Bruning, R., Dempsey, M. S., Kauffman, D. F., McKim, C., and Zumbun, S. (2013). Examining dimensions of self-efficacy for writing. *J. Educ. Psychol.* 105, 25–38. doi: 10.1037/a0029692
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C., et al. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *J. Educ. Psychol.* 100, 525–534. doi: 10.1037/0022-0663.100.3.525
- Chen, J., and Zhang, L. J. (2019). Assessing student-writers' self-efficacy beliefs about text revision in EFL writing. *Assess. Writ. Q.* 40, 27–41. doi: 10.1016/j.asw.2019.03.002
- Chen, J., Zhang, L. J., Wang, X., and Zhang, T. T. (2021). Influences of SRSD revision instruction on English-as-a-foreign-language (EFL) students' self-efficacy for text revision: a mixed-methods study. *Front. Psychol.* 12, 1–16. doi: 10.3389/fpsyg.2021.670100
- Chen, J., Zhang, L. J., and Parr, J. M. (2022). Improving EFL students' text revision with the self-regulated strategy development (SRSD) model. *Metacogn. Learn.* 17, 191–211. doi: 10.1007/s11409-021-09280-w
- Comrey, A. L., and Lee, H. B. (2016). *A first course in factor analysis (2nd)*. East Sussex, England: Psychology Press.
- Dörnyei, Z., and Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing (2nd)*. London: Routledge.
- Fabrigar, L. R., Wegener, D. T., Mac Callum, R. C., and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychol. Methods* 4, 272–299. doi: 10.1037/1082-989X.4.3.272
- Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educ. Psychol.* 53, 258–279. doi: 10.1080/00461520.2018.1481406
- Graham, S., and Harris, K. R. (1989). Improving learning disabled students' skills at composing essays: self-instructional strategy training. *Except. Child.* 56, 201–214. doi: 10.1177/001440298905600305
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). *Multivariate data analysis (8th)*. Boston, MA: Cengage.
- Hayes, J. (1996). "A new framework for understanding cognition and affect in writing" in *The science of writing: Theories, methods, individual differences and applications*. eds. C. M. Levy and S. Ransdell (Hillsdale, NJ: Lawrence Erlbaum), 1–29.
- Hayes, J., and Flower, L. (1980). "Identifying the organization of writing processes" in *Cognitive processes in writing*. eds. L. Gregg and E. Steinberg (London: Routledge), 3–30.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Writ. Commun.* 29, 369–388. doi: 10.1177/0741088312451260
- Horstmann, K. T., Knaut, M., and Ziegler, M. (2020). "Criterion validity" in *Encyclopedia of personality and individual differences*. eds. V. Zeigler-Hill and T. K. Shackelford (New York, NY: Springer International Publishing), 1–3.
- Huang, Y., and Zhang, L. J. (2020). Does a process-genre approach help improve students' argumentative writing in English as a foreign language? Findings from an intervention study. *Read. Writ. Q.* 36, 339–364. doi: 10.1080/10573569.2019.1649223
- Jacobs, H. L., Zingraf, S. A., Wormuth, D. R., Hartfiel, V. F., and Hughey, J. B. (1981). *Testing ESL Composition: A practical approach. English Composition Program*. New York, NY: Newbury House. <https://eric.ed.gov/?id=ed217708>
- Kirsch, I. (1985). Self-efficacy and expectancy: old wine with new labels. *J. Pers. Soc. Psychol.* 49, 824–830. doi: 10.1037/0022-3514.49.3.824
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. (4th, ed.). New York, NY: The Guilford Press.
- Li, H. H., and Zhang, L. J. (2022). Investigating effects of small-group student talk on the quality of argumentation in Chinese tertiary EFL students' individual writing. *Front. Psychol.* 13, 1–14. doi: 10.3389/fpsyg.2022.868045
- Limpo, T., Filipe, M., Magalhães, S., Cordeiro, C., Veloso, A., Castro, S. L., et al. (2020). Development and validation of instruments to measure Portuguese third graders' reasons to write and self-efficacy. *Read. Writ. Q.* 33, 2173–2204. doi: 10.1007/s11145-020-10039-z
- Matthews, P. H. (2010). Factors influencing self-efficacy judgments of university students in foreign language tutoring. *Mod. Lang. J.* 94, 618–635. doi: 10.1111/j.1540-4781.2010.01057.x
- Muthén, B., and Muthén, L. (2018). Mplus. In Linden W. J. van der (Eds.), *Handbook of item response theory* (pp. 507–518). London: Chapman and Hall/CRC Press.
- McCarthy, P., Meier, S., and Rinderer, R. (1985). Self-efficacy and writing: a different view of self-evaluation. *Coll. Compos. Commun.* 36, 465–471. doi: 10.2307/357865
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behav. Res. Methods Instrum. Comput.* 32, 396–402. doi: 10.3758/BF03200807
- Ong, J., and Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *J. Second. Lang. Writ.* 19, 218–233. doi: 10.1016/j.jslw.2010.10.003
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Read. Writ. Q.* 19, 139–158. doi: 10.1080/10573560308222
- Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Meas. Eval. Couns. Dev.* 39, 239–249. doi: 10.1080/07481756.2007.11909801
- Pajares, F., Hartley, J., and Valiante, G. (2001). Response format in writing self-efficacy assessment: greater discrimination increases prediction. *Meas. Eval. Couns. Dev.* 33, 214–221. doi: 10.1080/07481756.2001.12069012
- Pajares, F., and Johnson, M. J. (1994). Confidence and competence in writing: the role of self-efficacy, outcome expectancy, and apprehension. *Res. Teach. Engl.* 28, 313–331.
- Pajares, F., and Johnson, M. J. (1996). Self-efficacy beliefs and the writing performance of entering high school students. *Psychol. Sch.* 33, 163–175. doi: 10.1002/(sici)1520-6807(199604)33:2<163::aid-pits10>3.0.co;2-c
- Pajares, F., Miller, M. D., and Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *J. Educ. Psychol.* 91, 50–61. doi: 10.1037/0022-0663.91.1.50
- Pajares, F., and Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *J. Educ. Res.* 90, 353–360. doi: 10.1080/00220671.1997.10544593
- Pajares, F., and Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemp. Educ. Psychol.* 24, 390–405. doi: 10.1006/ceps.1998.0995
- Pajares, F., and Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: a function of gender orientation? *Contemp. Educ. Psychol.* 26, 366–381. doi: 10.1006/ceps.2000.1069
- Prat-Sala, M., and Redford, P. (2012). Writing essays: does self-efficacy matter? The relationship between self-efficacy in reading and in writing and undergraduate students' performance in essay writing. *Educ. Psychol.* 32, 9–20. doi: 10.1080/01443410.2011.621411
- Rahimi, M., and Zhang, L. J. (2021). Effects of an engaging process-genre approach on student engagement and writing achievements. *Read. Writ. Q.* 38, 487–503. doi: 10.1080/10573569.2021.1982431
- Raykov, T., and Marcoulides, G. A. (2008). *Introduction to applied multivariate analysis*. London: Routledge.
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educ. Psychol.* 26, 207–231. doi: 10.1080/00461520.1991.9653133
- Schunk, D. H., and Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Read. Writ. Q.* 23, 7–25. doi: 10.1080/10573560600837578
- Shell, D. F., Colvin, C., and Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: grade-level and achievement-level differences. *J. Educ. Psychol.* 87, 386–398. doi: 10.1037/0022-0663.87.3.386
- Shell, D. F., Murphy, C. C., and Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *J. Educ. Psychol.* 81, 91–100. doi: 10.1037/0022-0663.81.1.91
- Sun, T., and Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System* 90, 102221–102217. doi: 10.1016/j.system.2020.102221
- Tavakoli, H. (2012). *A dictionary of research methodology and statistics in applied linguistics*. Tehran: Rahnamā Press.

- Teng, L. S., Sun, P. P., and Xu, L. (2018). Conceptualizing writing self-efficacy in English as a foreign language contexts: scale validation through structural equation modeling. *TESOL Q.* 52, 911–942. doi: 10.1002/tesq.432
- Usher, E. L., and Pajares, F. (2008). Self-efficacy for self-regulated learning. *Educ. Psychol. Meas.* 68, 443–463. doi: 10.1177/0013164407308475
- Wang, C., and Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Q.* 51, 931–947. doi: 10.1002/tesq.355
- Wingate, U., and Tribble, C. (2012). The best of both worlds? Towards an English for academic purposes/academic literacies writing pedagogy. *Stud. High. Educ.* 37, 481–495. doi: 10.1080/03075079.2010.525630
- Yang, L. F., and Zhang, L. J. (2023). Self-regulation and student engagement with feedback: a case study of Chinese EFL student writers. *J. Engl. Acad. Purp.* 63, 101226–101214. doi: 10.1016/j.jeap.2023.101226
- Yang, L. F., Zhang, L. J., and Dixon, H. R. (2023). Understanding the impact of teacher feedback on EFL students' use of self-regulated writing strategies. *J. Second. Lang. Writ.* 60, 101012–101015. doi: 10.1016/j.jslw.2023.101015
- Yang, W., Lu, X., and Weigle, S. C. (2015). Different topics, different discourse: relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *J. Second. Lang. Writ.* 28, 53–67. doi: 10.1016/j.jslw.2015.02.002
- Yoon, H. J. (2021). Challenging the connection between task perceptions and language use in L2 writing: genre, cognitive task complexity, and linguistic complexity. *J. Second. Lang. Writ.* 54:100857. doi: 10.1016/j.jslw.2021.100857
- Zhang, L. J. (2022). "L2 writing: toward a theory-practice praxis" in *Handbook of practical second language teaching and learning*. ed. E. Hinkel (London: Routledge), 331–343.
- Zhang, L. J., and Cheng, X. L. (2021). Examining the effects of comprehensive written corrective feedback on L2 EAP students' performance: a mixed-methods study. *J. Engl. Acad. Purp.* 54, 1–15. doi: 10.1016/j.jeap.2021.101043
- Zhang, T. T., and Zhang, L. J. (2021). Taking stock of a genre-based pedagogy: sustaining the development of EFL students' knowledge of the elements in argumentation and writing improvement. *Sustainability* 13, 1–25. doi: 10.3390/su132111616
- Zimmerman, B. J. (2000). Self-efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25, 82–91. doi: 10.1006/ceps.1999.1016
- Zimmerman, B. J., and Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *Am. Educ. Res. J.* 31, 845–862. doi: 10.3102/00028312031004845
- Zimmerman, B. J., and Cleary, T. J. (2006). Adolescents' development of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In *Adolescence and education: Self-efficacy beliefs of adolescents*. eds. F. Pajares and T. Urdan (Greenwich, CT: Information Age), 45–69.
- Zumbrunn, S., Broda, M., Varier, D., and Conklin, S. (2020). Examining the multidimensional role of self-efficacy for writing on student writing self-regulation and grades in elementary and high school. *Br. J. Educ. Psychol.* 90, 580–603. doi: 10.1111/bjep.12315

Appendix A: writing prompts

1. Computers and the Internet have improved the efficiency and quality of learning for university students in China. Do you agree or disagree with the statement? Support your position with reasons. You should write at least 150 words.
2. Describe one of your unforgettable English learning experiences. You should write at least 150 words.

Appendix B: genre-based L2 writing self-efficacy scale

In this part, we would like you to help us by answering the following questions concerning your writing self-efficacy. Please give your answers sincerely, as only this will guarantee the success of the investigation. Thank you very much for your help.

In the following section, we would like you to tell us how much you agree or disagree with the following statements by simply ticking (✓) a number from 1 to 7. We are interested in your real situation and attitudes. Please do not leave out any of the items.

Not at all true of me	Not true of me	Slightly not true of me	Neutral	Slightly true of me	True of me	Very true of me
1	2	3	4	5	6	7

For example:

I like English movies. 1 2 3 4 5 6 7.

2. I can correctly use all parts of speech (e.g., nouns, verbs, adjectives, etc.) in writing.	1 2 3 4 5 6 7
3. I can write a simple English sentence with grammatical structure.	1 2 3 4 5 6 7
4. I can write compound and complex English sentences with grammatical structure.	1 2 3 4 5 6 7
5. I can write a good English paragraph with topic sentence or main idea.	1 2 3 4 5 6 7
6. I can write an English composition with a clear organisation or structure.	1 2 3 4 5 6 7
7. I can realise my goal to improve my English writing.	1 2 3 4 5 6 7
8. I can think of my goals before English writing.	1 2 3 4 5 6 7
9. I can think of different ways to help me to plan before English writing.	1 2 3 4 5 6 7
11. I can get an excellent grade for writing in the English course.	1 2 3 4 5 6 7
12. I can understand the most difficult writing material presented in the English course.	1 2 3 4 5 6 7
13. I can understand the basic concepts of writing taught in the English course.	1 2 3 4 5 6 7
14. I can understand the most complex material of writing presented by the instructor of the English course.	1 2 3 4 5 6 7
19. I can write an English narrative that includes several things that happened to the characters.	1 2 3 4 5 6 7
21. I can write an English narrative that describes clearly how the events develop.	1 2 3 4 5 6 7
22. I can write an English argumentation that includes arguments.	
23. I can write an English argumentation that includes sufficient evidences to support the arguments.	1 2 3 4 5 6 7



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Giulia Vettori,
University of Florence, Italy
Ester Trigo-Ibáñez,
Universidad de Cádiz, Spain

*CORRESPONDENCE

Jamie L. Metsala
✉ Jamie.Metsala@mnsu.ca

RECEIVED 27 March 2023

ACCEPTED 16 May 2023

PUBLISHED 15 June 2023

CITATION

Metsala JL (2023) Longitudinal contributions of morphological awareness, listening comprehension, and gains in word reading fluency to later word- and text-reading fluency. *Front. Educ.* 8:1194879. doi: 10.3389/feduc.2023.1194879

COPYRIGHT

© 2023 Metsala. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Longitudinal contributions of morphological awareness, listening comprehension, and gains in word reading fluency to later word- and text-reading fluency

Jamie L. Metsala*

Faculty of Education, Mount Saint Vincent University, Halifax, NS, Canada

This study examined the contributions of morphological awareness, listening comprehension, and early gains in word reading fluency to later outcomes in word- and text-reading fluency. There were 83 participants in second and third grade who were followed across 18 months. Gains in word reading fluency across the first six months predicted both word- and text-reading fluency one year later, beyond variance accounted for by initial word reading fluency, phonological awareness, rapid naming, and two oral language skills. Initial morphological awareness predicted reliable additional variance in word- and text-reading fluency 18 months later. The contribution of listening comprehension was specific to outcomes in text reading fluency. In the last analyses, listening comprehension, but not morphological awareness, predicted unique variance in final text reading fluency beyond final word reading fluency. Findings are discussed in terms of the developmental time-course of reading fluency and the roles of the two oral language skills examined.

KEYWORDS

reading development, reading fluency, morphological awareness, listening comprehension, longitudinal research

Introduction

A hallmark of skilled reading is the automatic recognition of words as one reads passages (Stanovich, 2000). This seemingly effortless word reading frees cognitive resources for the more attention demanding, meaning-making cognitive activities involved in comprehending texts (LaBerge and Samuels, 1974; Perfetti, 2007). Text reading fluency, defined here as the accuracy and rate of reading sentences and passages (see also, Kim, 2015; Kim et al., 2021), has been shown to contribute to individual differences in reading comprehension for typically developing young readers (e.g., Kim et al., 2021) and for those with reading disabilities (e.g., Torgesen and Hudson, 2006; Metsala and David, 2022). Models of text reading fluency and its development, in turn, highlight a critical role for word reading fluency, the accuracy and speed of reading words in a list format (e.g., Klauda and Guthrie, 2008; Hudson et al., 2009, 2012; Norton and Wolf, 2012). Comprehensive models of reading fluency have also proposed a role for linguistic processes, such as morphological awareness and listening comprehension (e.g., Wolf and Katzir-Cohen, 2001; Hudson et al., 2009; Kim, 2015; Kirby and Bowers, 2017).

Verbal efficiency theory identified both accuracy and rate as instrumental to word reading fluency (Perfetti, 1988). Theories of word reading development delineate how readers acquire this efficient word reading over time, with a focus on early phonological decoding skills and then on the increasingly internalized orthographic representations of words and word parts (Ehri, 1998, 2014; Share, 2008). Word reading fluency has been reported to be the single largest contributor to text reading fluency for typically developing and impaired readers (Torgesen et al., 2001; Torgesen and Hudson, 2006; Hudson et al., 2012). A considerable body of research has focused on the association between phonological awareness and rapid naming in the development of word reading fluency (e.g., Papadopoulos et al., 2016; Landerl et al., 2019). There has been less research on a potential role for oral language skills in the development of word- and text-reading fluency (Shechter et al., 2018; Manolitsis et al., 2019); therefore, the contributions proposed in some models have not been adequately tested.

Several questions concerning reading fluency and the skills that contribute to it in the early elementary grades remain largely unanswered. Kim (2015) suggested that some emergent skills studied in the context of reading development may contribute uniquely to word reading fluency and others to text reading fluency. Furthermore, the developmental time-course of stable individual differences and continuing malleability in reading fluency is not well delineated. One goal of this study is to test whether gains in word reading fluency across six months of second and third grade, predict word- and text-reading fluency outcomes one year later, beyond initial word reading fluency and several literacy related skills. A second goal is to examine whether morphological awareness and listening comprehension at study outset, predict variance in word- and text-reading fluency 18 months later, beyond initial word reading fluency and reading-related skills.

The role of early skill and gains in word reading fluency to later fluency outcomes

Word- and text-reading fluency increases sharply following students' entry into second and third grade (Hasbrouck and Tindal, 2017). This steep incline for typical readers has been proposed as one reason why readers with dyslexia remain woefully behind in fluency, even when remediation closes achievement gaps in accuracy and comprehension (Torgesen and Hudson, 2006; Metsala and David, 2017). The window for getting readers who struggle back on a solid trajectory for word- and text-reading fluency appears to close early, perhaps as early as the first couple years of elementary school (Juel, 1988; Torgesen et al., 2001). It seems important then, to understand the contribution of individual differences in word reading fluency beyond grade 1 to later fluency outcomes. A critical question is whether fluency trajectories are more or less consolidated quite early for developing readers, or whether there are significant individual differences in continuing gains in word reading fluency that affect later fluency outcomes. In a somewhat related approach, Petscher and Kim (2011) reported that growth rates in oral reading fluency across first grade accounted for the most variance in third grade reading comprehension; however, when these students were in grades 2 and 3, the strongest relationship to third grade comprehension was their initial status in oral reading fluency, rather than growth. Although they examined comprehension outcomes, their study supports the

notion that critical individual differences in fluency may be solidified quite early in reading acquisition. The current study set out to investigate whether gains in word reading fluency over roughly six months of second and third grade account for individual differences in word and text fluency one year later, beyond variance accounted for by initial word reading fluency and by several additional reading-related skills.

Morphological awareness and reading fluency

The role of foundational language skills has been emphasized in multicomponent models of reading comprehension; for example, the Reading Systems Framework recognizes that knowledge of vocabulary, syntax, and morphology directly influence comprehension processes, such as the readers' construction of mental models (Perfetti and Stafura, 2014; Stafura and Perfetti, 2017; for a focus on morphology, see Levesque et al., 2021). Alongside orthography and phonology, comprehensive models of reading fluency have proposed additional roles for semantic, syntactic, and morphological knowledge systems (Wolf and Katzir-Cohen, 2001; Hudson et al., 2009, 2012; Kim, 2015). One focus of the current study is on morphological awareness, an individuals' ability to recognize and manipulate the smallest meaningful units in words (Carlisle, 2000). Findings have been relatively consistent concerning the unique contribution of morphological awareness to word reading accuracy (e.g., Kirby et al., 2012; cf. McBride-Chang et al., 2005) and reading comprehension (e.g., Deacon and Kieffer, 2018; Metsala et al., 2021). Theoretical models also propose a role for morphological awareness in word- and text-reading fluency. Awareness of morphemes is implicated in interactive models for which word reading fluency is a result of connections among orthographic, phonological and semantic representations (Seidenberg and McClelland, 1989; Wolf and Katzir-Cohen, 2001). Moreover, morphological awareness may play a role in text reading fluency as one component of top-down linguistic influences potentially related to spreading activation facilitating reading sentences and passages (Hudson et al., 2009).

Despite these theoretical accounts, findings have been somewhat inconsistent in examinations of the associations between individual differences in morphological awareness and word reading fluency. Consider findings concerning the prediction of concurrent skills. Morphological awareness did not predict unique variance in first- and second-grade students' concurrent English word reading fluency with phonological awareness also in the equation (Apel et al., 2013). In contrast, other studies reported that morphological awareness predicted unique variance in concurrent English word reading fluency for first or third grade students, also with phonological awareness controlled (Kirby et al., 2012; Manolitsis et al., 2019).

Similar inconsistencies are found in longitudinal research across these early elementary school years. Students' fall morphological awareness predicted unique variance in spring word reading fluency for second grade English and French readers, with phonological awareness and rapid naming controlled, but this was not the case for young readers in Greek (Desrochers et al., 2018; see also, Georgiou et al., 2008; Diamanti et al., 2017; Giazitzidou et al., 2023 for similar findings with young Greek students; cf. Giazitzidou and Padelidi, 2022). Kirby et al. (2012) also found that morphological awareness predicted English word reading fluency for students followed from

second to third grade; although the relationship from first to third grade was not significant, as was the case for younger readers learning to read in Korean (Kim, 2015).

Only a handful of studies have controlled for initial word reading fluency in this literature. For English (but not for Greek) readers, morphological awareness at the end of grade 2 predicted word reading fluency at the beginning of grade 3 beyond initial word reading fluency and the other controls in the model (Manolitsis et al., 2019); however, even for the English readers in this sample, unique contributions to word reading fluency were not supported from the end of grade 1 to beginning of grade 2, nor across the grade 2 academic year. Similarly, morphological awareness did not account for unique variance in Greek word reading fluency from first to second grade after controlling for either initial word reading fluency or for several other reading related skills (Manolitsis et al., 2017).

Theoretical models also propose a role for morphological awareness in text reading fluency (e.g., Wolf and Katzir-Cohen, 2001), but little research has examined this directly (cf. Kirby et al., 2012; Kim and Wagner, 2015). Kirby et al. (2012) found that grade 2 morphological awareness predicted variance in third grade text reading fluency. For young children learning to read in Korean, morphological awareness did not predict text reading fluency one year later, with initial word reading fluency and other reading related skills in the model (Kim, 2015).

Inconsistent findings concerning the association between skills and reading may be explained, in part, by the nature of the orthographic system under investigation (Ziegler et al., 2010; Lee et al., 2022). Morphological awareness appears to play a more prominent role in word reading fluency in English and similar relatively opaque orthographies than in more transparent orthographies (Shechter et al., 2018). However, as reviewed here, inconsistencies remain among studies on learning to reading in English, the language examined in the current study. Furthermore, while longitudinal studies have routinely controlled for phonological awareness, there is less consistency for including initial word reading fluency or rapid naming; the former allows a look at gains over a specified period of time and, the latter has been shown to be uniquely associated with word reading fluency (e.g., Wolf and Bowers, 1999; Shechter et al., 2018; Landerl et al., 2019). Given the inconsistencies in the literature, the limited number of studies concerning text reading fluency, and the theoretical reasons to expect a role for morphological awareness in word- and text-reading fluency (Shechter et al., 2018), these relationships were examined in this study. In particular, the contribution of morphological awareness to later word- and text-reading fluency were tested, after variance accounted for by initial word reading fluency, gains in such, phonological awareness, and rapid naming were taken into account.

Listening comprehension and reading fluency

Linguistic comprehension, or understanding spoken sentences and discourse, is a major component of reading comprehension (Hoover and Gough, 1990). Research examining whether comprehension of spoken language also contributes to reading fluency is quite limited. Kim (2015) found that Korean listening comprehension was not uniquely associated with concurrent text reading fluency when the sample was younger (mean age was 5 yrs.,

2 mos). One year later, however, listening comprehension did predict concurrent text reading fluency beyond word reading fluency. In a longitudinal study with young English readers, first-grade students' fall listening comprehension predicted spring text reading fluency beyond fall word reading fluency; however, second-grade listening comprehension was not uniquely associated with text reading fluency one year later (Kim et al., 2021). This finding suggests that this relationship may be developmentally limited; that is, listening comprehension may facilitate text reading fluency only for young children whose single word reading is still relatively slow or inefficient. On the other hand, theories that propose automatic activation among semantically related words during reading or propose a top-down influence reflecting linguistic comprehension would lead us to expect an ongoing role for listening comprehension in text reading fluency (e.g., Wolf and Katzir-Cohen, 2001; Jenkins et al., 2003; Hudson et al., 2009). This is consistent with the findings that first through fourth grade students' listening comprehension contributed to variance in concurrent English text reading fluency beyond single word reading fluency (Kim and Wagner, 2015). Furthermore, bivariate correlations were small to moderate for listening comprehension and later word- and text reading fluency, although unique longitudinal associations were not examined. Further research concerning associations of listening comprehension with later fluency is needed.

The present study

The present study contributes to understanding aspects of fluency development in English reading by addressing three primary questions: (1) Do gains in word reading fluency over 6 months of second and third grade account for variance in word- and text-reading fluency one year later, after controlling for initial word reading fluency, phonological awareness, and rapid naming (as well as the two language skills)? It was hypothesized that gains in word reading fluency over this period would influence final outcomes, even beyond this stringent set of controls. (2) Do each of initial morphological awareness and listening comprehension account for unique variance in word- and text-reading fluency outcomes, over and above the control variables in this study? It was hypothesized that morphological awareness would account for variance in both fluency outcomes. Consistent with theoretical models (Kim, 2015), it was predicted that the effects of listening comprehension would be specific to text reading fluency. (3) Do each of the two oral language skills predict unique variance in final outcomes in text reading fluency beyond variance accounted for by final outcomes in word reading fluency? This question allows an examination of the effects of each oral language skill on text reading fluency that is not accounted for by an association with final outcomes in word reading fluency. It was predicted that listening comprehension would have a unique association with text reading fluency, but it is not clear if this would be the case for morphological awareness.

Methods

Participants

Second and third grade students were recruited as part of a longitudinal study on oral language and reading in English.

Eighty-three students completed a battery that included word- and text- reading fluency measures. Initial testing occurred when the students were in late fall of their second or third grade year (Initial time; Late fall-Year 1). After roughly six months, students again completed the word reading fluency measure (Spring-Year 1). Final fluency outcomes were measured roughly 18 months after study inception (Final time; Spring-Year 2; [Figure 1](#) shows the time line for measure administration).

Thirty-nine of these participants were in second grade at the beginning of the study (mean age 7;6 years; range 6;11–8;8; 20 females) and 44 in third grade at study onset (mean age of 8;4 years; range 7;10–8;10; 18 females). All but four students spoke English as their first language; these four students started to learn English as an additional language between 2 to 4 years of age and were included in all analyses.

The children were recruited from six public schools in suburban neighborhoods in an Eastern Canadian province. The neighborhoods were made-up largely of working- and middle-class families (district provided data showed the means of families served by the schools in the study did not differ from the district more generally concerning the prevalence of families with post-secondary education or categorized as falling within a low-income bracket; $M_s = 49\%$ and 7% for the district, respectively). The province and school district follow a balanced literacy curriculum ([Nova Scotia, 2019](#)).

Procedure

The study received ethical clearance from Mount Saint Vincent University's Research Ethics Board. All participants were included whose parents or legal guardians signed consent forms. The young participants gave assent to participate at the beginning of each session. All students were tested individually in quiet rooms within their schools.

Instruments

All standardized measures were administered following directions in each test manual.

Reliability coefficients reported for all standardized measures were above 0.80. The morphological awareness measure was an experimental measure, and is described below.

Phonological awareness: initial testing time

The Elision subtest of the Comprehensive Test of Phonological Processing - II ([Wagner et al., 2013](#)) required children to repeat a spoken word omitting the indicated sound(s) (e.g., "Say dog, now say it again without the /g/").

Rapid automatized naming: initial testing time

Naming speed was measured by the RAN Digit subtest of the Comprehensive Test of Phonological Processing - II ([Wagner et al., 2013](#)). Children are asked to name the stimuli as accurately and quickly as possible across the rows on a page. The manual reports raw scores as a function of number of errors and the time taken.

Word reading fluency: all three testing times

The Test of Word Reading Efficiency (TOWRE-2; [Torgesen et al., 2012](#)) required children to read as many words, in list format, as they could in 45 s. The raw score is number of words read correctly.

Gains in word reading fluency. Gains in word reading fluency across the 6 months of Year 1 were calculated by subtracting Time 1 raw scores on this subtest from T2 raw scores. Thus, these scores are the number of additional words each child read in 45 s at Time 2 than at Time 1.

Text reading fluency: final testing time

The Oral Reading Fluency subtest of the Woodcock Reading Mastery Tests - 3rd Edition ([Woodcock, 2011](#)) required children to read a series of short passages aloud. Their fluency score is based on the number of errors and the time taken for each passage.

Listening comprehension: initial testing time

Listening comprehension was measured with the Understanding Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals - 5th edition (CELF-5; [Wiig et al., 2013](#)). The examiner read aloud a series of paragraphs, and asked the student open ended questions after each paragraph.



FIGURE 1

Time line for administering measures across the 18 months of this study. PA, phonological awareness; RN, rapid naming; MA, morphological awareness; LC, listening comprehension; WRF, word reading fluency; TRF, text reading fluency.

Morphological awareness: initial testing time

A two-part, inflectional morphology task was used. The first twelve items required each child to provide the correct morphological form of a word that completed a spoken sentence. These items were from the Word Structure subtest of the Clinical Evaluation of Language Functioning–5 (Wiig et al., 2013). Five additional items required each child to correct mistakes in spoken sentences. All correct sentences required a manipulation of morphemes to correct the subject-verb agreement errors (e.g., *The cats plays and Todd sleep*; correct *The cat plays and Todd sleeps* or *The cats play and Todd sleeps*). These items stressed the meta-linguistic component of morphological awareness. Tasks with inflectional morphology were used to further tap into meta-cognitive awareness, as children this age have largely mastered implicit inflectional morphology (Robertson and Deacon, 2019). The Cronbach's alpha for this experimental task was 0.70.

Results

All variables in the study were normally distributed and there were no multivariate outliers (Tabachnick and Fidell, 2013). Raw scores are used in all analyses. Only one data point was missing, and this participant's data was not included in all analyses that included morphological awareness. Descriptive statistics are shown in Tables 1, 2 for all measures in this study. As can be seen from these Tables, the sample's standardized mean scores fell within the average range for all reading fluency measures and the additional reading-related skills, and the standard deviations appear similar to those on the standardized measures.

Bivariate correlations are shown in Table 3. Each of phonological awareness, rapid naming, morphological awareness, and listening comprehension are correlated with each measure of word- and text-reading fluency. The word reading fluency gain scores (based on the first six months of the study) are negatively correlated with initial word reading fluency. This indicates that those who had lower word reading fluency in the fall, made greater gains in this skill over the following 6 months. This is similar to previous findings for text reading fluency; students with lower fluency in the fall of grade 1 had greater growth across that academic year (Petscher and Kim, 2011).

A series of hierarchical regressions were used to answer the first two research questions. First, to test whether gains in word reading fluency and initial morphological awareness predicted unique

variance in final word reading fluency, a hierarchical regression was conducted with initial word reading fluency, phonological awareness and rapid naming entered in Step 1 (Table 4, Regression 1). This step accounted for 77% of the total variance, with initial word-reading fluency as the only significant predictor. Gains in word reading fluency accounted for an additional 6% of the variance as Step 2. Finally, morphological awareness predicted an additional, small but significant 1% of the variance in final word reading fluency.

The same hierarchical regression was conducted but with listening comprehension entered in Step 3 (Table 4, Regression 2). Listening comprehension was not a significant predictor of unique variance in final word reading fluency.

Similar analyses were used to examine whether morphological awareness, listening comprehension, and gains in word reading fluency predicted unique variance in final text-reading fluency. In the first regression, Step 1 accounted for 67% of the variance, with initial word reading fluency as the only significant predictor (Table 5, Regression 1). In Step 2, gains in word reading fluency accounted for an additional 7% of the total variance. In Step 3, morphological awareness predicted an additional 2% of the variance in text reading fluency. In the next hierarchical regression, listening comprehension comprised Step 3, predicting an additional 2% of the variance in final text reading fluency (Table 5, Regression 2).

The final regressions examined whether each of the two oral language skills would remain significant predictors of final text reading fluency, after accounting for final word reading fluency. First, final word reading fluency accounted for 77% of the total variance in text-reading fluency. When entered as Step 2, morphological awareness did not predict additional variance in text reading fluency (Table 5, Regression 3). In contrast, when entered as Step 2, listening comprehension accounted for an additional 2% of the variance in text reading fluency (Table 5, Regression 4).

Discussion

The purpose of this study was to examine unique contributions of English-speaking second and third grade students' initial morphological awareness and listening comprehension, as well as early gains in word reading fluency, to outcomes in word- and text-reading fluency. The sample of students in this study showed distributions on word- and text-reading fluency measures that

TABLE 1 Reading fluency measures: means and standard deviations.

Measure	Grade2 (n =39)		Grade3 (n =44)	
	Mean	SD	Mean	SD
Word Reading Fluency (T1-RS)	39.85	17.17	53.93	16.64
Word Reading Fluency (T1-SS)	97.23	16.49	97.41	17.28
Word Reading Fluency (T2-RS)	47.79	17.05	57.77	14.37
Word Reading Fluency (T2-SS)	98.03	17.90	95.50	16.21
Word Reading Fluency (T3-RS)	56.85	15.73	63.82	13.00
Word Reading Fluency (T3-SS)	95.13	18.48	94.66	15.35
Text Reading Fluency (T3-RS)	32.03	14.39	38.84	12.19
Text Reading Fluency (T3-SS)	99.18	15.89	99.27	13.94

RS, raw score; SS, standard score; T1, Time 1 (i.e., Initial Time); T2, Time 2 (i.e., 6 months into study); T3, Time 3 (i.e., Final time).

TABLE 2 Oral language and reading-related measures: means and standard deviations.

Initial-time measure	Grade2 (<i>n</i> = 39)		Grade3 (<i>n</i> = 44)	
	Mean	SD	Mean	SD
Phonological awareness (RS)	18.03	5.14	21.45	6.33
Phonological awareness (SS)	8.85	2.03	9.00	2.52
Rapid naming (RS)	23.62	5.08	20.34	4.47
Rapid naming (SS)	9.46	1.80	9.30	1.89
Listening comprehension (RS)	12.85	3.20	14.80	2.47
Listening comprehension (SS)	9.69	2.20	10.11	2.38
Morphological awareness (RS)	10.85	2.47	12.77	2.48

RS, raw score; SS, standard score. Mean for standardized tests = 10.

TABLE 3 Bivariate correlations among raw scores for study variables.

Measures		1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	Grade	—								
2.	PA	0.29	—							
3.	RAN	−0.33	−0.37	—						
4.	MA	0.37	0.47	−0.25*	—					
5.	List. Comp.	0.32	0.27	−0.18 ⁺	0.44	—				
6.	WRF (T 1)	0.39	0.63	−0.63	0.46	0.32	—			
7.	WRF (T 2)	0.31	0.59	−0.58	0.39	0.25*	0.93	—		
8.	ΔWRF	−0.31	−0.29	0.29	−0.31	−0.26*	−0.45	−0.09 ⁺	—	
9.	WRF (T 3)	0.24	0.53	−0.56	0.45	0.28	0.88	0.91	−0.16 ⁺	—
10.	TRF (T 3)	0.25	0.56	−0.47	0.47	0.37	0.82	0.86	−0.12 ⁺	0.88

All correlations significant at $p < 0.01$, unless otherwise indicated; * $p < 0.05$. ⁺ $p > 0.10$; PA, phonological awareness; RAN, rapid naming; MA, morphological awareness; List. Comp., listening comprehension; WRF, word reading fluency; TRT, text reading fluency; Δ, change; T1, Time 1/Initial; T2, Time 2/6 months into study; T3, Time 3/Final.

appeared similar to the normative, standardization samples (see Table 1). The study controlled for initial word reading fluency, phonological awareness, and rapid naming. Findings highlight the importance of continuing gains in word reading fluency over 6 months of second and third grade to fluency outcomes near the end of their third and fourth grade year. Initial word reading fluency did, however, account for the bulk of the variance in later fluency outcomes. Students' initial morphological awareness also predicted outcomes in both word- and text-reading fluency 18 months later; however, the unique contribution of morphological to text reading fluency was accounted for by its association with final word reading fluency. On the other hand, there was a consistent unique effect of listening comprehension on text reading fluency. Each finding in this study is discussed in turn.

The contributions of earlier gains in word reading fluency to fluency outcomes

The findings in this study, coupled with previous research, help to delineate the developmental time course of individual differences in English word reading fluency. For students with reading impairments beyond first grade, it has proven difficult to move the needle on standardized measures of reading fluency, while effective interventions prior to the end of first grade tend to normalize fluency outcomes

(Torgesen et al., 2001). These findings suggest that trajectories for English reading-fluency development may be solidified quite early in this process of learning to read. The current study lends support to the notion that fluency skills are determined quite early in the reading process. By the late fall of second and third grade, word reading fluency accounted for the bulk of the variance in word- and text-reading fluency outcomes 18 months later (77 and 68%, respectively). At the same time, additional, relatively sizable variance in both fluency outcomes were accounted for by gains in word reading fluency over six months of second and third grade (6–7%). Importantly, students who started off lower in word reading fluency made greater gains on this skill than those who were higher at study onset (similar to past findings for text reading fluency; Kim et al., 2010). This means that relative standing in word reading fluency remains at least somewhat malleable across the second and third grade years.

Neither phonological awareness nor rapid naming predicted unique variance in word- or text-reading fluency outcomes for this relatively opaque orthography. Conceivably, there may be a direct role for these skills on word reading fluency in participants younger than those in the current study; but, the roles of these two skills may be developmentally limited in English reading. Findings do not support the notion that phonemic awareness proficiency for more advanced phonological awareness tasks such as manipulation and deletion (the later measured in the current study), has a unique role or is critical to ongoing growth in fluency across these elementary

TABLE 4 Hierarchical regression analyses predicting outcomes in final word reading fluency.

Step	Predictor	ΔR^2	β	Final β
Regression 1				
1	Time 1 WRF	0.77	0.88**	0.97**
	PA		−0.04	−0.07
	RAN		−0.04	−0.04
2	Change WRF	0.06	0.28**	0.29**
3	Time 3 Morph. Awareness	0.01		0.11*
Regression 2				
1	Time 1 WRF	0.77	0.88**	0.99**
	PA		−0.04	−0.04
	RAN		−0.04	−0.03
2	Change SWRF	0.06	0.28**	0.29**
3	Time 3 List. Comp	0.00		0.05

* $p < 0.05$; ** $p \leq 0.01$; WRF, Word reading fluency; PA, phonological awareness; RAN, Rapid Naming; List. Comp., Listening Comprehension; T1, Time 1/Initial; T3, Time 3/Final.

TABLE 5 Hierarchical regression analyses predicting outcomes in final text reading fluency.

Step	Predictor	ΔR^2	β	Final β
Regression 1				
1	T1 WRF	0.67	0.80**	0.90**
	PA		0.07	0.03
	RAN		0.05	0.05
2	Change WRF	0.07	0.30**	0.32**
3	Time 1 Morph Awareness	0.02		0.16*
Regression 2				
1	T1 WRF	0.67	0.80**	0.91**
	PA		0.07	0.06
	RAN		0.05	0.05
2	Change WRF	0.07	0.30**	0.33**
3	T1 Listening Comp.	0.02		0.16*
Regression 3				
1	Time 3 WRF	0.77	0.88**	0.84**
2	Time 1 Morph Awareness	0.01		0.10
Regression 4				
1	T3 WRF	0.77	0.88**	0.84**
2	T1 Listening Comp.	0.02		0.13*

* $p < 0.05$; ** $p \leq 0.01$; WRF, word reading fluency; PA, phonological awareness; RAN, rapid naming; T1, Time 1/Initial; T3, Time 3/Final.

grades (Kilpatrick, 2020). The related recommendation for a prolonged focus on oral-only phonemic awareness instruction in order to set a solid trajectory for fluency (Heggerty, 2003; Kilpatrick, 2015) conflicts with recommendations from the National Reading

Panel (2000) (see Brady, 2020, for updated review and discussion) and are not suggested by the correlational findings in this study. Rather, the current study suggests a more direct focus on increasing efficiency of word reading skills in second and third grade may increase later reading fluency outcomes. This may not have been a target of instruction within the context of the balanced literacy curriculum for students in the current study; that is, balanced literacy tends not to place a significant focus on instruction for efficient, context-free recognition of words (Spear-Swerling, 2019).

The contribution of morphological awareness to word- and text-reading fluency

Multicomponent models of reading comprehension have outlined direct roles for foundational oral language knowledge on comprehension processes (e.g., Stafura and Perfetti, 2017). Models of reading fluency have also proposed a role for morphological knowledge directly to fluency (Hudson et al., 2009, 2012; Kim, 2015); fluency in turn, also contributes to reading comprehension (Kim et al., 2021). In this study, students' morphological awareness at the beginning of second and third grade was a predictor of gains in word reading fluency 18 months later. This finding builds on some past research by controlling for phonological awareness, rapid naming, and the autoregressive variable. Theoretical accounts propose several mechanisms through which morphological knowledge may influence word reading fluency. In Ehri's Phase Theory, efficient word reading comes to rely on increasingly larger chunks of internalized orthographic patterns, including morphemes (Ehri, 2005). This association seems congruent with Hudson et al.'s (2009) model, which proposes this higher-level orthographic knowledge as one influence on reading fluency (see also, Wolf and Katzir-Cohen, 2001). Levesque et al. (2021) propose an additional mechanism; children actively use their knowledge of morphemes to analyze and decode words. In turn, morphemes as meaningful units may trigger "top-down semantic activation enabling faster and more accurate word reading" (Levesque et al., 2021, p. 14). In support, Nunes et al. (2012) found that the degree to which children used morphological units in decoding and spelling predicted later outcomes in both word- and text-reading fluency. Morphological awareness thus appears to play a role in developing word reading fluency, and this may become more critical with the increase of morphologically complex words in texts with rising grades.

Morphological awareness was also found to predict unique variance in later text reading fluency, beyond initial status and early gains in word reading fluency, phonological awareness and rapid naming. This adds to a relatively small body of research examining this association from a longitudinal perspective. At first glance, it appears that morphological awareness, perhaps as one component of the top-down linguistic processes, facilitates text reading fluency. However, after controlling for final outcomes in word reading fluency, morphological awareness no longer predicted unique variance in text reading fluency. Kirby et al. (2012) found a unique contribution from grade 2 morphological awareness to grade 3 text reading fluency; however, initial or final status in word reading fluency were not included in their model. Kim (2015) found that morphological awareness did not predict later Korean reading fluency when word

reading fluency was in the model, albeit Korean is a more transparent orthography. Although the current findings suggest that the association between morphological awareness and text reading fluency are fully accounted for by its influence on gains in word reading fluency – this in no way diminishes its importance to text reading fluency, given the magnitude of the association between final outcomes in word- and text-reading fluency ($r=0.88$). The unique role found in this study for morphological awareness on word reading fluency appears consistent with this aspect of the Morphological Pathway's model (Levesque et al., 2021).

The contributions of listening comprehension to word- and text-reading fluency

Understanding oral language is an important component of reading comprehension (Hoover and Gough, 1990). This study contributes to the limited body of research on a potential role for listening comprehension in children's fluency development. Children's listening comprehension predicted unique variance in later text reading fluency, but not word reading fluency, beyond that accounted for by measures of earlier word reading fluency, phonological awareness, and rapid naming. Furthermore, this unique contribution remained significant when final outcomes in word reading fluency were entered into the equation. Previous research addressing this relationship has been somewhat sparse, but the findings from this study are consistent with multi-component models of fluency development (e.g., Hudson et al., 2009, 2012). For one such model, it is proposed that elements associated with reading (and listening) comprehension (e.g., knowledge, vocabulary, passage- and social-context) have a top-down influence on text reading fluency (Hudson et al., 2009). Their model does not incorporate direct or indirect effects of these comprehension processes on word reading fluency, consistent with the specific association found in the current study. Meaning making processes involved in oral language comprehension thus appear to facilitate text reading fluency, even over these extended years of elementary school. While the Simple View of Reading proposes a role for linguistic comprehension to reading comprehension (Hoover and Gough, 1990), the current findings are consistent with somewhat more complex models which support its role in text reading fluency (Kim, 2015). That is, listening comprehension may directly influence comprehension as a facet of semantic linguistic processes (Stafura and Perfetti, 2017), and may have an indirect effect on comprehension through its potential contribution to reading fluency. This proposal is also consistent with Kim and Wagner's (2015) findings that the relationship from listening comprehension to reading comprehension in English, was partially mediated by text reading fluency for second through fourth grade students.

Limitations

The current study adds to the growing research examining the time course of fluency development and the role of two dimensions of oral language. This study has limitations which must be considered

when interpreting the findings. First, the measure of morphological awareness included inflectional but not derivational morphology. This was judged appropriate as inflectional knowledge is more stable for the developmental period examined, and thus individual differences may more heavily tap meta-linguistic knowledge (see also Robertson and Deacon, 2019). However, derivational knowledge has been measured in this age range and further research is needed to test the generalizability of the current findings for this additional aspect of morphology. Second, the current study defined text reading fluency in terms of accuracy and speed (see also Kim, 2015; Kim et al., 2021). Other researchers include prosody or expression as a component of text reading fluency. The unique contributions of oral language measures may vary depending on how text reading fluency is operationalized. The current study lends support for a unique contribution of listening comprehension to text reading fluency, and this association might be expected to be stronger when prosodic elements are included in oral reading fluency measures. Third, the sample size was relatively small and thus somewhat limited the data analytic approach. Finally, the study was conducted within the context of a balanced literacy instructional approach (for discussion, see Spear-Swerling, 2019). Future research is needed to test the proposed direct and indirect contributions of these oral language skills to reading fluency outcomes and to test the generalizability to instructional contexts that may include more direct focus on word reading accuracy and efficiency.

Conclusion

Longitudinal research findings on the associations examined in this study have been inconsistent or sparse. The current findings contribute to a better understanding of the developmental time course of fluency skills in these early elementary years and of the contributions of oral language skills to reading fluency. Individual differences in word reading fluency appear quite stable from early in second and third grade. This has implications for thinking about early word reading instruction; approaches that focus on both accuracy and efficiency for components of word reading instruction from the earliest grades may be critical (Lane and Contessa, 2022). At the same time, the significant gains across six months of second and third grade influenced fluency outcomes near the end of the students' third- and fourth-grade years, and gains were largest for those with lower initial scores. Continued targeted instruction in word reading fluency across this period may help change trajectories for those off to a slower start. Findings also suggested that phonological awareness and rapid naming may have a developmentally limited role to play in reading fluency and the findings do not support an ongoing focus across these grades on oral-only phonemic awareness instruction. Finally, consistent with Kim's (2015) proposal, findings supported specificity in the unique associations of reading-related oral language skills with fluency. Morphological awareness appears to contribute primarily to gains in word reading fluency, consistent with mechanisms outlined in some models (Levesque et al., 2021). In turn, word reading fluency is the major determinant of text reading fluency. Conversely, listening comprehension was uniquely associated to outcomes in text reading fluency, supporting a potential role for top-down linguistic processes in text reading fluency (Hudson et al., 2009, 2012).

Data availability statement

The datasets presented in this article are not readily available because the consent forms signed for this study preclude making the data for this manuscript openly available. Requests to access the datasets should be directed to JM: Jamie.Metsala@msvu.ca.

Ethics statement

The studies involving human participants were reviewed and approved by University Research Ethics Board, Mount Saint Vincent University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This research was funded by an Inter-University Research Network Grant from the Nova Scotia Department of Education and

Early Childhood Development and by funding from the Gail and Stephen Jarislowsky Research Chair in Learning Disabilities Endowment at Mount Saint Vincent University.

Acknowledgments

The author wishes to thank participating students, their teachers and principals, as well as university researchers and the team who worked with the students, for their contributions to this research.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Apel, K., Diehm, E., and Apel, L. (2013). Using multiple measures of morphological awareness to assess its relation to reading. *Top. Lang. Disord.* 33, 42–56. doi: 10.1097/TLD.0b013e318280f57b
- Brady, S. (2020). A 2020 perspective on research findings on alphabets (phoneme awareness and phonics): Implications for instruction (expanded version). *Read. League J.* 1, 20–28.
- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing* 12, 169–190. doi: 10.1023/A:1008131926604
- Deacon, S. H., and Kieffer, M. (2018). Understanding how syntactic awareness contributes to reading comprehension: Evidence from mediation and longitudinal models. *J. Educ. Psychol.* 110, 72–86. doi: 10.1037/edu0000198
- Desrochers, A., Manolitsis, G., Gaudreau, P., and Georgiou, G. (2018). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Read. Writ.* 31, 1695–1719. doi: 10.1007/s11145-017-9772-y
- Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., and Protopapas, A. (2017). Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Front. Psychol.* 8, 20–39. doi: 10.3389/fpsyg.2017.02039
- Ehri, L. C. (1998). "Grapheme-phoneme knowledge is essential for learning to read words in English" in *Word recognition in beginning literacy*. eds. J. L. Metsala and L. C. Ehri (Mahwah, NJ: Lawrence Erlbaum), 3–40.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902_4
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21. doi: 10.1080/10888438.2013.819356
- Georgiou, G. K., Parrila, R., and Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *J. Educ. Psychol.* 100, 566. doi: 10.1037/0022-0663.100.3.566
- Giazitidou, S., Mouzaki, A., and Padelidu, S. (2023). Pathways from morphological awareness to reading fluency: the mediating role of phonological awareness and vocabulary. *Read. Writ.*, 1–23. doi: 10.1007/s11145-023-10426-2
- Giazitidou, S., and Padelidu, S. (2022). Contribution of morphological awareness to reading fluency of children with and without dyslexia: evidence from a transparent orthography. *Ann. Dyslexia* 72, 509–531. doi: 10.1007/s11881-022-00267-z
- Hasbrouck, J., and Tindal, G. (2017). An update to compiled ORF norms (No. 1702). Technical report. Available at: https://intensiveintervention.org/sites/default/files/TechRpt_1702ORFNorms%20FINAL.pdf (Accessed March 26, 2023).
- Heggerty, M. (2003). *Phonemic awareness: The skills that they need to help them succeed!* Oak Park, IL: Literacy Resources Inc.
- Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160. doi: 10.1007/BF00401799
- Hudson, R. F., Pullen, P. C., Lane, H. B., and Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Read. Writ. Quar. Overc. Learn. Difficult.* 25, 4–32. doi: 10.1080/10573560802491208
- Hudson, R. F., Torgesen, J. K., Lane, H. B., and Turner, S. J. (2012). Relations among reading skills and sub-skills and text-level reading proficiency in developing readers. *Read. Writ. Interdiscip. J.* 25, 483–507. doi: 10.1007/s11145-010-9283-6
- Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., and Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *J. Educ. Psychol.* 95, 719–729. doi: 10.1037/0022-0663.95.4.719
- Juel, C. (1988). Learning to read and write: a longitudinal study of 54 children from first through fourth grades. *J. Educ. Psychol.* 80, 437–447. doi: 10.1037/0022-0663.80.4.437
- Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. Hoboken, NJ: John Wiley & Sons.
- Kilpatrick, D. A. (2020). How the phonology of speech is foundational for instant word recognition. *Persp. Lang. Liter.* 46, 11–15.
- Kim, Y. S. G. (2015). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Read. Res. Q.* 50, 459–481. doi: 10.1002/rtrq.107
- Kim, Y. S. G., Petscher, Y., Schatschneider, C., and Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *J. Educ. Psychol.* 102, 652–667. doi: 10.1037/a0019643
- Kim, Y. S. G., Quinn, J. M., and Petscher, Y. (2021). What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation. *Dev. Psychol.* 57, 718–732. doi: 10.1037/dev0001167
- Kim, Y. S. G., and Wagner, R. K. (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. *Sci. Stud. Read.* 19, 224–242. doi: 10.1080/10888438.2015.1007375

- Kirby, R. J., and Bowers, P. N. (2017). "Morphological instruction and literacy: binding phonological, orthographic, and semantic features of words" in *Theories of reading development*. eds. K. Cain, D. L. Compton and R. K. Parrila (Amsterdam: John Benjamins), 437–461.
- Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., and Parrila, R. (2012). Children's morphological awareness and reading ability. *Read. Writ.* 25, 389–410. doi: 10.1007/s11145-010-9276-5
- Klauda, S. L., and Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *J. Educ. Psychol.* 100, 310–321. doi: 10.1037/0022-0663.100.2.310
- LaBerge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/00100285(74)90015-2
- Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., et al. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Sci. Stud. Read.* 23, 220–234. doi: 10.1080/10888438.2018.1510936
- Lane, H., and Contessa, V. (2022). *UFLI foundations: an explicit and systematic phonics program*. Sun Prairie, WI: Ventris Learning.
- Lee, J. W., Wolters, A., and Grace Kim, Y. S. (2022). The Relations of Morphological Awareness with Language and Literacy Skills Vary Depending on Orthographic Depth and Nature of Morphological Awareness. *Rev. Educ. Res.*, 1–31. doi: 10.3102/00346543221123816
- Levesque, K. C., Breadmore, H. L., and Deacon, S. H. (2021). How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development. *J. Res. Read.* 44, 10–26. doi: 10.1111/1467-9817.12313
- Manolitsis, G., Georgiou, G. K., Inoue, T., and Parrila, R. (2019). Are morphological awareness and literacy skills reciprocally related? Evidence from a cross-linguistic study. *J. Educ. Psychol.* 111, 1362–1381. doi: 10.1037/edu0000354
- Manolitsis, G., Grigorakis, I., and Georgiou, G. K. (2017). The longitudinal contribution of early morphological awareness skills to reading fluency and comprehension in Greek. *Front. Psychol.* 8, 1–14. doi: 10.3389/fpsyg.2017.01793
- McBride-Chang, C., Cho, J. R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., et al. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *J. Exp. Child Psychol.* 92, 140–160. doi: 10.1016/j.jecp.2005.03.009
- Metsala, J. L., and David, M. D. (2017). The effects of age and sublexical automaticity on reading outcomes for students with reading disabilities. *J. Res. Read.* 40, S209–S227. doi: 10.1111/1467-9817.12097
- Metsala, J. L., and David, M. D. (2022). Improving English reading fluency and comprehension for children with reading fluency disabilities. *Dyslexia* 28, 79–96. doi: 10.1002/dys.1695
- Metsala, J. L., Sparks, E., David, M., Conrad, N., and Deacon, S. H. (2021). What is the best way to characterise the contributions of oral language to reading comprehension: listening comprehension or individual oral language skills? *J. Res. Read.* 44, 675–694. doi: 10.1111/1467-9817.12362
- National Reading Panel (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development, National Institutes of Health.
- Norton, E. S., and Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annu. Rev. Psychol.* 63, 427–452. doi: 10.1146/annurev-psych.120710-100431
- Nova Scotia (2019). English Language Arts P-6: At a Glance. Available at: English Language Arts P-6 at a glance (2019).pdf (novascotia.ca)
- Nunes, T., Bryant, P., and Barros, R. (2012). The development of word recognition and its significance for comprehension and fluency. *J. Educ. Psychol.* 104, 959–973. doi: 10.1037/a00
- Papadopoulos, T. C., Spanoudis, G. C., and Georgiou, G. K. (2016). How is RAN related to reading fluency? A comprehensive examination of the prominent theoretical accounts. *Front. Psychol.* 7:1217. doi: 10.3389/fpsyg.2016.01217
- Perfetti, C. A. (1988). "Verbal efficiency in reading ability" in *Reading research: advances in theory and practice*. eds. M. Daneman, G. E. Mackinnon and G. T. Waller, vol. 6 (Cambridge, MA: Academic Press), 109–143.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383. doi: 10.1080/10888430701530730
- Perfetti, C., and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Sci. Stud. Read.* 18, 22–37. doi: 10.1080/10888438.2013.827687
- Petscher, Y., and Kim, Y. S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *J. Sch. Psychol.* 49, 107–129. doi: 10.1016/j.jsp.2010.09.004
- Robertson, E. K., and Deacon, S. H. (2019). Morphological awareness and word-level reading in early and middle elementary school years. *Appl. Psycholinguist.* 40, 1051–1071. doi: 10.1017/S0142716419000134
- Seidenberg, M. S., and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96, 523–568. doi: 10.1037/0033-295X.96.4.523
- Share, D. L. (2008). "Orthographic learning, phonological recoding, and self-teaching" in *Advances in child development and behavior*. ed. R. V. Kail, vol. 36 (Amsterdam: Elsevier Academic Press), 31–82.
- Shechter, A., Lipka, O., and Katzir, T. (2018). Predictive models of word reading fluency in Hebrew. *Front. Psychol.* 9:1882. doi: 10.3389/fpsyg.2018.01882
- Spear-Swerling, L. (2019). Structured literacy and typical literacy practices: Understanding differences to create instructional opportunities. *Teach. Except. Child.* 51, 201–211. doi: 10.1177/0040059917750
- Stafura, J. Z., and Perfetti, C. A. (2017). "Integrating word processing with text comprehension" in *Theories of reading development*. eds. K. Cain, R. K. Parrila and D. L. Compton (John Benjamins), 9–31.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Tabachnick, B., and Fidell, L. (2013). *Using Multivariate Statistics, 6th Edn* London: Pearson.
- Torgesen, J. K., and Hudson, R. (2006). "Reading fluency: Critical issues for struggling readers" in *Reading fluency: The forgotten dimension of reading success*. eds. S. J. Samuels and A. Farstrup (Newark: International Reading Association), 130–158.
- Torgesen, J. K., Rashotte, C. A., and Alexander, A. (2001). "Principles of fluency instruction in reading: Relationships with established empirical outcomes" in *Dyslexia, fluency, and the brain*. ed. M. Wolf (London: York Press), 333–356.
- Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (2012). *Test of word reading efficiency (TOWRE-2). 2nd Edn*. Austin, TX: Pro-Ed.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., and Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing (CTOPP-2) 2nd Edn*. Austin, TX: Pro-Ed.
- Wiig, E. H., Semel, E., and Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals (CELF-5) 5th Edn*. London: NCS Pearson.
- Wolf, M., and Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *J. Educ. Psychol.* 91, 415–438. doi: 10.1037/0022-0663.91.3.415
- Wolf, M., and Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Sci. Stud. Read.* 5, 211–239. doi: 10.1207/S1532799XSSR0503_2
- Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests (WRMT-3) 3rd Edn*, London: Pearson.
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Fásca, L., et al. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychol. Sci.* 21, 551–559. doi: 10.1177/0956797610363406



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga, Spain

REVIEWED BY

Antonella Marchetti,
Catholic University of the Sacred Heart, Italy
Maria Carmen Usai,
University of Genoa, Italy

*CORRESPONDENCE

Giulia Vettori
✉ giulia.vettori@unifi.it

RECEIVED 24 March 2023

ACCEPTED 30 May 2023

PUBLISHED 16 June 2023

CITATION

Bigozzi L, Incognito O, Mercugliano A,
De Bernart D, Botarelli L and Vettori G (2023) A
study on the emergence of sound-sign
correspondence in Italian-speaking
5-year-old pre-schoolers.
Front. Educ. 8:1193382.
doi: 10.3389/feduc.2023.1193382

COPYRIGHT

© 2023 Bigozzi, Incognito, Mercugliano, De
Bernart, Botarelli and Vettori. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

A study on the emergence of sound-sign correspondence in Italian-speaking 5-year-old pre-schoolers

Lucia Bigozzi, Oriana Incognito, Alice Mercugliano,
Diletta De Bernart, Lucia Botarelli and Giulia Vettori*

Department of Education, Languages, Interculture, Literatures and Psychology, University of Florence, Florence, Italy

This study investigates the emergence of sound-sign correspondence in Italian-speaking 5-year-old pre-schoolers. There are few experimental studies on the precursors of reading and writing skills and those existing mainly focus on letter knowledge or logographic processing of words in pre-schoolers. This paper evaluates and compares 5-year-old children's use of the logographic processing or the use of sound-sign processing to decode target words in original and modified versions. Furthermore, we verify whether pre-schoolers' type of reading words (logographic versus sound-sign processing) vary in accordance with children's socio-cultural differences (i.e., type of school and socio-cultural information from parents). This study tested 94 children (M-age = 5 years and 8 months) at the end of the last year of preschool. Six stimulus logos were used to evaluate children's ability to decode words and the type of decoding (logographic or sound-sign processing). The Chi-square results confirm that the achievement of the correspondence between sound-sign at the base of reading and writing has already started in preschool. Our findings shed light on a significant proportion of pre-schoolers who can already read words via sound-sign processing or show the emergence of notational awareness, while the others still rely on logographic processing. Moreover, the results show that pre-schoolers' notational awareness is related to socio-cultural characteristics pertaining to schools and families. These findings suggest that 5 years is an important age for the disentanglement between logographic and sound-sign correspondence in pre-schoolers and provide useful implications for theory and practice.

KEYWORDS

literacy acquisition, sound-sign correspondence, emergent literacy, pre-schoolers, notational awareness

1. Introduction

In studying the development of children's learning to read and write, most of the developmental theories on literacy acquisition refer to the importance of the so-called emergent literacy period (Pinto et al., 2012, 2017; Bigozzi et al., 2016a,b) which significantly supports later formalised learning to read and write. The emergent literacy period is identified as a preschool phase during which children are exposed to different forms of written language although this is not explicitly taught to them. Frith's (1986) hypothesis states that initially children have access to the meaning of a written word using its visual

form, without any phonological mediation, in what is called the 'logographic phase'. Thus, at the very beginning of their contact with written language, most children seem to be able to read a small number of words, those belonging to their everyday life context (Frith, 1986). Frith's theorisation has produced a conspicuous number of research studies to clarify several issues, for example the effective utility of the logographic phase for literacy acquisition and the existence of a logographic phase in all children (Bastien-Toniazzi and Jullien, 2001), the existence of language-specific effects (Aro and Wimmer, 2003; Spencer and Hanley, 2004; Marinelli et al., 2015), also given the great heterogeneity among subjects of the same age (Rayner et al., 2012).

A further line of research points out pre-schoolers' sensitivity to extra-linguistic environments as central factors in achieving the meaning of words (Mol and Bus, 2011). Other studies suggest that children store an overall shape in their memory (Seymour and Elder, 1986) or just certain letters (Ellis and Large, 1988; Kuhn et al., 2010) or even salient graphic features of the word's shape, such as the presence of ascenders and descenders (Ehri, 2020), that are used as a cue to recognise written words.

In addition to these different perspectives on pre-schoolers' sensitivity to written language features, Harris and Coltheart (1986) theorised that even in the first phase of literacy acquisition, which the authors call the visual vocabulary phase and place it when a child is four years old, children are able to read a small number of words aloud using a direct procedure (Harris and Coltheart, 1986).

In particular, the authors focus their attention on a single case of a four-year-old girl, Alice, (in the midst of a visual vocabulary stage) who was able to read about 30 words, some of which she had been taught while others were learned spontaneously. To find out whether this reading ability was based on a direct procedure or on visual form recognition, one of the words the child could read was used. The word "Harrods" was presented to her in a written form different from what she was used to (word in capital letters without logos). Alice was able to correctly read the word presented in an unusual graphic form, showing that she was using a directed reading process. This procedure is not related to grapheme-phoneme correspondence: words are recognised as specific sequences of letters and not in their global form. The child was also subjected to a second experiment: Alice was presented with a list of words similar to a word she already knew, (also on the list), but modified in the initial or final part, and she was able to correctly identify the known word. Here again, Alice used all the letters in the sequence to read the word and distinguish it from the others presented in the list (Harris and Coltheart, 1986). We also find this in the study by Masonheimer et al. (1984) where children, despite the alphabetical features of words being altered, read the same words through a correspondence between visual aspects and pronunciation with which, on the basis of their prior experience, they usually associate them. However, they do not consider the changes made and thus commit errors in reading the words XEPSI and PEPSI. Even though the letter P is replaced with an X, the children pay attention to the graphic aspects and continue to associate them with the form corresponding to the word PEPSI with which they have always been used to associating it, regardless of what is actually represented. In this example, we can also detect the effect of the law of the "good gestalt": when children are in the condition to read a new piece of writing, they are inclined to point to what is the 'form' and therefore the most common,

most balanced, simplest word that corresponds to it on the basis of its graphical aspects, without thinking of the possible presence of variations that would make it less homogeneous (Kanizsa, 1983). Thus on the basis of its visual characteristics, it is more harmonious to read the word XEPSI as PEPSI rather than in the actual corresponding form.

On the other hand, a part of the studies showed that preschool children already possess knowledge of the nature and conceptual meaning of a writing system at an early stage of development: drawing, numeracy, and literacy are all key components of emergent understanding of symbolic systems (Whitehurst and Lonigan, 1998; Rohde, 2015). All of these aspects influence the formal learning of conventional literacy processes (Hand et al., 2022) and can, therefore, be considered strong predictors of them (Missall et al., 2007; Bigozzi et al., 2016b).

Children already at preschool age possess the ability to differentiate between different symbolic systems (e.g., letters from numbers) and to use different knowledge and attitudes specific to each domain (Yamagata, 2007). Studies have shown that possession of these early skills leads to improved reading achievement in elementary school (Bishop, 2003). All these symbolic systems, therefore, allow the expression of mental representations (drawing and writing leave visible marks, unlike speaking or reading) (Hall et al., 2015) but differ in some aspects. Drawing can be described as a process characterised by recurring graphic patterns and certain rules that must be followed (Nicholls and Kennedy, 1992). Writing, on the other hand, is characterised by greater restrictions, as the set of units represents a closed system in which nothing can be added without drastically changing the meaning (Zhang and Treiman, 2021).

Pre-schoolers' sensitivity to the characteristics of signs and letters is well-documented in the literature (Neumann et al., 2012). Preschool children, in order to best develop conceptual knowledge of their writing system, reflect upon the different symbolic systems used to represent meanings (e.g., written language, numerical language, and drawing), make spontaneous attempts to represent words in print (Milburn et al., 2017; Ouellette and Sénéchal, 2017) and try to systematically match the sounds included in words with signs that are not necessarily conventional (which should be distinguished from the letters of the alphabet) (Treiman and Kessler, 2013; Puranik et al., 2014). In particular, pre-schoolers' notational awareness is a key component of the Emergent Literacy Model by Pinto et al. (2009, 2018; Pinto and Incognito, 2022) for Italian children. Notational awareness indicates pre-schoolers' ability to understand and operate the correspondence between sound-sign. Invented reading and invented spelling activities demonstrate the emergence and mastering of letter-speech sound integrations by young children before schooling. This complex skill, typical of preschool children, integrates phonological awareness (Ouellette and Sénéchal, 2017; Albuquerque and Alves Martins, 2022) with grapho-motor skills (Mohamed and O'Brien, 2022) and visual attention (Valdois et al., 2019). For example, "invented spelling" a developmental stage in which children attempt to merge the phonological and orthographic features of a word through the use of "invented" signs, written productions that, while not yet letters, include some of the properties of the writing system (Albuquerque and Alves Martins, 2022). An analysis of the literature suggests that, at later developmental ages, reading fluency (Bigozzi et al., 2016b) and the presence of reading and

spelling disorders (Bigozzi et al., 2016a) are significantly affected by children's early notational awareness. Previous studies show that socio-cultural factors affect children's literacy skills. The study by Incognito and Pinto (2021) demonstrates that growing up in a rich environment in terms of literacy materials and experiences contributes to pre-schoolers' emergent literacy skills development. Further research evidence shows the influence of home literacy on other important skills, such as word recognition (Evans and Shaw, 2008) and that children whose parents have higher levels of education and occupation interact with more literacy resources and receive more informal literacy activities, such as shared reading (Wang and Liu, 2021).

1.1. Aims and hypothesis

The aim of this study was to investigate the emergence and development of sound-sign correspondence in pre-schoolers.

Specifically, the aims were the following:

- 1) To evaluate and compare 5-year-old children's use of logographic processing or the use of sign-sound processing to decode target logo words and whether there are differences between words written in Italian (transparent orthography) and words written in English (opaque orthography);
- 2) To verify whether pre-schoolers' type of reading words (logographic versus sign-sound processing) vary in accordance with children's socio-cultural differences (i.e., type of school and socio-cultural information from parents).

Regarding the first aim, we expect that 5 years is an important age for the disentanglement between logographic and sound-sign correspondence in pre-schoolers. In 5-year-old children referable to the logographic period, we expect to observe the significant emergence of notational awareness in pre-schoolers when using the sound-sign processing in decoding logo words.

Regarding the second aim, we hypothesised that pre-schoolers' performances vary in accordance with children's socio-cultural differences, linked to the type of preschool and socio-cultural information from parents.

2. Materials and methods

2.1. Participants

The research involved 94 children, 49 boys and 45 girls (M-age = 5 years and 8 months), all attending three different preschools located in a city of central Italy. The schools involved in the research are three municipal preschools in the municipality of Arezzo and located in different areas of the city: School 1 (25 pre-schoolers – 14 males, 11 females), School 2 (32 pre-schoolers – 17 males, 15 females), School 3 (37 pre-schoolers – 18 males, 19 females). The children in School 1 had grown up in a more culturally elevated environment than the children in the other schools. Table 1 shows sample distribution across schools. All the children showed typical development, were born in Italy, and spoke Italian as their mother tongue. At the time of the study, no participant had a diagnosis of

TABLE 1 Percentages of the educational qualifications of the children's parents who participated in the research.

Parents' educational level	School 1	School 2	School 3
Degree	78%	59%	13%
High school diploma	22%	23%	42%
Middle school diploma	0	18%	45%

School 1: $N = 25$ children (14 males, 11 females); School 2: $N = 32$ children (17 males, 15 females); School 3: $N = 37$ children (18 males, 19 females).

physical or mental disability or had begun a diagnostic evaluation process or, according to teachers, presented a special educational need. The three preschools have a catchment area where the parents come from different socio-economic backgrounds. The association between children's responses and the type of preschool and socio-cultural information from parents were analysed. We checked that all schools included in this study adhere to the standard national curriculum in accordance with the Italian national guidelines issued by the Ministry of Education. Children attending the last year of the preschool class were exposed to different forms of written language (words, posters, signs and books, etc.) but not explicitly taught the grapho-phonological code. Data collection in classrooms took place at the end of the school year during the month of May.

2.2. Procedure and measures

Three administrations were performed within one-week of time. In the first administration, at the first meeting with the children, the work was presented collectively, explaining to them that this work was not used to judge them but to help the teachers who they would meet at primary school to decide on the best way for them to learn to read and write; it was also explained to them that such work had to be done individually in a quiet place beside the classroom. Following the general illustration phase, each child was presented individually with the target word and asked the question: "Could you tell me what is written on this card?"

The task material consists of six stimulus words chosen from a larger list of words (including toy brands and cartoon characters after having checked the copyrights) through a recognition task performed in a pilot study on 23 children (13 females and 10 males) not included in the final sample used in this research. We subjected 23 children to 14 target logos of the following 14 words: Walt Disney, Nutella, Kinder, Fiesta, Playmobil, Barbie, Action Man, Nemo, Madagascar, Shrek, Estathe, Gormiti, Winx. Those stimuli belong to specific categories such as snacks, names of toys or cartoons, easily recognisable by the children as every-day and familiar stimuli. From these stimuli presented to the pilot sample, we chose the most recognised and familiar ones, therefore the six words used in the research were the following logos:













1.  indicating one of the main characters of the computer-cartoon called "Finding Nemo" and "Finding Dory";
2.  indicating the computer-animated survival comedy film;
3.  indicating the computer-animated comedy film;


TABLE 2 Number of children who read the words correctly in their original form and their percentages.


Original version of target word	Total number of children who can read correctly	% Correct reading in School 1	% Correct reading in School 2	% Correct reading in School 3
	83 (88%)	92	93	81
	85 (90%)	92	90	89
	93 (99%)	100	100	97
	86 (91%)	88	93	91
	94 (100%)	100	100	100
	94 (100%)	100	100	100
Total % of children who can read correctly for each school		92%	94%	81%

The last three columns show the percentages of children who read correctly for each school.

4.  indicating a brand of tea;
5.  indicating a toy brand;
6.  indicating an animated fantasy series.

Only those children recognising the target word at the first administration were subsequently presented with the modified versions of the word as described below.

In the second administration, the six modified versions of target words with a change in the central letters (e.g., ) were presented to those pre-schoolers who had previously recognised the original graphic version of the word and they were asked to say what was written in them.

In the third and last administration, pre-schoolers were asked to read the six modified versions of target words with the change to initial/final letters (e.g., ) and then in the graphic form in capital letters without logos (i.e., NEMO, MADAGASCAR, SHREK, ESTATHÉ, GORMITI, WINX).

For the modified versions of the target words in the central letters and in the initial/final letters, the child's types of reading were coded as described in the following:

- **NOTATIONAL AWARENESS (NA)**: assigned when the child reads the modified word correctly via a “Sound-sign processing when reading”, e.g., child reads the word correctly as it is written and reads the modified word “Sadagascar”. This indicates the child's ability to read the word correctly in its original form and modified versions.
- **EMERGENT NOTATIONAL AWARENESS (ENA)**: when presented to children the original version or modified versions of the target word, e.g., child realises that there is something dissonant in the modified words and does not read by saying, for example, that they do not recognise the word because “it has a little letter that is not the right one”. This indicates the child's sensitivity to the variation of the pattern of letter or graphic form denoting the early emergence of notational awareness (e.g., the modified version “NOMO” of the original word “NEMO”).

- **LOGOGRAPHIC RECOGNITION (LR)**: assigned when the modified word was read as it was correctly written, e.g., child reads “Madagascar” where “Sadagascar” was written. This indicates that children continue to use logographic processing in reading the modified versions of words as if they were written correctly in their basic form (thus not noticing the changes made and relying only on the graphical aspects, thus committing reading errors).

For the modified versions of the target words in the graphic form in capital letters without logos, the children's types of reading were coded as described in the following:

- **YES**: when children read
- **NO**: when children do not read

2.3. Data analysis

To perform the analyses on the logos modified in the middle letters and in the beginning and ending letters, we distinguished two categories of scores: (1) children who read through logographic recognition, and (2) children who had notational awareness or emergent notational awareness of the word.

Preliminarily, the sample distribution across schools and parental educational level were calculated. Regarding the first aim, frequencies and percentages for the type of reading manifested by the children when reading the original and modified versions of words were calculated, also in relation to the three different preschools. The Chi-Square Goodness-of-Fit Test was used to compare observed and expected frequencies in each category. The proportion of cases expected in each group of the categorical variable can be equal or unequal (Balakrishnan et al., 2013). In view of this, we expect there to be differences in the variables considered for our sample. Values with $p < 0.05$ were considered significant.

Then, a Chi-square test was used to verify if there are differences between the words with a transparent orthography (i.e., Italian) and the words with an opaque orthography (i.e., English).

Regarding the second aim, Chi-square tests were used to verify the association between the type of reading presented by the children in relation to the type of school and socio-cultural information from parents. In addition, Bonferroni's method with adjustment of p value for comparison of column proportions was used. Values with $p < 0.05$ were considered significant. Further corrections for multiple comparisons are not necessary if there are three groups being compared, since the degrees of freedom have already been taken into account (Shaffer, 1986). In the case of significance, Phi or Cramer's V values were calculated. Phi is a measure of the strength of association between two categorical variables in a 2×2 contingency table. Whereas, Cramer's V is an alternative to Phi in tables larger than 2×2 . Both coefficients range between 0 and 1 with no negative values. Thus, a value close to 0 means that there is no association. However, a value above 0.25 is considered a very strong relationship for Cramer's V. In line with Liebetrau's (1983) interpretations, the following cut-offs can be defined: > 0.25 Very strong; > 0.15 Strong; > 0.10 Moderate; > 0.05 Weak; > 0 None or very weak.

3. Results

First, to examine the sample distribution, a Chi-square test was performed. The percentages of parental educational level show that *School 1* has a significantly higher percentage of children's parents with a university degree, in comparison with *School 2* and *School 3* [$\chi^2(4) = 31.82, p < 0.001$, Cramer's $V = 0.41$]. Table of percentages was reported in Participants section. At the preliminary administration of the original versions of the six logos, the results showed that almost all pre-schoolers correctly read the logos with a percentage ranging from 83 to 100%, as detailed in [Table 2](#). At this stage, pre-schoolers who read the logo correctly could be children in the logographic phase who read the logo correctly as a result of recognition of words based on visual cues or children who read the logo correctly as a result of sound-sign processing. This step was useful to verify that the majority of pre-schoolers could continue in the subsequent research steps that better clarify the type of processing used by pre-schoolers to read words.

Regarding pre-schoolers' ability to correctly read the modified versions of logos, [Table 3](#) reports the results of the frequency percentages of reading types for each modified version of the six target logos (NA - Notational Awareness; LR - Logographic recognition; ENA - Emergent Notational Awareness). The results of the Chi-Square Goodness-of-Fit Test show statistically significant differences in some categories.

Specifically, the results show that in the case where logos were changed in the central letters statistically significant differences in favour of reading with logographic recognition were found in the following stimuli: Madagascar [$\chi^2(1) = 6.86, p < 0.01$]; Shrek [$\chi^2(1) = 31.82, p < 0.001$]; Estathé [$\chi^2(1) = 3.76, p < 0.05$]; Gormiti [$\chi^2(1) = 13.79, p < 0.001$]; and, Winx [$\chi^2(1) = 22.51, p < 0.001$]. At the same time, statistically significant differences were found in favour of children correctly reading the word regardless of the logo or not reading because they notice the logo change in the following stimulus compared to those children reading in logographic: Nemo [$\chi^2(1) = 54.08, p < 0.001$].

On the other hand, in the case of logo modification in the initial and final letters, in all stimuli, it is observed that the proportion of children who read logographically does not differ statistically from those with notational awareness. Except for Shrek [$\chi^2(1) = 3.88, p < 0.05$]; in this case, frequency comparison shows that statistically, most children read through logographic recognition.

Finally, statistically significant differences were observed in reading the word capital letters of the logo in all stimuli: Madagascar [$\chi^2(1) = 4.76, p < 0.05$]; Shrek [$\chi^2(1) = 21.77, p < 0.001$]; Estathé [$\chi^2(1) = 16.79, p < 0.001$]; Gormiti [$\chi^2(1) = 10.89, p < 0.001$]; and, Winx [$\chi^2(1) = 14.09, p < 0.001$], except for the Nemo stimulus. In all cases, most of the children did not read the modified word in capital letters without the logo.

To verify if there are differences between the Italian words and English words, first, we selected the most frequently read Italian word and we compared it with the most frequently read English word. Among the words with transparent spelling, most children correctly read the word Nemo (41%), followed by the words: Madagascar (38%), Gormiti (33%) and Estathé (28%). Among the opaque spelling words, most children correctly read the word Winx (30%), followed by the word Shrek (26%). The two most easily recognised words were selected: Winx and Nemo. The results of





the Chi-square test show that there is a statistically significant difference between the two words [$\chi^2(1) = 59.15, p < 0.001$].

Regarding the second aim, the frequency and percentages of reading types of the modified versions of the six target logos (NA - Notational Awareness; LR - Logographic recognition; ENA - Emergent Notational Awareness; for modified word in capital letters: YES - reads correctly; NO - does not read) across the three preschool types of school are shown in [Table 3](#). Results of the Chi-square test with Bonferroni's method with adjustment of p value show that there is no significant association between school type and the reading of the original logo, nor between school type and the reading of the logo when the change is made in the middle letters. Regarding the reading of the logo with changes in the initial and final letters, there are no significant associations with school type except for the following stimuli: Madagascar [$\chi^2(4) = 14.86, p < 0.01$, Cramer's $V = 0.30$] and Winx [$\chi^2(4) = 12.43, p < 0.05$, Cramer's $V = 0.26$]. The results showed that in all cases the association was between reading of the logo with changes in initial and final letters with School 1.

On the other hand, the results show strong associations between school type and logo reading in word in capital letters. Specifically: Nemo [$\chi^2(2) = 7.70, p < 0.05$, Cramer's $V = 0.30$]; Madagascar [$\chi^2(2) = 10.27, p < 0.01$, Cramer's $V = 0.35$]; Shrek [$\chi^2(2) = 8.47, p < 0.05$, Cramer's $V = 0.30$]; Estathé [$\chi^2(2) = 11.02, p < 0.01$, Cramer's $V = 0.36$]; Gormiti [$\chi^2(2) = 11.28, p < 0.01$, Cramer's $V = 0.35$]; and, Winx [$\chi^2(2) = 12.62, p < 0.01$, Cramer's $V = 0.37$]. The results showed that in all cases the strongest association was between correct reading of word in capital letters with School 1, compared to those who do not read.

4. Discussion

This study investigates the emergence of sign-sound correspondence, named notational awareness, in Italian-speaking 5-year-old pre-schoolers. Studies in the Italian language system (Pinto et al., 2012, 2017; Bigozzi et al., 2016a,b) and in other orthographies (Treiman and Kessler, 2013; Puranik et al., 2014) indicate that sign-sound correspondence achievement is a key skill for the development of children's reading and writing skills. In accordance with the literature (Garcia-Mila et al., 2004), we assumed that notational awareness needs to be constructed by young children along a process that sees the production, reflection, and interpretation of one's own notations (e.g., invented spelling activities) and notations provided by the surrounding world (e.g., preschool and home environment) and others (e.g., educators and parents). The literature (Frith, 1986) states that children around age 5 are in the logographic phase. Our results contribute to further explore the logographic phase as a window into the emergence and early development of sound-sign processing when decoding words in 5-year-old pre-schoolers acquiring a transparent language like Italian and when notational awareness fully takes over.

Regarding the first aim of evaluating and comparing the type of reading in Italian-speaking pre-schoolers, it is interesting to observe that the results show a consistent part of pre-schoolers able to use the sign-sound correspondences when decoding the modified versions of logos. The correct reading of words with modification in the central letters is around 29% of the sample for the target words , ,  and goes as high as 36% for the word . Words with a lower percentage

of correct reading are **SHREK** and **Win2** with a correct reading percentage of around 20%. The literature (see Frith, 1986) informs us that children around the age of 5 are in the logographic phase given that they are able to recognise words by relying on visual and contextual cues without decomposing the words into smaller units and spell on a letter-by-letter basis. Our results add to the inspection of the logographic phase by showing a more complex pattern of reading profiles. Our results, indeed, showed the




large presence of 5-year-old pre-schoolers with a high notational awareness that allows sign-sound processing before school entry because they read the word correctly despite still having retained the logo without being fooled by the logo, thus attributing the correct sound to each sign or other pre-schoolers who report that they cannot read because they realise that some letters have changed and therefore still do not get fooled by the logo showing an emergent level of notational awareness.

TABLE 3 Frequency rates for the reading type for each version of the target word (original or modified) and frequencies and percentages according to the school attended.

Word list number	Original version		Modified version – central letter			Modified version – initial or final letter			Modified version – word in capital letters without logo	
	NA	ENA	NA	ENA	LR	NA	ENA	LR	YES	NO
Word list 1	NEMO		NOMO			MEMO			NEMO	
%	88	12	38	9	26	46	6	48	41	59
School 1	23/25 92%	2/25 8%	13/23 56.6%	2/23 8.7%	8/23 34.8%	13/23 56.5%	3/23 13.1%	7/23 30.4%	14/23 60.9%	9/23 39.1%
School 2	30/32 93.7%	2/32 6.2%	8/30 26.7%	4/30 13.3%	18/30 60%	14/30 46.6%	2/30 6.6%	14/30 46.6%	13/30 43.3%	17/30 56.6%
School 3	30/37 81.1%	7/37 18.9%	9/30 30%	2/30 6.6%	19/30 63.3%	11/30 36.6%	0/30 0%	19/30 63.3%	7/30 23.3%	23/30 76.6%
Word list 2	MADAGASCAR		MADAGASCAR			SODAGASCAR			MADAGASCAR	
%	90	10	30	6	64	38	16	46	38	62
School 1	23/25 92%	2/25 8%	12/23 52.2%	1/23 4.3%	10/23 43.5%	14/23 60.9%	1/23 4.3%	8/23 34.8%	14/23 60.9%	9/23 39.1%
School 2	29/32 90.6%	3/32 9.4%	7/29 24.1%	2/29 6.9%	20/29 69%	11/29 37.9%	2/29 6.9%	16/29 55.2%	12/29 41.4%	17/29 58.6%
School 3	33/37 89.2%	4/37 10.8%	6/33 18.9%	2/33 6.06%	24/33 72.7%	7/33 21.2%	10/3330.3%	15/33 45.4%	6/33 18.9%	26/33 78.8%
Word list 3	SHREK		SHREK			CHREK			SHREK	
%	99	1	20	2	77	34	5	60	26	74
School 1	25/25 100%	0/25 0%	25/25 100%	1/25 4%	18/25 72%	11/25 44%	3/25 12%	11/25 44%	11/25 44%	14/25 56%
School 2	32/32 100%	0/32 0%	32/32 100%	0/32 0%	25/32 78.1%	11/32 34.4%	2/32 6.2%	19/32 59.4%	9/32 28.1%	23/32 71.9%
School 3	36/37 97.3%	1/37 2.7%	36/37 97.3%	1/36 2.7%	29/36 80.5%	10/36 27.7%	0/36 0%	26/36 72.2%	4/36 11.1%	32/36 88.8%
Word list 4	ESTATHE		ESTATUE			ESTATHI			ESTATHÈ	
%	91	9	29	10	61	35	9	55	28	72
School 1	22/25 88%	3/25 12%	11/22 50%	2/22 9.09%	9/22 40.9%	12/22 54.5%	1/22 4.54%	9/22 40.9%	12/22 54.5%	10/22 45.4%
School 2	30/32 93.7%	2/32 6.2%	7/30 23.3%	4/30 13.3%	19/30 63.3%	10/30 33.3%	3/30 10%	17/30 56.7%	7/30 23.3%	23/30 76.7%
School 3	34/37 91.8%	3/37 8.1%	7/34 20.5%	3/34 8.8%	24/34 70.6%	8/34 23.5%	4/34 11.7%	22/34 64.7%	5/34 14.7%	29/34 85.3%
Word list 5	GORTITI		GORTITI			MORMITI			GORMITI	
%	100	0	29	2	69	39	2	59	33	67
School 1	25/25 100%	0/25 0%	11/25 44%	1/25 4%	13/25 52%	13/25 52%	0/25 0%	12/25 48%	13/25 52%	12/25 48%
School 2	32/32 100%	0/32 0%	7/32 21.9%	0/32 0%	25/32 78.1%	13/32 40.6%	1/32 3.1%	18/32 56.2%	13/32 40.6%	19/32 59.4%
School 3	36/37 97.3%	1/37 2.7%	9/37 24.3%	1/37 2.7%	27/37 73%	11/37 29.7%	1/37 2.7%	25/37 67.6%	5/37 13.5%	32/37 86.5%



(Continued)


TABLE 3 (Continued)


Word list number	Original version		Modified version – central letter			Modified version – initial or final letter			Modified version – word in capital letters without logo	
	NA	ENA	NA	ENA	LR	NA	ENA	LR	YES	NO
Word list 6									WINX	
%	100	0	20	5	75	31	13	56	30	70
<i>School 1</i>	25/25 100%	0/25 0%	8/25 32%	3/25 12%	14/25 56%	12/25 48%	2/25 8%	11/25 44%	12/25 48%	11/25 44%
<i>School 2</i>	32/32 100%	0/32 0%	6/32 18.7%	0/32 0%	26/32 81.2%	12/32 37.5%	6/32 18.7%	14/32 43.7%	12/32 37.5%	20/32 62.5%
<i>School 3</i>	37/37 100%	0/37 0%	5/37 13.5%	2/37 5.4%	30/37 81.1%	5/37 13.5%	4/37 10.8%	28/37 75.7%	4/37 10.8%	33/37 89.2%

The last column of the modified version of the target word in terms of graphic cues was not inserted given that not having the logo children cannot read the word via logographic processing.

For an effective use of the sound-sign processing children must represent the information in symbols and operate a matching between grapheme and phoneme (Ravid and Tolchinsky, 2002). In transparent orthographies like Italian, this sound-sign matching is biunivocal despite what happens in opaque orthographies in which the grapheme level does not match the oral version of the word (Caravolas, 2013).

Regarding the modified version  in the central letters of the target word , our results showed a high proportion of pre-schoolers who have overcome the logographic processing and realise that one or more graphemes have been modified by using the sound-sign processing. There is also an interesting number of children showing the emergence of the sound-sign correspondence; indeed, some pre-schoolers realise that there is something dissonant in the modified words and do not read by saying for example that they do not recognise the word because “it has a little letter that is not the right one” or “this letter I don’t know how to say...”, thus still noticing a change and informing us about their paying attention to the individual letters that make up the word. It could be also possible that the four letters, fully capitalised may facilitate young children’s recognition and sound-sign processing given the decreased cognitive demands required for decoding words (Ober et al., 2020).

Regarding the modified version in the initial/final letters, the results of the comparisons of Italian-speaking 5-year-old pre-schoolers’ reading type confirm the pattern previously described given that there is an important proportion of pre-schoolers reading correctly or not reading, thus manifesting the sound-sign correspondence achieved or at its emergence. A child who, smiling with amusement at the reading of the word  in the modified form, said: “Mormiti!....that’s wrong....you can’t write Gormiti with an M”, demonstrating that his attention was entirely devoted to the letters. Our results are in line with Coltheart’s (1981) theory according to which 5-year-old children operate for the recognition of words as a particular sequence of letters, thus making them capable of detecting changes made at the alphabetical level. Those pre-schoolers using sound-sign processing also confirm the tendency in easily recognising the initial and final sounds of words (Pinto, 2003), while other pre-schoolers show a still dominant logographic recognition via visual cues even when the word is altered in its salient initial or final components. This is particularly evident for the word list “Shrek”, in fact, our results pointed out that most children read its modified version in initial/final letters as the

original logo, thus logographic processing remains dominant in this case. One explanation may be that the characteristics of the word  which has a lower-case letter and is not an Italian word that does not facilitate young children’s recognition and sound-sign processing. Our results show that children performed sound-sign processing significantly more when reading the word Nemo that is a disyllabic word that follows the sound-sign transposition as in Italian in comparison to the word Winx that contains letters not present in the Italian alphabet.

Finally, regarding the modified words written in capital letters without the logo, the results confirmed that a great proportion of 5-year-old children can read the word via sound-sign processing, while others who cannot read indicate that they are developing notational awareness. As highlighted in the literature, it is better to conceive the achievement of the sound-sign correspondence as a developmental process during which several phases mark the differential emergence of representational systems (Yamagata, 2007).

For what concerns the second aim, we assumed that preschool curriculum and home literacy practices may enhance pre-schooler’s achievement of the sound-sign correspondence which is at the basis of later reading and writing acquisitions (Incognito et al., 2021; Bigozzi et al., 2023, in press). We thus decided to analyse the existence of differences in type of reading in the light of socio-cultural characteristics pertaining to preschools and parents’ levels of education. A preliminary inspection of data showed that *School 1* reported a significantly higher percentage of children’s parents with a university degree, in comparison with *School 2* and *School 3*, meanwhile *School 3* reported a significantly higher percentage of high-school diploma and middle-school diploma in comparison to *School 2* and *School 1*. The results of the comparisons showed an association between correct reading type and *School 1* in comparison to *School 2* and *School 3*. Especially in the last phase of the research, the results of *School 1* stand out positively in terms of the number of correct answers, in fact more than half of the children in this school read the word written in block letters correctly. The results of the χ^2 test show an association between *School 1* and the ability to read the modified words written in capital letters without the logo. A closer look at data shows that in *School 1* the great majority of parents have a *University Degree* and the other part have a *High-School Diploma*, while no parents have a *Middle-School Diploma*. In *School 2* the number of graduates decreases even more dramatically in *School 3* where 45% of parents have a *High-School*

Diploma and only 13% have a *University Degree*. Therefore, these results confirm that there is an influence of the socio-cultural level of parents. Growing up in a rich environment in terms of literacy materials and experiences contribute to pre-schoolers' emergent literacy skills development (Incognito and Pinto, 2021) and word recognition (Evans and Shaw, 2008).

These findings, while exploratory and preliminary in nature, extend our limited understanding about the emergence of sign-sound processing in Italian-speaking 5-year-old pre-schoolers. The sign-sound correspondence achievement in pre-schoolers is critical for later reading and writing development (Tijms et al., 2020). Our results have educational implications by highlighting that literacy acquisition experience may promote young children's awareness of sound-sign correspondence. As suggested in the literature, there is a high probability that high levels of education in parents are associated with higher home literacy environments and practices. The results from the current study indicate that the home literacy environment and practices promote children's early literacy skills, such as notational awareness. Thus, parents should encourage children in joint reading and writing activities at home or invented reading and invented spelling practices, which require many important skills such as the use of grapheme-phoneme correspondence or the early letter knowledge and their word meaning. Previous research conducted in the Italian context suggests that helping pre-schoolers to link letters with their corresponding sounds benefits their notational awareness (Incognito et al., 2021; Bigozzi et al., 2023, in press). Also, as suggested by Coltheart (1981), it is important to expose the child to words and logos to facilitate their knowledge and storing in memory. Finally, the results from the current study indicate that implementing research-based programmes, such as Promoting Sound Sign Achievement (PASSI; Pinto et al., 2018), can be effectively fostered during the preschool years of notational awareness, particularly with educators' guidance and peers' involvement in small group activities.

4.1. Limitations and future research

The children observed in this study were a group of 5-year-old children. Future longitudinal studies which will include a larger sample of young children are needed. A longitudinal research design will help to gain a more comprehensive view of the phase in which children disentangle logographic and sound-sign processing. Moreover, given the results indicating that children's type of reading varied in relation to sociocultural status pertaining to type of school and parents' educational level, it is important for future research to be able to generalise to other populations of children the results from this study. In this respect, a more detailed examination of children's home literacy environment and practices would be useful to verify similarity across cultures

and orthographies in terms of how parents support children's sound-sign correspondence achievement. The disentanglement between logographic and sound-sign correspondence in pre-schoolers might show different patterns of development in a cross-linguistic perspective. The acquisition of the English language, which is a language with an opaque orthography whose letter-sound correspondence relations do not have a coherent pattern, presents different difficulties and different developmental trajectories compared to Italian, which, being a language with a transparent orthography, is characterised by regular letter-sound correspondences.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by University of Florence. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Albuquerque, A., and Alves Martins, M. (2022). Invented spelling as a tool to develop early literacy: The predictive effect on reading and spelling acquisition in Portuguese. *J. Writ. Res.* 14, 113–131. doi: 10.17239/jowr-2022.14.01.04
- Aro, M., and Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Appl. Psycholinguist.* 24, 621–635. doi: 10.1017/S0142716403000316

- Balakrishnan, N., Voinov, V., and Nikulin, M. S. (2013). *Chi-squared goodness of fit tests with applications*. Cambridge, MA: Academic Press.
- Bastien-Toniazzo, M., and Jullien, S. (2001). Nature and importance of the logographic phase in learning to read. *Read. Writ.* 14, 119–143. doi: 10.1023/A:1008138931686
- Bigozzi, L., Tarchi, C., Caudek, C., and Pinto, G. (2016a). Predicting reading and spelling disorders: A 4-year cohort study. *Front. Psychol.* 7:337. doi: 10.3389/fpsyg.2016.00337
- Bigozzi, L., Tarchi, C., Pezzica, S., and Pinto, G. (2016b). Evaluating the predictive impact of an emergent literacy model on dyslexia in Italian children: A four-year prospective cohort study. *J. Learn. Disabil.* 49, 51–64. doi: 10.1177/0022219414522708
- Bigozzi, L., Vettori, G., and Incognito, O. (2023). The role of kindergartners' home literacy environment and emergent literacy skills on later reading and writing skills in primary school: A mediational model. *Front. Psychol.* 14:1113822. doi: 10.3389/fpsyg.2023.1113822
- Bishop, A. G. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screenings. *Learn. Disabil. Q.* 26, 189–200. doi: 10.2307/1593651
- Caravolas, M. (2013). "Learning to spell in different languages: How orthographic variables might affect early literacy," in *Handbook of orthography and literacy*, eds R. M. Joshi and P. Aaron (Milton Park: Routledge), 511–526. doi: 10.1111/j.1469-7610.2006.01725.x
- Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Lang.* 15, 245–286.
- Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Read. Res. Q.* 55, 45–60.
- Ellis, N., and Large, B. (1988). The early stage of reading: A longitudinal study. *Appl. Cogn. Psychol.* 2, 47–76.
- Evans, M. A., and Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Can. Psychol.* 49, 89–95. doi: 10.1037/0708-5591.49.2.89
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Ann. Dyslexia* 36, 67–81. doi: 10.1007/BF02648022
- Garcia-Mila, M., Marti, E., and Teberosky, A. (2004). Emergent notational understanding: Educational challenges from a developmental perspective. *Theory Pract.* 43, 287–294. doi: 10.1207/s15430421tip4304_7
- Hall, A. H., Simpson, A., Guo, Y., and Wang, S. (2015). Examining the effects of preschool writing instruction on emergent literacy skills: A systematic review of the literature. *Lit. Res. Instr.* 54, 115–134.
- Hand, E. D., Lonigan, C. J., and Puranik, C. S. (2022). Prediction of kindergarten and first-grade reading skills: Unique contributions of preschool writing and early-literacy skills. *Read. Writ.* 1–24. doi: 10.1007/s11145-022-10330-1
- Harris, M., and Coltheart, M. (1986). *Language processing in children and adults: An introduction*. Routledge & Kegan Paul, London. Trad. It. L'elaborazione del linguaggio nei bambini e negli adulti. Bologna: Il Mulino.
- Incognito, O., and Pinto, G. (2021). Longitudinal effects of family and school context on the development of emergent literacy skills in preschoolers. *Curr. Psychol.* 42, 9819–9829. doi: 10.1186/s12913-016-1423-5
- Incognito, O., Bigozzi, L., Vettori, G., and Pinto, G. (2021). Efficacy of two school-based interventions on notational ability of bilingual pre-schoolers: A group-randomized trial study. *Front. Psychol.* 12:686285. doi: 10.3389/fpsyg.2021.686285
- Kanizsa, G. (1983). *Perception and language*. Bologna: Il Mulino.
- Kuhn, M. R., Schwanenflugel, P. J., and Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Read. Res. Q.* 45, 230–251.
- Liebetrau, A. M. (1983). *Measures of association*, Vol. 32. Newbury Park, CA: Sage.
- Marinelli, C. V., Romani, C., Burani, C., and Zoccolotti, P. (2015). Spelling acquisition in English and Italian: A cross-linguistic study. *Front. Psychol.* 6:1843. doi: 10.3389/fpsyg.2015.01843
- Masonheimer, P. E., Drum, P. A., and Ehri, L. C. (1984). Does environmental print identification lead children into word reading? *J. Read. Behav.* 16, 257–271.
- Milburn, T. F., Hipfner-Boucher, K., Weitzman, E., Greenberg, J., Pelletier, J., and Girolametto, L. (2017). Cognitive, linguistic and print-related predictors of preschool children's word spelling and name writing. *J. Early Child. Lit.* 17, 111–136. doi: 10.1177/1468798415624482
- Missall, K., Reschly, A., Betts, J., McConnell, S., Heistad, D., Pickart, M., et al. (2007). Examination of the predictive validity of preschool early literacy skills. *School Psychol. Rev.* 36, 433–452. doi: 10.1080/02796015.2007.12087932
- Mohamed, M. B. H., and O'Brien, B. A. (2022). Defining the relationship between fine motor visual-spatial integration and reading and spelling. *Read. Writ.* 35, 877–898. doi: 10.1007/s11145-021-10165-2
- Mol, S. E., and Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychol. Bull.* 137, 267–296.
- Neumann, M. M., Hood, M., Ford, R. M., and Neumann, D. L. (2012). The role of environmental print in emergent literacy. *J. Early Child. Lit.* 12, 231–258. doi: 10.1177/1468798411417080
- Nicholls, A. L., and Kennedy, J. M. (1992). Drawing development: From similarity of features to direction. *Child Dev.* 63, 227–241.
- Ober, T. M., Brooks, P. J., Homer, B. D., and Rindskopf, D. (2020). Executive functions and decoding in children and adolescents: A meta-analytic investigation. *Educ. Psychol. Rev.* 32, 735–763. doi: 10.1007/s10648-020-09526-0
- Ouellette, G., and Sénéchal, M. (2017). Invented spelling in kindergarten as a predictor of reading and spelling in Grade 1: A new pathway to literacy, or just the same road, less known? *Dev. Psychol.* 53, 77–88. doi: 10.1037/dev0000179
- Pinto, G. (2003). *Il suono, il segno, il significato. Psicologia dei processi di alfabetizzazione*. Roma: Carocci.
- Pinto, G., and Incognito, O. (2022). The relationship between emergent drawing, emergent writing, and visual-motor integration in preschool children. *Infant Child Dev.* 31:e2284. doi: 10.1002/icd.2284
- Pinto, G., Bigozzi, L., Gamannossi, B. A., and Vezzani, C. (2009). Emergent literacy and learning to write: A predictive model for Italian language. *Eur. J. Psychol. Educ.* 24, 61–78.
- Pinto, G., Bigozzi, L., Gamannossi, B. A., and Vezzani, C. (2012). Emergent literacy and early writing skills. *J. Genet. Psychol.* 173, 330–354. doi: 10.1080/00221325.2011.609848
- Pinto, G., Bigozzi, L., Tarchi, C., and Camilloni, M. (2018). Improving conceptual knowledge of the Italian writing system in kindergarten: A cluster randomized trial. *Front. Psychol.* 9:1396. doi: 10.3389/fpsyg.2018.01396
- Pinto, G., Bigozzi, L., Vezzani, C., and Tarchi, C. (2017). Emergent literacy and reading acquisition: A longitudinal study from kindergarten to primary school. *Eur. J. Psychol. Educ.* 32, 571–587. doi: 10.1007/s10212-016-0314-9
- Puranik, C. S., Petscher, Y., and Lonigan, C. J. (2014). Learning to write letters: Examination of student and letter factors. *J. Exp. Child Psychol.* 128, 152–170.
- Ravid, D., and Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *J. Child Lang.* 29, 417–447. doi: 10.1017/S0305000902005111
- Rayner, K., Pollatsek, A., Ashby, J., and Clifton, Jr, C. (2012). *Psychology of reading*. London: Psychology Press.
- Rohde, L. (2015). The comprehensive emergent literacy model: Early literacy in context. *Sage Open* 5:2158244015577664.
- Seymour, P. H. K., and Elder, L. (1986). Beginning reading without phonology. *Cogn. Neuropsychol.* 3, 1–37.
- Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *J. Am. Statist. Assoc.* 81, 826–831.
- Spencer, L. H., and Hanley, J. R. (2004). Learning a transparent orthography at five years old: Reading development of children during their first year of formal reading instruction in Wales. *J. Res. Read.* 27, 1–14. doi: 10.1111/j.1467-9817.2004.00210.x
- Tijms, J., Fraga-González, G., Karipidis, I. I., and Brem, S. (2020). The role of letter-speech sound integration in normal and abnormal reading development. *Front. Psychol.* 11:1441. doi: 10.3389/fpsyg.2020.01441
- Treiman, R., and Kessler, B. (2013). Learning to use an alphabetic writing system. *Lang. Learn. Dev.* 9, 317–330.
- Valdois, S., Roulin, J., and Bosse, M. L. (2019). Visual attention modulates reading acquisition. *Vision Res.* 165, 152–161. doi: 10.1016/j.visres.2019.10.011
- Wang, L., and Liu, D. (2021). Unpacking the relations between home literacy environment and word reading in Chinese children: The influence of parental responsive behaviors and parents' difficulties with literacy activities. *Early Child. Res. Q.* 56, 190–200. doi: 10.1016/j.jecresq.2021.04.002
- Whitehurst, G. J., and Lonigan, C. J. (1998). Child development and emergent literacy. *Child Dev.* 69, 848–872.
- Yamagata, K. (2007). Differential emergence of representational systems: Drawings, letters, and numerals. *Cogn. Dev.* 22, 244–257. doi: 10.1016/j.cogdev.2006.10.006
- Zhang, L., and Treiman, R. (2021). Preschool children's knowledge of letter patterns in print. *Sci. Stud. Read.* 25, 371–382. doi: 10.1080/10888438.2020.1801690



OPEN ACCESS

EDITED BY

Olivia López Martínez,
University of Murcia, Spain

REVIEWED BY

María García-Serrano,
University of Salamanca, Spain
Joseph Magliano,
Georgia State University, United States

*CORRESPONDENCE

Andrea Strandberg
✉ andrea.strandberg@ki.se

RECEIVED 17 November 2022

ACCEPTED 28 June 2023

PUBLISHED 25 July 2023

CITATION

Strandberg A, Nilsson M, Östberg P and
Seimyr GÖ (2023) Eye movements are stable
predictors of word reading ability in young
readers.
Front. Educ. 8:1077882.
doi: 10.3389/feduc.2023.1077882

COPYRIGHT

© 2023 Strandberg, Nilsson, Östberg and
Seimyr. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Eye movements are stable predictors of word reading ability in young readers

Andrea Strandberg^{1*}, Mattias Nilsson¹, Per Östberg² and
Gustaf Öqvist Seimyr¹

¹Division of Eye and Vision, Department of Clinical Neuroscience, Karolinska Institutet (KI), Stockholm, Sweden, ²Division of Speech and Language Pathology, Department of Clinical Science, Intervention and Technology, Karolinska Institutet (KI), Stockholm, Sweden

During the first years of formal instruction in reading, there are developmental changes to the characteristics of children's eye movements that correspond to their progress. Generally, these changes are driven by improved text processing and a shift from reliance on sub-lexical to lexical processes. Currently, however, little is known about to what extent early eye movements during reading in ecological contexts account for variations in later word reading ability. In this paper we investigate this association in 164 children in first grade and 206 children in second grade. We recorded their eye movements during normal text passage reading in an unrestricted eye tracking set-up. We analyzed to what extent they account for variation in word reading ability 1 year post-recording, and make comparisons to concurrent predictions. Regression analysis revealed that eye movements accounted for approximately 60% of the variance in concurrent word reading ability and approximately 48% 1 year later. Mean fixation duration explained unique variance in reading ability and largely drives the correlation. Proportion of regressions was not a significant longitudinal predictor among the youngest readers. The difference between the concurrent and longitudinal predictions was greatest in the younger children, which was in line with our expectation. Findings are discussed in relation to current models of word reading. Our results suggest that eye movements are stable predictors of word reading ability. Ultimately, knowledge of what can be gleaned from early readers' natural eye movements about later word reading ability could help inform assessments of reading development in the educational setting, where the presence of digital assessment tools is growing.

KEYWORDS

eye movements, text reading, eye tracking, reading development, word reading ability

1. Introduction

Eye tracking is used to study a variety of topics within several areas of research, such as medicine, education and psychology (Henderson et al., 2015; McDonald et al., 2015; Ekstrand et al., 2021). Its temporal and spatial sensitivity makes it particularly suitable for investigations into the reading process (Carter and Luke, 2020). Indeed, eye movement research has taught us a great deal about the psychological, neurological and linguistic aspects of adult reading (Rayner and McConkie, 1976; Gidlöf et al., 2013; König, 2016; Zhang et al., 2021) while children's reading development long was sparsely studied. As there are differences to the mechanics of text processing depending on proficiency, the findings from adult studies are not always applicable

to young readers (Rayner, 1998; Reichle et al., 2013). However, reading is widely considered an indispensable skill and one of the most cognitively demanding tasks that we engage in on a daily basis. The first school years are critical to its' acquirement, which makes children's reading development an area of continued interest.

1.1. Adults' eye movements during reading

A great deal of eye movement research has contributed to our understanding of the reading process in skilled, mostly adult, readers (Rayner, 1975, 1998; Rayner and McConkie, 1976; Inhoff, 1990). Development of innovative experimental techniques has allowed researchers to investigate word processing and moment-to-moment allocation of attention while the individual reads a text (McConkie and Rayner, 1975; Rayner, 1975; Clifton et al., 2016). Throughout all orthographies studied to this date, adult's eye movements consistently follow a fixation-saccade pattern during reading. This means that when we read, our eyes make a series of stops and between those stops, horizontal movements from left to right across the text (in alphabetic writing systems). Those stops and fast gaze movements are what are referred to as fixations and saccades. During a fixation, the reader holds their gaze still over a word before moving, i.e., making a saccade, onto the next. A number of different variables, all tied to these aspects of online text processing, are analyzed within this field. These are either of spatial (e.g., the distance the eyes move during a saccade), numerical (e.g., the number or proportion of regressions) or temporal nature (e.g., the duration of a fixation; Rayner, 1998). They may also move the gaze backwards, which is called a regressive saccade or regression. Variations in eye movement characteristics reflect the relative ease with which the reader progresses through the text (Rayner, 1977; Yang, 2012). In general, fluctuations in those variables are interpreted as indications of text processing difficulty rather than the cause thereof (Rayner, 1977; Just and Carpenter, 1980).

In adult readers, a fixation typically lasts between 60 and 500 ms (milliseconds) with an average of 250 ms (Rayner, 1998). Fixations are directly affected by the length and the frequency of the fixated, word as well as by lexical predictability. The longer the word, the higher is the probability that the reader refixates it and low-frequency words are fixated for a longer time than high-frequency words (Rayner and McConkie, 1976; Rayner and Duffy, 1986; Liversedge and Findlay, 2000; Joseph et al., 2013). Word frequency effects are really the effects of lexical processing, a necessary operation to access the word's meaning (Reichle et al., 2013). Repeated exposure to print strengthens the links between form and meaning in memory. Over time, word recognition requires shorter and fewer fixations. This process is widely considered an important driving force of eye movements during reading (Rayner and McConkie, 1976; Just and Carpenter 1980; Inhoff, 1984; Rayner, 1986; Rayner and Duffy, 1986; Clifton et al., 2007; Reichle et al., 2009; Rau et al., 2014; Blythe et al., 2015; Tiffin-Richards and Schroeder, 2015).

Rayner et al. (Reichle et al., 2009) describes an initial stage of processing taking place during fixations, where our decoding system quickly determines whether the word is likely to be identified. When a certain threshold is reached, a saccade to the next word is programmed. Following that, the second stage entails lexical access to the word. In parallel, the reader extracts relatively low-level visual and linguistic information, such as letter shape, word length and

phonological cues, parafoveally to the right of the fixation. These estimations of the upcoming print help guide the eyes where to go next (Reichle et al., 1998; Rayner et al., 2004; Pollatsek et al., 2006). Saccades are affected by the processing difficulty experienced by the reader and struggling readers tend to make shorter saccades (Liversedge and Findlay, 2000; Hindmarsh et al., 2021). In English adults reading English text, a progressive saccade is approximately 7–9 letter spaces or 3–4 degrees in amplitude (Rayner, 1998; Liversedge and Findlay, 2000; Seassau and Bucci, 2013). Approximately 15–20% of all saccades are regressive, or regressions, which are associated to several functions. Regressions move the eye opposite of the reading direction, i.e., from right to left in alphabetic orthographies. They may resolve oculomotor targeting errors or let the reader review and reprocess a word with the purpose of validating or updating the reader's comprehension of the text (Inhoff et al., 2019). Adult studies have compellingly shown how cognitive processing is reflected in eye movements during reading. Research focusing on young individuals faces challenges relating to ethics, technology and the complex interplay of cognitive and visual processes during reading development. These factors have contributed to a gap in knowledge on how text processing and processing difficulties are represented in children's eye movements during reading. Nevertheless, progress has been made in the last decades, which we will turn to next.

1.2. Children's eye movements during reading

In the wake of important technical advancements, recent years have seen a growing body of eye tracking research on reading development (Sovik et al., 2000; Blythe et al., 2006, 2011, 2015; Häikiö et al., 2009; Joseph et al., 2013; Reichle et al., 2013; Blythe, 2014; Vorstius et al., 2014; Spichtig et al., 2016; Kornev et al., 2019) which have examined how progress is reflected in young readers' eye movements (Blythe, 2014). Young readers have a higher number of fixations and longer fixation durations than older, proficient readers. They make a higher number of short forward saccades and make regressive saccades more often than adults (Rayner, 1986; Inhoff et al., 2019). In general, regressions decrease in frequency as the child's reading improves and the occurrence of errors relating to both decoding and comprehension become rarer. Children's eye movements during reading largely resemble those of adults around 10–12 years of age (Rayner and Duffy, 1986; Blythe et al., 2009; Reichle et al., 2013; Blythe, 2014; Strandberg et al., 2022). Commonly, studies across age groups have informed researchers about this development (Blythe et al., 2009, 2011; Häikiö et al., 2009). Like adults', children's gaze durations are significantly longer for low than high-frequency words (Joseph et al., 2013). Word recognition processes involve a cognitive load which presumably is higher in beginning than skilled readers. This affects the preprocessing of upcoming letters and any other retrievable information from their perceptual span, which is smaller than that of adults (Rayner, 1986; Kim et al., 2019). Indeed, the gaze behavior of adult, struggling readers share several characteristics with that of very young readers (Prado et al., 2007; Barnes and Kim, 2016). Struggling readers tend to have longer fixations, a higher number of fixations and saccades, and make shorter forward saccades which is indicative of longer processing times (McConkie and Rayner, 1975; Rayner, 1998; Al Dahhan et al., 2014).

Hence, the number and duration of fixations decrease, refixations decrease, the number of saccades decreases while forward saccade amplitude increases, and the proportion of regressive saccades decrease over time despite increased text difficulty (Blythe, 2014; Strandberg et al., 2022). This account of young readers' eye movements is consistent with current developmental models of word reading (Blythe et al., 2011; Reichle et al., 2013; Blythe, 2014). Above, lexical processing is mentioned as a crucial route during reading. However, very young children largely depend on sub-lexical processes to decode text. Children initially learn how to read small units of print, while simultaneously trying to make the connections between orthography to phonology and semantics (Rau et al., 2014, 2015b). Over time, when the necessary links have been established, the quite demanding letter-by-letter processing that initiates reading acquisition gradually develops to automatized orthographic reading (LaBerge and Jay Samuels, 1974; Ehri, 2005, 2014).

The research that we have described here has developed our understanding of children's text processing and the knowledge that can be gained from their eye movements during reading. Some areas are still poorly understood, including what, if any, conclusions that can be drawn about future word reading ability based on early eye movements. We know that current reading skill is reflected in gaze patterns. However, it is unclear with what certainty we can estimate children's reading ability at a later point based on early reading eye movements.

1.3. Underpinnings of the present study

While the branch of reading research that focuses on text comprehension often deals with higher-level processes, the Lexical Quality Hypothesis developed by Perfetti and Hart (2001) emphasizes a lexical basis of comprehension. In this view, understanding of word meanings is central to global text comprehension. Perfetti (2010) describes a two-fold mechanism of progress; understanding a text lets children add specific semantic word knowledge to their vocabularies, while it also allows application of that new knowledge when reading other texts that contain those words. Several accounts of inter-reader differences focus on cognitive differences, which do not fully explain differences in knowledge of word meanings that remain throughout literacy development (Perfetti, 2010). Evidence that vocabulary training has effects on global comprehension (Beck et al., 1982) suggests that word knowledge should be taken into account when considering individual differences in reading. In this paper, we analyze a composition of tests that reflect fluency, accuracy and speed. Altogether, it is representative of abilities that support word level processes (Kamhi and Catts 2005). We assume word level processing to be a significant marker of reading ability and important for the continued reading development. While the assessment carried out highlights decoding and processing of words, rather than knowledge of words, there is arguably tight interplay between the skills during reading. We have chosen this stance based on a paradigm where word reading and comprehension are considered dissociated abilities, insofar that they are supported by partly different sets of underlying component skills. However, there is also an overlap between word level operations and text comprehension in developing readers (Hosp et al., 2019). In this study, we analyze word reading ability because smooth decoding is a crucial component of the reading process. It

allows rapid word identification and retrieval of word meanings (Perfetti, 2010). We assume that it to some extent reflects abilities that assist text comprehension as well.

Assessments of reading ability within the educational setting are largely dependent on offline measures (Kim et al., 2019). It often demands some response on the student's behalf, be it verbal or written. Additionally, assessments are essentially always multipart. While they may give the instructor an idea of the students' reading ability, they are time consuming. According to Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), more than 50 different tests are used to diagnose reading difficulties and none of them have been scientifically evaluated. On an international level, the report found scientific basis for very few assessments (Statens beredning för medicinsk utvärdering (SBU), 2014). Furthermore, nearly all screening protocols available require manual scoring and subjective assessment of the performance, which entails a risk of interrater variation. Finally, records show that the average age of diagnosis of reading difficulties in Sweden is 13 years, while early detection for reading difficulties is significant in terms of success of treatment (Denton et al., 2006; Zijlstra et al., 2021). Investigations of behavioral measures in this context have yielded some promising results. Eye movement patterns have been used to identify students with reading difficulties with high accuracy in a few studies (Rello and Ballesteros 2015; Ygge et al., 2016). The current study attempts to extend findings from this area to a diverse sample and a natural reading situation in a longitudinal design. Eye movement analysis could be used in the educational setting in the future, given a continued development of eye tracking technology toward accessibility and more manageable administration. It could potentially offer a fast analysis and an objective perspective in reading assessments in the academic or clinical setting. While most available offline tests of reading provide information about current reading ability, longitudinal coherence of the outcomes or predictive validity of the test materials is most often not evaluated. An ultimate goal for gaze behavior analysis could be to work as a complementary tool that provides quick and highly accurate predictions of word reading ability in young children (Rello and Ballesteros 2015; Ygge et al., 2016).

There is considerable variation to the methodological aspects of previous eye movement research on children's reading. Oftentimes, data is collected under controlled experimental conditions, most often in a research lab. Elements that minimize the influence of movements are common, such as head-tracking devices and chin- or head-rests. Other constraints concern how the stimuli are modeled and presented. Researchers can use gaze contingent techniques (such as the disappearing text or moving window paradigm; McConkie and Rayner, 1975, 1976) and present the participant with letters, words or sentences; one-by-one or sequentially. Restrained set-ups are associated to precise results and have contributed significant insight into the reading process. However, as is frequently the case where a phenomenon that naturally occurs in uncontrolled environments is studied in a controlled setting, the inherent characteristics of highly controlled set-ups may decrease the generalizability of findings. Their relevance for the educational setting, for example, where natural reading of print or on screens is practiced is theoretically somewhat limited. Unrestrained set-ups are associated with larger margins of error and more variability and are less common in this field of research. Hence, there is less knowledge about children's eye movements during reading in authentic contexts. Therefore, the

participants of the current study read normal text passages at the location of their school to resemble a normal reading situation and to represent the probabilistic, uncertain nature of the everyday environment.

Reading ability is crucial throughout life, while its foundation is established during early childhood. Despite progress during recent years, investigations into children's reading continue to be of importance. So far, little research investigates children's eye movements and reading ability over time (Spichtig et al., 2016; Kim et al., 2019) within the same subjects. We know that gaze behavior during reading is a good real-time measure of the relative ease with which the reader processes the text. If early eye movements can account for variation in later word reading skill in this setting, that insight could have theoretical and practical implications. First, it is a contribution to our understanding of eye movements of developing readers, which is still limited. The gap in the literature concerns developing readers in any context, and data from naturalistic settings is scarce. Few, if any, studies have investigated to what extent early eye movements during reading predict later word reading ability in children in this context. Ultimately, eye movement analysis could inform the identification of children needing extra support in early reading development. Encumbered word level processes can have implications for reading comprehension and the importance of early identification of children with poor word reading and poor predicted development cannot be overstated (Rayner et al., 2009; Perfetti, 2010). Further, the benefits of predicting word reading ability extend to a wider perspective, as improved assessment could assist specialized support also for proficient readers.

This study springs from a research project developing an eye-tracking and AI-based screening method for reading difficulties (RD) described by Ygge et al. (2016), intended for use in the school environment. All participants were recorded in their schools. Eye movement analysis was based on normal text reading and a word reading assessment was carried out to establish their reading proficiency. Thus, the present analysis examines data collected during this large-scale, data driven research project ($n=2,687$) which provides the methodological context of this paper (Strandberg et al., 2022).

In this study, we analyze to what extent children's eye movements during reading predict word reading ability tested the same year and compare the results to predictions of word reading skill one year later. To our best knowledge, it has not been done before with this approach and in this setting. It seems plausible that eye movements will explain some of the variance in word reading ability measured at the same occasion. It is more uncertain, however, whether it is possible to make long-term predictions about word reading ability, such as one year later. Presumably, the amount of variance explained by the eye movement measures decreases the longer the time that has passed between the time of recording and the estimation of word reading ability, although we currently do not know by how much. Therefore, it is of interest to compare the variance explained the first year to the second occasion one year later, in order to get a quantitative estimation of the predictive validity of eye movements. Due to a lack of studies of immediate relevance, it is unfeasible to build supported hypotheses based on findings directly linked to our questions.

However, previous results from this field lead us to anticipate a bigger difference in explained variance between the concurrent and longitudinal model in the group of youngest participants (i.e., those who were in first grade at the time of eye movement recording; referred to as sample A and B in the analysis and onward, for the sake of clarity). That means, we expect the variance in word reading ability that can be ascribed to the eye movement model to decrease more longitudinally among the younger readers. Primarily, because we anticipate the eye movement behavior of this group to be influenced by a transition from sub-lexical to lexical processing, which is known to affect the eye movement characteristics during reading (Reichle et al., 2013b). Also, we expect higher variability among the eye movement measures in first grade students and because the youngest students may be more susceptible to other factors that influence the reliability of the eye movement recording and/or word reading assessment. Finally, the duration of their access to sufficient instruction is more uncertain (whereas the second grade students have had at least one year of formal instruction, which we deem enough to establish the basic concepts of word reading, at least). We address the following three key questions:

(1) To what extent do eye movements recorded in first and second grade predict word reading ability at the time of recording and 1 year later? (2), Does the amount of explained variance in word reading ability decrease in the longitudinal model compared to predictions made the same year and if so, by how much? (3) Is the decrease larger in either sample?

2. Methods

2.1. Participants

The participants were elementary school students in Järfälla municipality in Stockholm, Sweden (see Table 1 for detailed information). They all went to public schools (all municipally governed elementary schools were enrolled in the study). The yearly income of Järfälla's population was slightly higher than the Swedish median. The unemployment rate was lower than the Swedish median percentage (SCB, 2023). The participants were singled out from the larger sample and included in the present analysis on the basis of having completed all tasks of the word reading assessment and having participated twice with a one-year-interval, either in first and second grade or second and third grade (2015 and 2016).

No children with intellectual disability ($IQ < 70$) or general learning disabilities were part of the sample, as they attend classes with

TABLE 1 Participant information.

	Sample A	Sample B
Girls ($n=$)	84	97
Boys ($n=$)	79	107
N/A	1	1
Total ($N=$)	164	205
Age M1 m (sd)	7.9	8.8
Age M2 m (sd)	8.8	9.8

M1 refers to the first cycle of data collection (spring semester of 2015) and M2 to the second cycle of data collection (spring semester of 2016).

TABLE 2 Details about language, background and Caregivers' educational level among elementary school students in Järfälla.

	% of all students		
	Swedish L2	FB	HE
Trosa	4.7	12	51
Järfälla	11.2	34	61
Sweden	10.6	24	57

Swedish L2, Swedish as a second language; FB, Foreign background (the participant was born abroad or was born in Sweden to caregivers born abroad); HE, higher education among caregivers (completed third cycle education or higher).

TABLE 3 Results on the National Tests (reading subtests only) in Swedish among third graders 2016.

	Fiction	Fact	Oral	Discourse
Trosa	91.9	93.3	93.2	96.2
Järfälla	94.9	94.9	96.6	97.3
Sweden	92.5	93.8	96.1	97.2

Percentage of participants whose test results corresponded to curriculum learning goals. Fiction, Silent reading of fiction; Fact, Silent reading of fact based text; Oral, Oral reading; Discourse, Verbal text discourse.

adjusted curriculum and syllabus. Readers with Swedish as first and second language cannot be differentiated in the current study. Public records indicate that the proportion of students with Swedish as a second language was marginally larger in Järfälla public elementary schools than the Swedish average (SCB, 2023; see Table 2 for more information).

The national tests are compulsory exams which Swedish children take in third, sixth and ninth grade of elementary and middle school, and again in high school. The purpose of the tests is to support teachers' assessments of students in certain subjects in accordance to goals stated in the national curricula. Information gleaned from the test results is also used for comparison of the relative attainment of students in different schools and areas. Downloadable information in English about the national test is available on the Swedish National Agency of Education's website (Skolverket, 2023). Indicative of the general reading performance among the third grade students in this study, Table 3 shows the third grade students' 2016 results on the reading subtests of the national tests in Järfälla, as well as the national average.

2.2. Ethical considerations

The research protocol of the current study is approved by the Central Ethical Review Board (Ö 13/2015). Prior to data collection, the supervising researchers (Mattias Nilsson and Gustaf Öqvist Seimyr) met with the principals, teachers and caregivers involved at parent-teacher meetings and informed them about the study. Further information and forms for written consent were sent to the residence of all caregivers. All children enrolled in first, second and third grade were offered to participate, provided that parental consent was acquired. Moreover, their personal consent was obtained at the time of testing, and they were informed they may terminate their participation at any moment. No formal exclusion criteria were applied, but all participants were enrolled in schools with standard curriculum and syllabus, i.e., no students in special educations programs participated in the study.

2.3. Apparatus and procedure

In order to maintain authentic conditions for data collection, the testing and recording were carried out in the child's every-day environment. The participants' eye movements were recorded during normal text passage reading in an unrestrained, binocular eye tracking set-up. Thus, no devices that limit head movements were used.

The participants completed five different tasks: Word Chains, Rapid Automatized naming (RAN), word reading, pseudo word reading, and text reading (the two text passages constituting the text reading assignment are considered as part of one and the same task). Eye movements were not recorded during the first assignment (Word Chains), which was conducted in the classroom in accordance with the instruction handbook. The remainder of the assignments (RAN, word reading, pseudo word reading and text reading) was completed individually, together with an experiment leader in a separate room. The participants were called out one by one from their classroom activity. In the adjacent room where the recording took place, they were seated at approximately 70 centimeters from the monitor. Before the start of the experiment, a five-point calibration procedure was performed on the monitor of a Tobii T120 eye tracker (120 Hz; Tobii Technology AB, Danderyd, Sweden) with a 1280*1024 screen resolution. The children were given step-by-step information in the beginning assignment and throughout the assignment at relevant times. The personal consent of each participating child was obtained. The tasks were presented sequentially and run with the software Optoscope (version 3.0.0.19). Participants could terminate their participation at any time. Where applicable, data recorded before this point was saved for analysis. Note that while eye movements were recorded during each task (except Word Chains) only eye movements during text reading are analyzed in the present paper. Sound was recorded with an external USB microphone (Samsung GoMic). All tests (RAN, word reading, pseudo word reading, text reading) were scored and analyzed as described below and constituted parts of the word reading assessment.

2.4. Stimulus material

The word reading assessment in its entirety is found in [Supplementary Figures S1–S10](#). Below, the word reading assessment is described in chronological order (e.g., the first task was Word Chains and so forth).

2.4.1. Word chains

Word Chains is a standardized test used to assess children's visual word recognition. It requires the participant to separate words from each other when presented in series of strings (i.e., written consecutively, without blank spaces). The test involves 80 three-word-strings. The word strings consist of nouns, adjectives and verbs intermingled and the task is to mark the boundaries between words with a pencil. The participants were instructed to work through as many strings as possible within the time limit of 2 min. This assignment was carried out in the class room setting and is a part of a special edition of Läskedjor-2 (Jacobson 2015).

TABLE 4 Text passage metrics.

Grade	W (n)	WL (m)	LW (n)	S (n)	SL (m)	% long words	LIX	TTR %	OVIX	OVR %
1st										
Text A	20	3.6	0	5	4	0	4	80	41.7	92.55
Text B	20	3.6	0	5	4	0	4	80	41.7	92.55
2nd										
Text A	42	3.8	1	7	6	2.38	8	76.19	53.22	92.72
Text B	42	3.8	1	7	6	2.38	8	76.19	53.22	92.72
3rd										
Text A	60	3.85	6	11	5.45	10	15	68.33	46.04	90.7
Text B	60	3.85	6	11	5.45	10	15	68.33	46.04	90.7

W, words; WL, word length, (n) letters/(n) words; LW, long words > 6 letters; S, sentences; SL, sentence length (n) words/(n) sentences; % of long words, long words/words*100; LIX, SL + percent long words; TTR, type/token ratio; OVIX, $\log(\text{tokens})/\log[2\text{-}\log(\text{types})/\log(\text{tokens})]$; OVR, $\log(\text{types})/\log(\text{tokens})$.

LIX is an abbreviation of läsbarhetsindex in Swedish, which translates to readability index in English. Type/token ratio is a measure of lexical density. OVIX is an abbreviation of ordvariationsindex in Swedish which takes consideration of total text length and translates to word variation index in English. OVR is an abbreviation of ordvariationsratio in Swedish, which translates to word variation ratio.

2.4.2. Rapid automatized naming

This and the subsequent tasks (described below) were presented on the Tobii T120 screen. The alphabetical, serial format rapid naming task was based on the Comprehensive Test of Phonological Processing (CTOPP-II; Wagner et al., 2013) which taps into the reader's processing speed and fluency. It consists of 4 rows of 9 letters each, in black on white background. Participants were to name the presented letters consecutively as fast as possible. Due to time limitations, only a letter naming task was included in the present study. The responses were recorded with an external USB-microphone (Samson, Go Mic). The total naming time was recorded and reanalyzed in terms of named letters per minute.

2.4.3. Word and pseudo word reading

The tests were not part of any standardized test protocol, but were developed in accordance with the principles of Test of Word Reading Efficiency (TOWRE; Torgesen et al., 2012) which is a word reading test in English. The outcomes were intended to reflect word recognition accuracy and phonological decoding. 64 words were presented in rows of 8×8 in black on white background. They were ordered by increasing length and decreasing frequency, starting with some of the most common two-letter words and ending with nine-letter words. The pseudo words were presented in the same manner, ordered by increasing length. The pseudo words were constructed based on the real words according to a system developed by Gustaf Öqvist Seimyr. The objective was to create pseudo words that were similar to real words in terms of phonetic complexity. The consonants n, l and m were replaced by l, m and n, respectively. Vowels were replaced in according to place of articulation: (a->o, o->å, å->u, u->a), and (e->i, i->y, y->ö, ö->ä, ä->e). This procedure was repeated if a word was still semantically coherent after the first transposition. The participants were instructed to read as many words/pseudo words out loud as possible. Their responses were recorded and the number of accurate items within the time limit of 30 s per test was noted.

2.4.4. Text reading

The participants were asked to read two short fictional texts in Swedish. The texts consisted of one-passage paragraphs presented in black text on white background. We wanted to incite the gaze behavior

associated with a natural reading situation. Therefore we considered it suitable to adapt the level of difficulty to what the participants normally would read. This has implications for group comparisons. However, by varying the stimuli we avoid unbalanced results which using identical text passages for all students would lead to (disproportionate processing difficulties for participants in first grade or ceiling effects among participants in third grade). In collaboration with a special education teacher, the authors developed three sets of texts adapted to grade and presumed approximate reading ability, guided by the amount of reading instruction received. Children start receiving formal training in basic skills necessary for reading in the last year of preschool (at approximately 6 years of age). Despite this, there is great variation in reading and writing competence among children who start first grade. In Strandberg et al., 2022 this text reading task (as well as the remainder of the word reading assignment) was used in a sample of 2,876 children and descriptive data, correlation coefficients and linear mixed models are reported. The text passages are described in further detail in Table 4. The texts all dealt with everyday topics. The participants were informed on beforehand that one question about the content would follow upon them finishing each text. The questions were of simple yes- or no character and the answer was evident in the texts (no inferences were required on the student's behalf). The purpose of the questions was to encourage attentive reading. The participants' answers did not provide information pertinent to assess reading comprehension. Thus, their answers are not analyzed further in the present study. The participants were encouraged to read silently if possible, but reading aloud was allowed when requested by the child. Reading speed was measured by the number of read words per minute and the result of both texts was averaged. The eye movement measures in this study are based on this tasks.

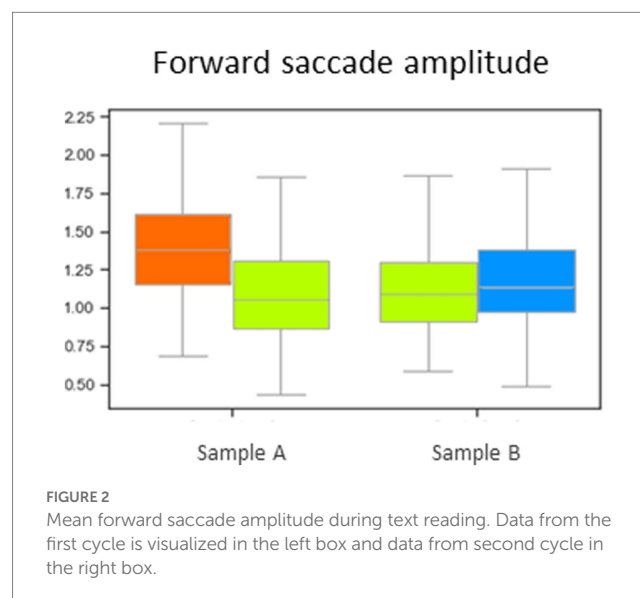
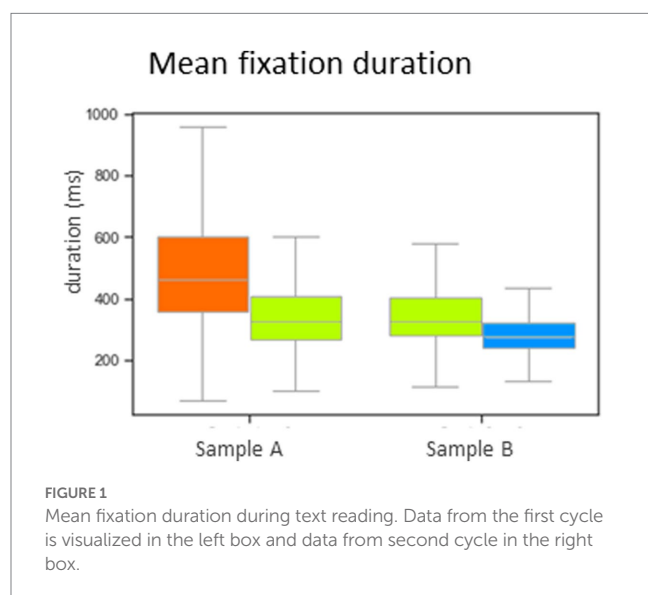
2.5. Analysis

Eye movements were recorded and analyzed binocularly. The eye movement data was analyzed in Optosphere (Version 2.1) (Nilsson Benfatto et al., 2016). The software classifies the eye movements as events, depending on their properties. A fixation was defined as the event of the eye remaining within an area corresponding to the fovea

TABLE 5 Descriptive statistics of eye movement variables.

	Sample A, m (sd)	Sample B, m (sd)
	1	2
Mean fixation duration (milliseconds)	519 (185)	349 (102)
Forward saccade amplitude (degrees)	0.74 (0.19)	0.95 (0.26)
Proportion of regressions (ratio)	0.33 (0.05)	0.21 (0.06)

The variables are based on eye movement recordings made in first grade in Sample A and second grade in group Sample B (e.g., recorded during first cycle of data collection). Correlations between eye movement variables in each sample.

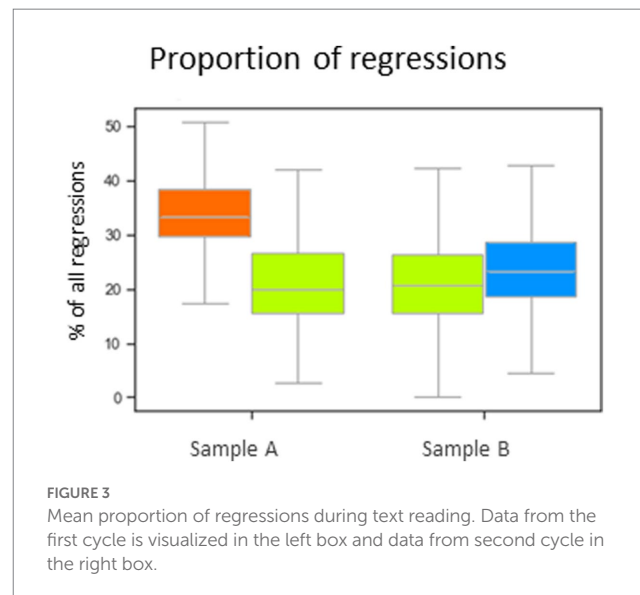


for at least 50 ms. Saccades were classified according to whether they occur within or outside the perceptual span and are identified as progressive (45–224 degrees) or regressive (224–45 degrees) depending on their directional angle. Fixations were analyzed in terms of mean fixation duration. Saccades were analyzed according to directional angle, amplitude and proportion of regressive saccades (Ygge et al., 2016). See Table 5 and Figures 1–3 for the descriptive statistics and distribution of eye movement data and Table 6 for the descriptive statistics of the word reading assessment.

The sample was separated in two groups based on which grade the participant attended. Throughout the remainder of this paper, we refer to the group of children who participated in first and again in second grade as sample A, while the group of children who participated in second and third grade is referred to as sample B. At the outset of data analysis, the entire sample consisted of 561 participants. The data was filtered according to a three-step procedure. Eye movement recordings were removed if the amount of noise exceeded 25% ($n = 178$). Next, subjects were excluded based on whether the difference in reading speed (WPM) between the two text conditions exceeded the outlier threshold values ($n = 8$). Outlier values in the eye movement data were then identified with Tukey's interquartile range with the scale factor set to 3.0. Following the complete filtering procedure, 164 participants remained in sample A and 206 participants remained in sample B.

2.5.1. Outcome variable

To generate a representative measurement of reading ability, we combined the outcomes of the reading assessment in one,



composite variable. We standardized the results through z-transformation and combined them using simple averaging. The tests are described under Materials, above.

We carried out this analysis based on the assumption that the tests (1) assess important constructs of reading ability, (2) to some extent are conceptually linked and (3) provide a global representation of reading ability. Moreover, they are relevant constituents of reading assessments. Our intention was to facilitate interpretation of results pertaining to multiple variables

TABLE 6 Descriptive statistics of word reading assessment.

Grade	Sample A, m (sd)		Sample B, m (sd)	
	1	2	2	3
Rapid naming	79 (18)	100 (20)	97 (19)	112 (20)
Word chains	10 (6)	18 (7)	17 (6)	24 (7)
Pseudo word reading	16 (7)	23 (6)	22 (6)	26 (6)
Word reading	28 (11)	39 (10)	39 (10)	46 (9)
Reading speed	56 (29)	104 (39)	103 (40)	139 (40)

Rapid naming is estimated in letters per second and reading speed is estimated in words per minute (during text reading) while the remaining tests are evaluated in number of correct responses.

TABLE 7 Intercorrelations word reading assessment in first grade (sample A).

	1	2	3	4	5
1. Rapid naming		0.37***	0.43***	0.52***	0.51***
2. Word chains			0.58***	0.65***	0.72***
3. Pseudo word reading				0.86***	0.75***
4. Word reading					0.88***
5. Reading speed					

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 8 Intercorrelations word reading assessment in second grade (sample A).

	1	2	3	4	5
1. Rapid naming		0.22**	0.45***	0.55***	0.43***
2. Word chains			0.49***	0.60***	0.58***
3. Pseudo word reading				0.77***	0.55***
4. Word reading					0.77***
5. Reading speed					

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

and to generate a robust, conceptually valid measure. The interaction of eye movement variables and each of the reading tests was analyzed (see [Supplementary Table S1](#)). The main finding of this prelude, exploratory analysis is that the predictor variables account for less variance in three of the outcome variables (RAN, word chains, and pseudo word reading) when those tests are handled separately. The predictor variables account for slightly less of the variance in word reading accuracy and approximately the same amount of variance in reading speed as in the composite measure. Largely, these findings support the integrity of the composite measure. They indicate that combining these outcome variables does not compromise nor exaggerate the predictive validity of the model. There are no notable differences between the individual tests and the composite measure in terms of longitudinal differences.

The variance inflation factor (VIF) was calculated in order to investigate the influence of multicollinearity among the predictor variables. The variance inflation factor was < 2.5 for all predictors in both samples and multicollinearity should therefore only affect the variance of the regression coefficients to a limited extent ([Thompson et al., 2017](#)).

A multiple linear regression was calculated to predict and account for the variance in the composite word reading ability score (OV) based on eye movement measures (PV). Eye movements recorded in first grade were used as predictive variables for word reading ability assessed the same year (i.e., assessed at the same occasion as eye movement recording took place) and 1 year post-recording (i.e., when the students were in second grade). The same analysis was carried out in sample B - eye movements recorded in second grade were used as predictive variables for word reading ability assessed the same year and 1 year post-recording.

3. Results

Inter-/correlations of the word reading assessment (including the composite measure) are described in [Tables 7–12](#).

Below, we introduce the models of eye movement variables predicting word reading ability recorded at the same occasion. Second, we address predictions of word reading ability 1 year post eye movement recordings.

TABLE 9 Intercorrelations word reading assessment in second grade (sample B).

	1	2	3	4	5
1. Rapid naming		0.37***	0.51***	0.59***	0.49***
2. Word chains			0.58***	0.66***	0.68***
3. Pseudo word reading				0.82***	0.7***
4. Word reading					0.81***
5. Reading speed					

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 10 Intercorrelations word reading assessment in third grade (sample B).

	1	2	3	4	5
1. Rapid naming		0.31***	0.42***	0.54***	0.35***
2. Word chains			0.48***	0.54***	0.59***
3. Pseudo word reading				0.76***	0.60***
4. Word reading					0.68***
5. Reading speed					

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 11 Correlations word reading assessment in first x second grade (sample A).

	1	2	3	4	5	6
1. Rapid naming	0.58***	0.38***	0.31***	0.43***	0.39***	–
2. Word chains		0.74***	0.38***	0.51***	0.56***	–
3. Pseudo word reading			0.72***	0.73***	0.61***	–
4. Word reading				0.8***	0.74***	–
5. Reading speed					0.79***	–
6. Composite variable						0.85***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 12 Correlations word reading assessment in second x third grade (sample B).

	1	2	3	4	5	6
1. Rapid naming	0.61***	0.42***	0.45***	0.57***	0.43***	
2. Word chains		0.79***	0.47***	0.51***	0.61***	
3. Pseudo word reading			0.82***	0.73***	0.59***	
4. Word reading				0.82***	0.75***	
5. Reading speed					0.79***	
6. Composite variable						0.88***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.1. Concurrent analysis: predictions of word reading ability based on eye movements recorded the same year

3.1.1. Sample A

A significant regression equation was found [$F(3, 160) = 80.64$, $p < 0.001$]. Eye movements recorded in first grade accounted for 59% of the variance in word reading ability score in first grade ($R^2 = 0.6$, $R^2_{Adjusted} = 0.59$). Mean fixation duration and forward saccade amplitude were significant predictors of word reading ability score. The model shows that an increase of mean fixation duration and proportion of regression is associated to a decrease in outcome variable, while an

increase of saccade amplitude is linked to a higher word reading ability score (see Table 13).

3.1.2. Sample B

A significant regression equation was found [$F(3, 202) = 107$, $p < 0.001$]. Eye movements recorded in second grade accounted for 61% of the variance in word reading ability score in second grade ($R^2 = 0.61$, $R^2_{Adjusted} = 0.61$). Each variable in the model was a significant predictor of word reading ability score. The model shows that an increase of mean fixation duration and proportion of regression is associated to a decrease in outcome variable, while an increase of saccade amplitude is linked to a higher word reading ability score (see Table 14).

TABLE 13 Summary of multiple regression analysis: reading ability in first grade predicted by eye movements variables in first grade.

	Estimate	Std. error	t-value	P (> t)
(Intercept)	115.01	13.02	8.83	1.74e-15 ***
Mean fixation duration	−0.10	0.01	−10.72	< 2e-16 ***
Forward saccade amplitude	29.13	9.45	3.08	0.00242 **
Proportion of regressions	−82.33	29.63	−2.78	0.00612 **

Outcome variable: Composite measure of reading assessment outcomes in first grade. Multiple R-squared: 0.69. Adjusted R-squared: 0.59. $p < 0.001$. Significance codes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 14 Summary of multiple regression analysis: reading ability in second grade predicted by eye movement variables in second grade.

	Estimate	Std. error	t-value	P (> t)
(Intercept)	148.82	14.14	10.52	< 2e-16***
Mean fixation duration	−0.2	0.02	−9.73	< 2e-16***
Forward saccade amplitude	57.36	8.49	6.76	1.47e-10***
Proportion of regressions	−150.11	29.13	−5.15	6.09e-07***

Note. Outcome variable: composite measure of reading assessment outcomes in second grade. Multiple R-squared: 0.61. Adjusted R-squared: 0.61. $p < 0.001$. Significance codes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 15 Summary of multiple regression analysis: reading ability in second grade predicted by eye movement variables in first grade.

	Estimate	Std. Error	t-value	P (> t)
(Intercept)	161.93	20.56	7.88	4.86e-13***
Mean fixation duration	−0.11	0.01	−7.64	1.84e-12***
Forward saccade amplitude	38.47	14.92	2.6	0.0108*
Proportion of regressions	−82.61	46.78	−1.76	0.0793

Outcome variable: composite measure of reading assessment outcomes in second grade. Predictor variables are based on recordings from first grade. Multiple R-squared: 0.45. Adjusted R-squared: 0.44. $p < 0.001$. Significance codes: $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 16 Summary of multiple regression analysis: word reading ability in third grade predicted by eye movement variables in second grade.

	Estimate	Std. Error	t-value	P (> t)
(Intercept)	175.24	16.38	10.7	< 2e-16***
Mean fixation duration	−0.19	0.02	−7.86	2.28e-13 ***
Forward saccade amplitude	43.9	9.83	4.5	1.33e-05 ***
Proportion of regressions	−100.76	33.75	−2.96	0.00318 **

Outcome variable: composite measure of reading assessment outcomes in third grade. Predictor variables are based on recordings from second grade. Multiple R-squared: 0.48. Adjusted R-squared: 0.48. $p < 0.01$. Significance codes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.2. Longitudinal analysis: predictions of word reading ability based on eye movements recorded one year earlier

3.2.1. Sample A

A significant regression equation was found [$F(3, 160) = 43.41$, $p < 0.001$]. Eye movements recorded in first grade accounted for 44% of the variance in word reading ability score among second grade students ($R^2 = 0.45$, $R^2_{Adjusted} = 0.44$). Mean fixation duration and forward saccade amplitude were individually significant predictors of word reading ability score. The model shows that an increase of mean fixation duration and proportion of regression is associated to a decrease in outcome variable, while an increase of

saccade amplitude is linked to a higher word reading ability score (see Table 15).

3.2.2. Sample B

A significant regression equation was found [$F(3, 202) = 63.23$, $p < 0.001$]. Eye movements recorded in second grade accounted for 48% of the variance in word reading ability score in third grade ($R^2 = 0.49$, $R^2_{Adjusted} = 0.48$). All eye movement measures in the model were significant predictors of word reading ability. The model shows that an increase of mean fixation duration and proportion of regression is associated to a decrease in outcome variable, while an increase of saccade amplitude is linked to a higher word reading ability score (see Table 16).

Eye movements recorded in first grade predicted 59% of the variance in the children's word reading ability the same year, while it accounted for 44% of the variance one year later. In sample B, eye movements recorded in second grade explained 61% of the variance in word reading ability measured the same year, and 48% one year later. Thus, the explained variance decreases with 15 percentage points between first and second grade, and with 13 percentage points between second and third grade. Further analysis showed that r^2 is substantially dependent on mean fixation duration in each instance while forward saccade amplitude and proportion of regressions contribute to the models to a lesser extent. As a unique predictor, mean fixation duration accounts for on average 53% of the variance in concurrent word reading ability and 42% on year post-recording. Thus, the exclusion of the forward saccade amplitude and proportion of regression two did influence r^2 negatively, albeit to a small extent. The initial model including all three variables was deemed to best predict word reading ability score.

The texts which constituted the text reading task were adapted to an approximated level of reading proficiency associated with each grade (e.g., participant age). Consequently, the participants read different texts in the first vs. second cycle of data collection. In terms of the present investigation, the eye movement variables and reading speed (WPM, which was entered in the composite variable) are based on the same text in the analysis of concurrent word reading ability. Exclusion of reading speed from the composite measure would eliminate the risk of interdependence between those variables. Theoretically, however, it would make for a less sound representation of natural reading. It would narrow the scope of the composite variable to a representation mainly of lexical level word reading (of word lists), which is unintuitive from a logical and conceptual standpoint. The objective of this paper is to investigate to what extent eye movements account for variation in word reading ability, and reading speed is typically indicative of proficiency (Sovik et al., 2000). We ran an additional analysis to determine the effects of exclusion of reading speed from the composite variable. The exclusion of reading speed affected r^2 in such a way that eye movements were on average 9 percentage points less predictive of word reading ability ($R^2_{\text{Adjusted}} = 0.53$ in sample A and 0.49 in sample B). As participants had graduated and enrolled in the subsequent grade by the time of the second cycle of data collection, the overlap does not exist in the longitudinal models. Thus, based on an unrelated text, eye movements accounted for on average 48% of the variation in word reading ability measured 1 year later. Based on conceptual assumptions and the fact that there only was a nine percentage point difference related to the inclusion or exclusion of reading speed in the concurrent model, we opted to incorporate it as a measure in the composite variable of word reading ability.

4. Discussion

The objective of this study was to investigate to what extent eye movements during reading predict later word reading ability in young readers. We were interested in examining this relationship in a large and naturally diverse sample under as ecological circumstances as possible. With this in mind, we analyzed the extent to which three basic eye movement measures predicted the variance in word reading ability, both concurrently and longitudinally.

4.1. General discussion of descriptive eye movement findings

A pattern of quite robust associations between the composite variable and the eye movement variables emerge in our data. In summary, an increase of the composite variable was associated to shorter fixations, longer forward saccades, and fewer regressions during reading. This means that a higher result on the variable intended to reflect word reading ability was linked to changes in eye movements that are consistent with improved text processing (Blythe, 2014).

Turning to the relationship between the eye movement variables in this study and those reported in earlier work, we found some variation that appears to highlight the difference of approach between our study and the lion's share of past research. Compared to several prior studies (Rayner, 1986; Blythe et al., 2006, 2009; Reichle et al., 2013; Rau et al., 2015b) mean fixation duration was higher in both sample A and sample B. Increased variation among variables can be a product of a naturalistic set-up and diverse sample of participants. In this case, children with Swedish as a second language and children with rudimentary reading skill were included. Additionally, we collected the data at the location of the schools, to fulfill the intention to record the students as they normally would read. Finally, tasks with unlimited time and text-level stimuli, by their nature, involve more freedom in when and how to move the eyes than time-limited reading of, for example, single sentences or words. Typically, forward saccade length increases with age and reading development (Rayner, 1986; Blythe, 2014). In this case, forward saccade length was 0.74 degrees in first grade and 0.95 degrees in second grade (i.e., in sample A and B, respectively). This leads us to believe that our text passages were appropriately adapted to the participants' age and presumptive reading proficiency. The first grade students read texts containing line-breaks at the end of each sentence, whereas sentences appeared consecutively in the texts for second grade students. Words and sentences were longer and semantically more advanced for the children in second grade. We were careful so as not to exaggerate these factors, as we wanted the children to be able to process the texts, given that their reading ability was age adequate. Proportion of regression decreased, as expected according to the literature, although the prevalence and role of regressions vary depending on the context. Characteristics of the supposed incongruence that the reader experiences as well as the orthographic transparency influence if and how they reread a part of the text (Rau et al., 2015a; Inhoff et al., 2019). The significance of regressive saccades in this analysis is further discussed on page 17. Structured control conditions could reveal more about the influence of these and similar stimuli characteristics in future work.

4.2. Concurrent predictions of word reading ability

Eye movements accounted for more or less the same amount of variance in concurrent word reading ability in sample A and sample B (i.e., in first compared to second grade; the difference was 1 percentage point). Overall, average fixation duration was

the most important predictor of later and concurrent word reading ability. Saccade amplitude and proportion of regressions contributed to the models but to a lesser extent. The average proportion of regressions generally was within the expected range, but was not a significant predictor of longitudinal word reading ability in the youngest readers. This suggests that early on, when children often read out each word slowly in turn, the number of regressive saccades is not informative, or at least not significantly so, in terms of word reading ability. Further, it is important to consider the qualities of the variable. In this analysis, proportion of regressions was based on the recording as a whole and does not take details of the context in which they occur into account. The proportion of regressive saccades reveals nothing of the length of the regression or if they occur within or across several words. The properties of the word that the reader rereads are important to the interpretation of why the reader moves the gaze backwards (Inhoff et al., 2019). Word-based eye movement measures could enable a deeper look at that issue in a future analysis. Speculatively, the fact that the first grade students (mean age = 7.85) are in an early stage of reading acquisition plays a role. Their eye movement patterns (short forward saccades) suggest that they to some extent rely on small-unit decoding, which may lessen the need to reread text. Previous studies found that children reading in an orthographically opaque language (English) have higher rereading times than children reading in a transparent language (Rau et al., 2015b). Young readers of transparent orthographies tend to depend on sub-lexical processes for word recognition which may limit the need and prevalence of regressions. Regressions may therefore be of limited value to the assessment of word reading ability in very young readers, as small-unit decoding is usually reliable in a somewhat transparent language such as Swedish. Our results show that while the participants in first grade did make regressions to an expected extent, they did not account for variance in word reading ability to any significant degree. As small unit decoding is reliable, processing difficulties associated to RD might be reflected in variations in other variables in the youngest readers. Future work could further investigate what eye movement characteristics that differentiate students with RD from students who are slow to acquire the first steps of reading but over time progress to the level of readers with typical reading development.

4.3. Longitudinal predictions of word reading ability

We expected the eye movement variables to predict variance in word reading ability score to a lesser extent 1 year post-recording, compared to the models of eye movements and concurrent word reading ability. In line with that assumption, $R^2_{Adjusted}$ was lower in the longitudinal predictions in both samples of students. While the model predicts roughly 60% of the variance in concurrent word reading ability, the explained variance is on average 14 percentage points lower 1 year later. This means that eye movements accounted for less of the variation in word reading ability 1 year after the recording, however, we consider the decrease small seeing as it is based on an unrelated text and

a year has passed between predictions. We consider this finding as further indication (see page 15) that the predictions were not unduly biased by the fact that reading speed and eye movement measures were recorded from the same text stimuli in concurrent analysis. The fact that the decrease in explained variance is small when the predictor variables are based on an unrelated text speaks to their validity as reflections of global reading ability rather than to text-specific characteristics.

Considering the relation between the samples of participants, we found that the difference between the concurrent and longitudinal models are greatest in sample A (that is, in the group of students whose eye movements were recorded in first grade). Similarly, there is a somewhat greater difference in word reading assessment outcomes between first and second grade, in relation to second and third grade. Assuming that there is a tight interaction between word reading and lexical processing, this pattern can be mapped onto a shift of decoding strategy, and its' interplay with sub-lexical and lexical processes. As the young readers begin to integrate lexically based processes in their reading, eye movements in first grade may be reflective of a decoding strategy which is less used by second grade.

In future work it would be interesting to consider the influence of reading proficiency on predictions of word reading ability based on eye movements. In this area, interesting work has been carried out by Vorstius et al. (2014) who found that the relationship between eye movement variables and reading ability was weaker among poor readers, especially those with results below the 0.3 quantile. The predictability of word reading ability may decrease at different rates across individuals, why it would be valuable to compare the predicted values to the actual outcomes in children and analyze possible interactions with reading proficiency.

There are some implications worth noting with regards to the present study. Pursuant to our objectives, we analyzed data recorded in an authentic environment where simplicity and ecological validity was prioritized. This was part of an explorative approach which allowed a high number of students to participate. Further, we chose an inclusive approach regarding the sample because we wanted collect data from a typical group of elementary school students, aware that it has implications for the outcomes and conclusions in this study. Indeed, the results reflect variation among the variables that is consistent with limited experimental control and the findings should be considered with this in mind. There are several interesting variables that could be taken into account regarding the participants for a fine-grained analysis of the relationships reported here, such as mother tongue and presence of specific learning difficulties (such as RD). Further, both the model and materials for assessments should be developed for more precise predictions. In future work, additional data points would enable valid growth modeling. It would be interesting to include word-based eye movements, which we think could improve the models while also inform assessments about variations in underlying processes during reading. For educational purposes, both basic and word-based eye movements could be used to infer precise information about strengths and weaknesses in a child's reading. As an example, information about word recognition efficacy can be inferred from the relationship between fixation durations and word frequency.

Small unit decoding, which can encumber fluent reading, could be revealed through analysis of the saccade pattern and also at what level of complexity that the reader resorts to this strategy.

Nevertheless, the data in the present paper can reproduce previously well-known findings and relationships between eye movements, word reading ability and age differences. Here, we have shown that eye movements collected in a natural school setting with limited experimental control can account for a large portion of variance in word reading ability, not only concurrently, but also 1 year after the recordings were made. This suggests that eye tracking could be a potentially useful and efficient complementary technique for monitoring and predicting children's word reading development during the early school years. Children with a predicted poor developmental reading trajectory might then be given professional support early on and minimize the effect on school performance, motivation, and self-esteem.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Central Ethical Review Board (Ö 13/2015). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AS, MN, and GÖ contributed to the conception and design of the study, and performed the word reading assessment and eye tracking recordings. MN and GÖ processed the eye movement data. AS and MN performed the statistical analysis and wrote the draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

References

- Al Dahhan, N., Georgiou, G. K., and Hung, R. (2014). Eye movements of university students with and without reading difficulties during naming speed tasks. *Ann. Dyslexia*. 64:137–150. doi: 10.1007/s11881-013-0090-z
- Barnes, A. E., and Kim, Y. S. (2016). Low-skilled adult readers look like typically developing child readers: a comparison of reading skills and eye movement behavior. *Read. Writ.* 29, 1889–1914. doi: 10.1007/s11145-016-9657-5
- Beck, I., Perfetti, C., and McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *J. Educ. Psychol.* 74:506. doi: 10.1037/0022-0663.74.4.506
- Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Curr. Dir. Psychol. Sci.* 23, 201–207. doi: 10.1177/0963721414530145
- Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., and Hyönä, J. (2011). Reading disappearing text: why do children refixate words? *Vis. Res.* 51, 84–92. doi: 10.1016/j.visres.2010.10.003
- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., and Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vis. Res.* 46, 3898–3908. doi: 10.1016/j.visres.2006.06.006
- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., and Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vis. Res.* 49, 1583–1591. doi: 10.1016/j.visres.2009.03.015
- Blythe, H. I., Pagán, A., and Dodd, M. (2015). Beyond decoding: phonological processing during silent reading in beginning readers. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1244–1252. doi: 10.1037/xlm0000080
- Carter, B. T., and Luke, S. G. (2020). Best practices in eye tracking research. *Int. J. Psychophysiol.* 155, 49–62. doi: 10.1016/j.jpsycho.2020.05.010
- Clifton, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., et al. (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *J. Mem. Lang.* 86, 1–19. doi: 10.1016/j.jml.2015.07.004

Funding

This work was supported by the Sweden's innovation agency, VINNOVA, under Grant 2014–03459; the Sigvard and Marianne Bernadotte Research Foundation for Children's Eye Care and Ulla and Ingemar Dahlberg's Foundation.

Acknowledgments

The authors express their gratitude to the Sigvard and Marianne Bernadotte Research Foundation for Children's Eye Care, Ulla and Ingemar Dahlberg's Foundation, VINNOVA – Sweden's Innovation Agency and Trosa and Järfälla schools and municipalities for their support and contribution to this project.

Conflict of interest

MN and GÖ own equity in Optolexia, a company whose aim is to offer new technologies for the assessment of reading deficits in school-age children.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2023.1077882/full#supplementary-material>

- Clifton, C., Staub, A., and Rayner, K. (2007). "Eye movements in reading words and sentences," in *Eye Movements: A window on mind and brain*. eds. R. P. G. Van Gompel, M. H. Fischer, W. S. Murray and R. L. Hill (Oxford: Elsevier), 341–371.
- Denton, C. A., Fletcher, J. M., Anthony, J. L., and Francis, D. J. (2006). An evaluation of intensive persistent reading difficulties. *J. Learn. Disabil.* 39, 447–466. doi: 10.1177/00222194060390050601
- Ehri, L. C. (2005). Scientific studies of reading learning to read words: theory, findings, and issues learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21. doi: 10.1080/10888438.2013.819356
- Ekstrand, G., Carin, A., Benfatto, M. N., and Seimyr, G. Ö. (2021). Screening for Reading difficulties: comparing eye tracking outcomes to neuropsychological assessments. *Front. Educ.* 6, 1–13. doi: 10.3389/feduc.2021.643232
- Gidlöf, K., Wallin, A., Dewhurst, R., and Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: gaze behaviour during decision making in a natural environment. *J. Eye Mov. Res.* 6, 1–14. doi: 10.16910/jemr.6.1.3
- Häikiö, T., Bertram, R., Hyönä, J., and Niemi, P. (2009). Development of the letter identity span in reading: evidence from the eye movement moving window paradigm. *J. Exp. Child Psychol.* 102, 167–181. doi: 10.1016/j.jecp.2008.04.002
- Henderson, J. M., Choi, W., Luke, S. G., and Desai, R. H. (2015). Neural correlates of fixation duration in natural reading: evidence from fixation-related fMRI. *Neuroimage* 119, 390–397. doi: 10.1016/j.neuroimage.2015.06.072
- Hindmarsh, G. P., Black, A. A., White, S. L. J., Hopkins, S., and Wood, J. M. (2021). Eye movement patterns and reading ability in children. *Ophthalmic Physiol. Opt.* 41, 1134–1143. doi: 10.1111/opo.12854
- Hosp, M. K., Fuchs, L. S., Hosp, M. K., and Fuchs, L. S. (2019). Using CBM as an indicator of decoding, word Reading, and comprehension: do the relations change with grade? *School Psychology Review* 34, 9–26. doi: 10.1080/02796015.2005.12086272
- Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *J. Verbal Learn. Verbal Behav.* 23, 612–624. doi: 10.1016/S0022-5371(84)90382-7
- Inhoff, A. W. (1990). Integrating information across eye fixations in reading: the role of letter and word units. *Acta Psychol.* 73, 281–297. doi: 10.1016/0001-6918(90)90027-D
- Inhoff, A. W., Kim, A., and Radach, R. (2019). "Regressions during Reading." *Vision* 3:35. doi: 10.3390/vision3030035
- Jacobson, C. (2015). *Läskedjor-2. Hogrefe Psykologiförlaget AB*.
- Joseph, H. S. S. L., Nation, K., and Livensedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *Sch. Psychol. Rev.* 42, 207–222. doi: 10.1080/02796015.2013.12087485
- Just, M. A., and Patricia, A. C. (1980). "A theory of reading: From eye fixations to comprehension," *Psychological Review* 87:329–354. doi: 10.1037/0033-295X.87.4.329
- Kamhi, A., and Catts, H. (2005). *Language and Reading Disabilities*. Boston: Pearson.
- Kim, Y. S., Grace, Y. P., and Vorstius, C. (2019). Unpacking eye movements during oral and silent reading and their relations to reading proficiency in beginning readers. *Contemp. Educ. Psychol.* 58, 102–120. doi: 10.1016/j.cedpsych.2019.03.002
- König, P. (2016). Eye movements as a window to cognitive processes. *J. Eye Mov. Res.* 1–16. doi: 10.16910/jemr.9.5.3
- Kornev, A. N., Oganov, S. R., and Galperina, E. I. (2019). Development of the psychophysiological mechanisms in the comprehension of printed texts: eye tracking during text reading in healthy and dyslexic children aged 9–11. *Hum. Physiol.* 45, 249–255. doi: 10.1134/S0362119719030083
- LaBerge, D., and Jay Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2
- Livensedge, S. P., and Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends Cogn. Sci.* 4, 6–14. doi: 10.1016/S1364-6613(99)01418-7
- McConkie, G. W., and Rayner, K. (1975). The span of the effective stimulus during a fixation in Reading. *Percept. Psychophys.* 17, 578–586. doi: 10.3758/BF03203972
- McConkie, G. W., and Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bull. Psychon. Soc.* 8, 365–368. doi: 10.3758/BF03335168
- Mcdonald, M. A., Holdsworth, S. J., and Danesh-meyer, H. V. (2015). Eye movements in mild traumatic brain injury: clinical challenges. *J. Eye Mov. Res.* 15, 1–15. doi: 10.16910/jemr.15.2.3
- Nilsson Benfatto, M., Seimyr, G. Ö., Ygge, J., Pansell, T., Rydberg, A., and Jacobsson, C. (2016). Screening for dyslexia using eye tracking during reading. *PLoS ONE* 11. doi: 10.1371/journal.pone.0165508
- Perfetti, C. A. (2010). "Decoding, vocabulary, and comprehension" in *Bringing reading research to life*. eds. M. G. McKeown and L. Kucan (New York: Guilford), 291–302.
- Perfetti, C. A., and Hart, L. (2001). The lexical basis of comprehension skill. in *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. ed. D. S. Gorfein (American Psychological Association), 67–86.
- Pollatsek, A., Reichle, E. D., and Rayner, K. (2006). Tests of the E-Z reader model: exploring the interface between cognition and eye-movement control. *Cogn. Psychol.* 52, 1–56. doi: 10.1016/j.cogpsych.2005.06.001
- Prado, C., Dubois, M., and Valdois, S. (2007). The eye movements of dyslexic children during Reading and visual search: impact of the visual attention span. *Vis. Res.* 47, 2521–2530. doi: 10.1016/j.visres.2007.06.001
- Rau, A. K., Moeller, K., and Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: an eye-tracking study. *Sci. Stud. Read.* 18, 224–233. doi: 10.1080/10888438.2013.857673
- Rau, A. K., Moll, K., Snowling, M. J., and Landerl, K. (2015a). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *J. Exp. Child Psychol.* 130, 92–105. doi: 10.1016/j.jecp.2014.09.012
- Rau, A. K., Moll, K., Snowling, M. J., and Landerl, K. (2015b). Journal of experimental child effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *J. Exp. Child Psychol.* 130, 92–105. doi: 10.1016/j.jecp.2014.09.012
- Rayner, K. (1975). The perceptual peripheral span and cues in reading. *Cogn. Psychol.* 7, 65–81. doi: 10.1016/0010-0285(75)90005-5
- Rayner, K. (1977). Visual attention in reading: eye movements reflect cognitive processes. *Mem. Cogn.* 5, 443–448. doi: 10.3758/BF03197383
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *J. Exp. Child Psychol.* 41, 211–236. doi: 10.1016/0022-0965(86)90037-8
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372
- Rayner, K., Ashby, J., Pollatsek, A., and Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: implications for the E-Z reader model. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 720–732. doi: 10.1037/0096-1523.30.4.720
- Rayner, K., Jane, A., Chace, K. H., and Slattery, T. J. (2009). "Eye Movements as Reflections of Comprehension Processes in Reading." *Scientific Studies of Reading* 10:241–255. doi: 10.1207/s1532799xssr1003
- Rayner, K., and Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. *Mem. Cogn.* 14, 191–201. doi: 10.3758/BF03197692
- Rayner, K., and McConkie, G. W. (1976). What guides a reader's eye movements? *Vis. Res.* 16, 829–837. doi: 10.1016/0042-6989(76)90143-7
- Reichle, E. D., Livensedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., et al. (2013). Using E-Z reader to examine the concurrent development of eye-movement control and reading skill. *Dev. Rev.* 33, 110–149. doi: 10.1016/j.dr.2013.03.001
- Reichle, E. D., Pollatsek, A., Fisher, D. L., and Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychol. Rev.* 105, 125–157. doi: 10.1037/0033-295X.105.1.125
- Reichle, E. D., Warren, T., and McConnell, K. (2009). Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychon. Bull. Rev.* 16, 1–21. doi: 10.3758/PBR.16.1.1
- SCB. (2023). "Kommuner i siffror." Available at: <https://kommunsiffror.scb.se/?id1=0123&id2=null>
- Seassau, M., and Bucci, M.-P. (2013). Reading and visual search: a developmental study in normal children. *PLoS One* 8:e70261. doi: 10.1371/journal.pone.0070261
- Skolverket (2023). "Statistik." Available at: <https://www.skolverket.se/skolutveckling/statistik>
- Sovik, N., Arntzen, O., and Samuelstuen, M. (2000). Eye-movement parameters and reading speed – a study of Oral and silent reading performances of twelve-year-old children. *Read. Writ.* 13, 237–255. doi: 10.1023/A:1026495716953
- Spichtig, A. N., Hiebert, E. H., Vorstius, C., Pascoe, J. P., David Pearson, P., and Radach, R. (2016). The decline of comprehension-based silent reading efficiency in the United States: a comparison of current data with performance in 1960. *Read. Res. Q.* 51, 239–259. doi: 10.1002/rrq.137
- Statens beredning för medicinsk utvärdering (SBU). (2014). *Dyslexi hos barn och ungdomar: tester och insatser; en systematisk litteraturoversikt*. Stockholm: Statens beredning för medicinsk och social utvärdering.
- Strandberg, A., Nilsson, M., Östberg, P., and Seimyr, G. Ö. (2022). Eye movements during reading and their relationship to reading assessment outcomes in Swedish elementary school children. *J. Eye Mov. Res.* 15. doi: 10.16910/JEMR.15.4.3
- Thompson, C. G., Kim, R. S., Aloe, A. M., and Becker, B. J. (2017). Extracting the variance in flation factor and other multicollinearity diagnostics from typical regression results. *Basic Appl. Soc. Psychol.* 39, 81–90. doi: 10.1080/01973533.2016.1277529
- Tiffin-Richards, S. P., and Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vis. Res.* 113, 33–43. doi: 10.1016/j.visres.2015.05.008

- Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (2012). *Test of word reading efficiency*. 2nd Edn Pro-Ed.
- Vorstius, C., Radach, R., and Lonigan, C. J. (2014). Eye movements in developing readers: a comparison of silent and ORAL sentence reading. *Vis. Cogn.* 22, 458–485. doi: 10.1080/13506285.2014.881445
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., and Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing (CTOPP-II)*. 2nd Edn. Pro-Ed.
- Yang, S. N. (2012). Effects of processing difficulty on eye movements in Reading: a review of behavioral and neural observations. *J. Eye Mov. Res.* 5:1. doi: 10.16910/jemr.5.4.1
- Ygge, J., Seimyr, G. Ö., Benfatto, M. N., Pansell, T., Jacobson, C., and Rydberg, A. (2016). Screening for dyslexia using eye tracking during reading. *PLoS One* 11:e0165508. doi: 10.1371/journal.pone.0165508
- Zhang, G., Yuan, B., Hua, H., Lou, Y., Lin, N., and Li, X. (2021). Individual differences in first-pass fixation duration in reading are related to resting-state functional connectivity. *Brain Lang.* 213:104893. doi: 10.1016/J.BANDL.2020.104893
- Zijlstra, H., Van Bergen, E., Amsterdam, V. U., Regtvoort, A., and Van Der Leij, A. (2021). Prevention of reading difficulties in children with and without familial risk: short-and long-term effects of an early intervention. *Journal of Educational Psychology* 113:248–267. doi: 10.1037/edu0000489



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Octávio Moura,
University of Coimbra, Portugal
Denis Alamargot,
Université Paris-Est Créteil Val de Marne, France

*CORRESPONDENCE

Frederick K. S. Leung
✉ frederickleung@hku.hk

RECEIVED 22 September 2022

ACCEPTED 03 July 2023

PUBLISHED 28 July 2023

CITATION

Lu H, Chen X, Leung FKS and Zuo H (2023)
Reliability, validity, and measurement invariance
of a Chinese handwriting legibility scale among
primary students in central China.
Front. Psychol. 14:1050894.
doi: 10.3389/fpsyg.2023.1050894

COPYRIGHT

© 2023 Lu, Chen, Leung and Zuo. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Reliability, validity, and measurement invariance of a Chinese handwriting legibility scale among primary students in central China

Hong Lu¹, Xin Chen¹, Frederick K. S. Leung^{1*} and Haode Zuo²

¹Faculty of Education, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, ²College of Mathematical Science, Yangzhou University, Yangzhou, Jiangsu, China

Background: Chinese handwriting has a close relationship with spatial cognition, and the legibility dimension is prominent with its spatial-oriented characteristics. However, handwriting evaluation focusing on the detailed spatial aspects of the legibility dimension in the Chinese context is rare.

Aims and methods: We aimed to develop a Chinese Handwriting Legibility Scale (CHLS) and examine its reliability, validity, and measurement invariance among Chinese primary students of different grades. A total of 684 students aged 8–12 years were recruited from a mainstream primary school in central China and were asked to copy a Chinese template as legibly as possible within 4 min. The developed CHLS was used to assess these students' legibility performance.

Results: The seven-criteria CHLS favored content validity. The inter-rater reliability was good; however, the scoring instructions need to be refined. Principal component analysis (PCA) revealed a one-factor solution explaining 62.336% of the variance of the seven-criteria CHLS, and confirmatory factor analysis (CFA) confirmed its appropriateness. There was a high internal consistency ($\alpha = 0.902$). In terms of measurement invariance, the factor structures and loadings of the CHLS were consistent across students of different grades; however, significant intercept variations were detected between students of Grades 2 and 4.

Conclusion: CHLS may be effective for evaluating Chinese handwriting legibility performance in the Chinese primary school context in the central region. Students' Chinese handwriting legibility performance may have developmental specificity in different grades.

KEYWORDS

Chinese handwriting, legibility, spatial cognition, development and validation, primary school students

1. Introduction

As children progress through school, they are expected to write legibly at a reasonable speed. Traditionally, handwriting has been closely associated with keeping up with class work and examinations (Barnett et al., 2018) and has been described as a 'low-level' perceptual motor skill in the broader writing process (Berninger et al., 2002; Barnett et al., 2018). However, handwriting

activities, which dynamically integrate perceptual motor, cognition, tactile, and kinesthetic sensitivities (Feder and Majnemer, 2007), also involve high-level cognitive processes (Kao, 2000). The fact that the “central” cognitive processes and “peripheral” motor processes continuously interact during written word production (Berninger and Swanson, 1994; Graham and Weintraub, 1996; Purcell et al., 2011; Roux et al., 2013; Kandel et al., 2017; Zhang and Feng, 2017) supports this standpoint. Cognitive processes here generally refer to cognitive planning, working memory processes, and phonological and orthographic coding (e.g., McCutchen, 2000; Volman et al., 2006). Whereas motor processes denote the retrieval/production of written forms, planning and ordering of the sequence of letters/characters and execution of specific motor programs (Ellis, 1982). In the field of education, handwriting maneuvers are reportedly intertwined with domain-specific spatial cognition, such as mental rotation (e.g., Li et al., 1999, 2014; Sakamoto and Spiers, 2014), spatial visualization (e.g., Kao, 2000; Likhonov et al., 2018), visual motor integration (e.g., Maeland, 1992; Tseng and Murray, 1994; Weil and Cunningham Amundson, 1994), and visual-spatial working memory (e.g., Flaherty and Connolly, 1995; Demetriou et al., 2005; Kazi et al., 2012). The underlying rationale is that written scripts are rich in geometric patterns and visual-spatial features, and handwriting production requires the script components to be organized with appropriate proportioning and accurate spatial relationships (Lai, 2008). In this sense, owing to the limited working memory capacity of humans, handwriting automation is vital not only for meeting the needs of examinations and daily learning but also for its potential associations with spatial cognition.

Spatial cognition particularly permeates Chinese handwriting processing (Kao, 2000). Compared with alphabetic letters built on phonemic structures, Chinese characters are more thorough visual characters, relying on various visual configurations, and are typically described as a logographic system (Lai, 2008; Lam et al., 2011). Chinese characters possess a high nonlinear complexity, with all strokes and radicals packed into an imaginary uniformed square (Chow et al., 2003; Tan et al., 2005). Thus, Chinese characters demand a substantially higher visual discrimination of fine changes in the shapes, locations, and spatial arrangement of strokes (Tan et al., 2001; Chow et al., 2003; Lai, 2008). Specifically, to legibly write a Chinese character, the writer needs to pay visual attention to not only the spatial position of each stroke relative to the imaginary square frame but also to the spatial relationship among strokes. After gaining a certain perceptual understanding and memory of the complex spatial relations, the writer then needs “to mentally represent each stroke spatially, to copy each one accurately, and to learn the spatial relations by heart through practice” (Li and Nuttall, 2001, p. 16). According to Li et al. (1999), writing Chinese characters (rather than alphabetic letters, the processing of which is primarily linear and emphasizes smoothness and continuity; Lai, 2008) can provide learners with more opportunities to experience spatial relations in the Euclidean system. Besides, as suggested by the Chinese-character writing’s psycho-geometric theory (Kao, 2000), Chinese handwriting is not only an external projection and execution of the writer’s internal cognitive images of Chinese words but also functions to re-train and improve the writer’s visuospatial cognition. Numerous cross-cultural comparative studies uncovered that East Asian (e.g., Chinese, Japanese) students presented advantage in spatial abilities relative to their Western (e.g., North American and European) counterparts (e.g.,

Flaherty and Connolly, 1995; Li et al., 1999; Demetriou et al., 2005; Kazi et al., 2012; Sakamoto and Spiers, 2014) and claimed that the differences between the Chinese and alphabetic languages, especially *the writing systems and handwriting processes*, play a fundamental role in shaping this cognitive difference, providing empirical support for this theory.

Although not specified, the close association between Chinese handwriting and spatial cognition is mainly sourced from the dimension of handwriting legibility. Handwriting legibility refers to the clarity, quality, precision, and accuracy of a person’s handwriting production (e.g., Maeland, 1992; Gilboa et al., 2010); according to Tomchek and Schneck (2006, chap. 14), it represents “the degree of the handwriting produced in consideration of alignment and size on a line and spacing between letters and words in relation to each other as well as the organization of the whole page.” As Bo et al. (2014) noted, among the different dimensions of handwriting performance, handwriting legibility reflects more about the spatial characteristics of script layout, relative to handwriting speed and pressure which correspond more to the temporal-oriented characteristics of the handwriting process. Common indicators to operationalize handwriting legibility include spacing/spatial relationships (or spatial organization), size, formation, alignment, slant/direction, and baseline orientation (e.g., Volman et al., 2006; Parush et al., 2010; Rosenblum et al., 2010; Klein et al., 2011; Bo et al., 2014; Linda et al., 2014). These indicators imply that spatial thinking underlies the legible Chinese handwriting processing, highlighting the need to evaluate Chinese handwriting legibility based on its spatial-oriented characteristics. However, despite the significance of this handwriting performance and the fact that numerous students struggle with this area (Lam et al., 2011), little research has been done to concretize the scenarios of Chinese handwriting legibility. This is partly attributable to the absence of robust assessment tools. Most handwriting evaluation scales have been developed to assess the handwriting quality or to detect dysgraphia via the teacher’s overall impressions of children’s handwriting production in class (e.g., Tseng, 1993; Chang and Yu, 2005, 2012; Barnett et al., 2007, 2013, 2018; Rosenblum and Livneh-Zirinski, 2008). These scales provide a valuable overview of various aspects of handwriting, including, for example, global legibility, production speed, page layout, effort to read the script, the motivation/attitude to write, fatigue, writing alterations (attempts made to rectify the writing), pencil grip, and gross movement. However, originating from teachers’ overall impression, these evaluations are often highly subjective, scale poorly, and fail to afford an in-depth analysis of specific, especially spatial, aspects of handwriting performance.

Computerized assessments using digitizing tablets (with various supporting programs, e.g., Rosenblum et al., 2006; Rosenblum and Livneh-Zirinski, 2008; Li-Tsang et al., 2011, 2013; Lee et al., 2016; Mekyska et al., 2016; Pagliarini et al., 2017; Asselborn et al., 2018; Gargot et al., 2020) support the precise measurement of some spatial characteristics (size, spacing, etc.) of handwriting production across diverse handwriting tasks varying in cognitive demands and task lengths. Particularly, combining the temporal characteristics collected during handwriting processing, these technologies make the analysis of legibility, dynamics (e.g., velocity, acceleration, etc.), pressure and even pen tilt of handwriting product and process sophisticated and automatic, thus providing more comprehensive and quantitative information concerning students’ handwriting acquisition and performance/disabilities. This information includes but is not limited to the pattern

identification of potential future handwriting impairments at a very early age (e.g., Pagliarini et al., 2017), and the classification (or digital diagnosis) of dysgraphic children (e.g., Mekyska et al., 2016; Asselborn et al., 2018) even with the age effect considered (e.g., Gargot et al., 2020). However, despite the increasing availability of digital tablets and their advantages in dynamic evaluation, regarding legibility evaluation, these new technologies usually pay more attention to stroke accuracy but are limited in holistic legibility. In addition, they are not always accessible to students in regular classrooms, and their manipulation is not intuitive to teachers (Barnett et al., 2018). All these limitations determine that their application to legibility evaluation in the classroom context remains to be improved.

Several scales were developed to assess handwriting legibility in great detail using multiple criteria. Examples of such scales include Minnesota Handwriting Assessment (MHA; Reisman, 2004), Scale of Children's Readiness In PrinTing (SCRIPT; Weil and Cunningham Amundson, 1994), Hebrew Handwriting Evaluation (HHE; Erez et al., 1996), Concise Assessment Scale for Children's Handwriting (BHK; Hamstra-Bletz et al., 1987), the Persian Handwriting Assessment Tool (PHAT; Havaei et al., 2017), the Evaluation Tool of Children's Handwriting-Manuscript (ETCH-M; Amundson, 1995), and Tseng Handwriting Problem Checklist (Tseng, 1993). The specific criteria assessed by these scales include but are not limited to spacing/spatial relationships (the relative position of strokes/letters/characters), the size of strokes/letters/characters, stroke/letter/character formation (e.g., closure, superfluous/missing strokes, line quality, and slant/direction), alignment of letters/characters, and baseline orientation (e.g., out of grid/line, overshooting or undershooting the baseline, and inappropriate margins) (e.g., Volman et al., 2006; Parush et al., 2010; Rosenblum et al., 2010; Klein et al., 2011; Cheng-Lai et al., 2013; Bo et al., 2014; Linda et al., 2014). The consistency/uniformity of or variation/error in the writers' production relative to a standard typically represents the extent of legibility (Graham et al., 2006; Bo et al., 2014). Although such scales can appropriately assess the spatial characteristics of handwriting production, the extensive list of criteria covered by these scales makes their use in a classroom setting difficult and time-consuming (e.g., Asselborn et al., 2018). Besides, most of these criteria are applicable to alphabetic contexts, and their appropriateness in and contribution to Chinese handwriting legibility assessment are yet to be examined (Li-Tsang et al., 2013) (see a summary of existing handwriting evaluation instruments in [Supplementary material](#)).

In this context, both teachers and students need a tailored practical tool for evaluating the detailed spatial aspects of Chinese handwriting legibility. This tool may be helpful in describing and quantifying students' Chinese handwriting legibility performance and identifying Chinese handwriting difficulties that are specifically caused by spatial cognition deficits, which in turn assisting the development of corresponding supporting plans. This will also enable teachers and students to consider the relationships between Chinese characters and Chinese handwriting and spatial thinking, fostering their consciousness and autonomy of understanding Chinese character learning from a higher-level perspective. Accordingly, in this study, we aimed to develop a reliable and valid Chinese Handwriting Legibility Scale (CHLS) focusing on the detailed spatial aspects of the handwriting legibility dimension.

This study included Chinese students (aged 8–12 years) attending a primary school in central China who had undergone several years

of Chinese handwriting training and should have developed certain written communication skills. The students were from Grades 2, 4, and 6, which aligns with the main student groups assessed in related research (e.g., Tseng and Chow, 2000; Li-Tsang et al., 2013; Lee et al., 2016). This cross-grade sampling design was adopted not only because it supports the scale development and validation aiming for a broad application in primary school stages but also because it provides opportunities for examining the measurement invariance of the CHLS (of course constrained by central China). To the researchers' knowledge, measurement invariance of handwriting instruments has rarely been assessed in previous studies; hence, our study may provide novel insights into the generalizability of the CHLS and the possible developmental changes underlying the Chinese handwriting legibility performance among students of different grades in central China.

2. Methods

2.1. Development of the CHLS criteria and assessment of content validity

Following the guidelines of Li-Tsang et al. (2013), the measurement criteria were developed or selected based on the following four major concerns: (i) the criteria regularly adopted by schoolteachers or occupational therapists to evaluate students' (Chinese) handwriting, (ii) the recognized characteristics of children with handwriting problems, (iii) the logographic nature and visual-spatial properties of Chinese characters, and (iv) a review of the literature. Besides, the assessment was based on analytic and holistic impressions of handwriting production suggested by Lam et al. (2011) and Li-Tsang et al. (2013). The analytic evaluation approach focuses on judging or grading the quality of stroke-level (i.e., *within* character) handwriting features according to predetermined standards. By contrast, the holistic evaluation approach focuses on assessing the character-level (i.e., *between* characters) features of a written passage as compared with a group of pre-graded writing samples.

Accordingly, 10 legibility criteria were initially established to form the basis of CHLS: h1: spacing/spatial relationships between *strokes/radicals*; h2: spacing/spatial relationships between *characters*; h3: alignment of characters; h4: baseline orientation; H5: uniformity of *stroke/radical* size; h6: uniformity of *character* size; h7: number of strokes (no superfluous/missing strokes); h8: closure of stroke/radical; h9: line formation; and h10: direction. The instructions for scoring h1, h5, h7, h8, h9, and h10 emphasized decision at the analytic level, whereas those for scoring h2, h3, h4, and h6 corresponded to the legibility performance at the holistic level. A Likert scale ranging from 1 (poor) to 5 (good) with scoring instructions and examples was developed to assess each criterion. These were applied to 'copying' products gathered as part of the Smart Handwriting Analysis and Recognition Platform (SHARP) handwriting task assessment (Li-Tsang et al., 2022; see Section 2.5.1 below). China does not have a single, prescribed writing style; however, the basic horizontal and vertical stroke requirements tend to be invariant throughout students' development, as reflected in our study. The CHLS was applied to the writing with at least three lines of handwriting. The total scores of the initial version ranged from 10 to 50, with higher scores indicating better legibility. A scoring sheet was designed, with the summed score representing the global legibility score. The first author and two

TABLE 1 Number of boys and girls included in the study according to age.

Age (years)	Boys	Girls	Total
~8	115	94	209
~10	126	123	249
~12	126	100	226

research assistants independently scored 10 handwriting samples from Grade 4. This process refined the scoring sheet's wording and layout to enhance the overall ease of use.

Next, four experts from different professions (Chinese teaching, educational psychology, and mathematics education) evaluated the tool independently. Among them, two Chinese language teachers have rich expertise in teaching Chinese handwriting; another two researchers majoring in educational psychology, and mathematics education, respectively, have expertise in spatial cognition, and educational measurement and assessment; all these experts are native Chinese speakers and fluent in Chinese handwriting. An example handwriting sample that was previously scored using the CHLS was provided as a reference. With five additional products, the experts needed to apply the scale to these five and then fill out a feedback form (see details in [Supplementary material](#)) detailing their thoughts on the clarity of each criterion, the content breadth, and the degree to which they thought each criterion contributes to the construct of 'Chinese handwriting legibility'. The experts were also asked to make any additional remarks on the CHLS.

2.2. Inter-rater reliability and construct validity

To assess the inter-rater reliability, two new raters were invited to independently score the products from 20 s grade students (randomly selected); this process was repeated for the products from 20 fourth and sixth grade students as well. [Barnett et al. \(2018\)](#) suggested that the narrower age band focusing on each scoring round could eliminate any discrepancies caused by age and render the discrimination between samples easier. The new raters were both 4th grade Chinese teachers with rich Chinese teaching experience throughout the whole elementary school; and they were trained by the first author to use the CHLS. Hence the raters' standards in scaling can be unified to some extent. The first author scored all the handwriting products, and the total CHLS scores were divided into three categories in each grade: low, medium, and high (the mean minus/plus one standard deviation was set as the cut-off standard; [Cascio et al., 1988](#); [Barnett et al., 2018](#)). The inter-rater reliability was calculated by applying the Cohen's kappa coefficient between each new rater's scores and the first author's scores, and six inter-rater reliabilities were finally obtained.

To examine the construct validity, 50% of the participants' products were randomly selected and subjected to a principal component analysis (PCA) to determine the number of components assessed by the CHLS. The remaining 50% of the participants' products were subjected to confirmatory factor analysis (CFA) to further confirm the appropriateness of the factor structure explored using PCA. As boys' handwriting is usually poorer than girls' (e.g., [Graham et al., 1998](#); [Cui et al., 2012](#); [Wicki et al., 2014](#)), gender effects on each criterion and the

total CHLS scores were also recorded to support the validity of the CHLS.

2.3. Internal consistency reliability and measurement invariance across grades

The internal consistency reliability of the entire sample was calculated using Cronbach's alpha coefficient.

Measurement invariance (multi-group CFA), also with the entire sample, was assessed to examine the generalizability of the CHLS among the Grades 2, 4, and 6 students. A chi-square ratio (χ^2/df) of ≤ 3 ([Schermelleh-Engel et al., 2003](#)), comparative fit index (CFI) of ≥ 0.95 , Tucker–Lewis index (TLI) of ≥ 0.95 , standardized root mean squared residual (SRMR) of ≤ 0.08 , and root-mean-square error of approximation (RMSEA) of ≤ 0.06 ([Hu and Bentler, 1999](#)) were considered indicators of good model fits. For multi-group comparisons, the significance of $\Delta\chi^2$ was used to judge the change in model fit between the compared models ([Byrne, 2001](#)).

2.4. Participants

A total of 684 students (367 boys, 317 girls) aged 8–12 years were recruited from a mainstream public primary school in Jingmen city in Hubei Province, China. Hubei Province is located in the central region of China, with a medium level of economic development. In the first three quarters of 2022 GDP ranking of provinces, Hubei ranked eighth with 3729.89 billion yuan among the 31 provinces; on this basis, located in the central district of Hubei Province, in its first three quarters of 2022 GDP ranking of cities, Jingmen city ranked seventh with 156.061 billion yuan among the thirteen cities. In this sense, with the middle economic development level, the Jingmen district should be representative of the average quality of basic education and academic level of students in central China. Due to personal limitations, with convenience sampling, only one mainstream public primary school in the urban area in Jingmen city was involved in the study, then 4, 5, and 4 classes from Grades 2, 4, and 6 were randomly chosen as the sample classes, respectively.

Of the 684 students, 209 were from Grade 2, 249 were from Grade 4, and 226 were from Grade 6 ([Table 1](#)). All participants are right-handed. According to the school records and teacher feedback, none of the students had developmental delays; dyslexia; neurological deficits; physical or mental challenges; or behavioral and emotional issues and sensory processing disorders.

2.5. Measures

2.5.1. The SHARP handwriting task

As familiarity with and the complexity of the characters in the handwriting task can directly impact students' evaluated performance, handwriting tasks should resemble what students regularly write at school but with varied difficulties ([Li-Tsang et al., 2013](#)). The template used by the SHARP evaluation for Chinese handwriting ([Li-Tsang et al., 2022](#)) was adopted in the present study; this template includes 90 simplified Chinese characters with font size 26, font type 'KaiTi', and triple-line spacing displayed in nine columns of 10 characters on an

A4-sized paper. The template is designed on the one hand, based on how frequently the characters are used on a daily basis (Poon and Hong, 2003) – characters with low frequency were discarded from the selection so as to minimize the probability of students making errors because they had not learned that specific character before. On the other hand, the template covers all the six basic structures of Chinese characters (i.e., above–bottom, left–right, above–middle–bottom, left–middle–right, inside–outside, and independent) and 25 of the 30 basic stroke units (Law et al., 1998) to ensure the representativeness of the selected Chinese characters. Two 4th grade Chinese language teachers have also reviewed the template and confirmed its appropriateness for the Chinese primary school students in central China.

Contrary to the original experimental design requiring students to copy the characters on a digitized tablet without time limits, the present task required the students to copy the template on an A4-sized grid paper (top to bottom, left to right) as legibly as possible without compromising on speed within 4 min (but they did not need to copy the entire template). As demonstrated earlier, temporal characteristics reflect dynamics of handwriting process, signifying one critical dimension of handwriting evaluation; and the temporal pressure typically cause degradation of handwriting production (e.g., Gargot et al., 2020). Therefore, although this study does not involve the evaluation of handwriting dynamics, to ensure the potential comparability of the legibility results with prior work, time constraints were also adopted in this study. Additionally, requiring students to write with a certain speed can simulate the daily classroom setting, hence greatly reflecting their actual handwriting performance.

Within a time frame of 4 min, most of participants cannot finish the whole handwriting, and slow writers can copy at least first three lines. The similar research design can refer to existing handwriting research (e.g., Volman et al., 2006; Kaiser et al., 2009; Hellinckx et al., 2013; Van Hartingsveldt et al., 2015). The first author and the invited raters then assessed the handwriting legibility manually using the CHLS.

Moreover, to unify the assessment procedures, the environment set-up with the real-life handwriting context, and ergonomic factors such as the placement of handwriting materials, the use of writing accessories, lighting and noise, and writing posture were carefully monitored and controlled to ensure consistency among participants from different grades (Feder and Majnemer, 2007; Li-Tsang et al., 2013).

3. Data analysis

PCA and CFA were performed to determine and validate the factor structure of the CHLS, and the *t*-test was conducted to evaluate the gender differences in each criterion and the overall scores. Multi-group CFA was conducted to assess the measurement invariance of the CHLS. All data were analyzed using SPSS 26.0 and Mplus 8.3.

4. Results

4.1. Content validity

After rating the five products using the CHLS, the experts provided independent feedback based on their expertise in Chinese

handwriting and spatial cognition, practical experience working with students, and development and usage of other assessment instruments. The experts thought that most of the criteria were clearly presented, and further remarked that more scoring examples would be helpful to assist new raters in understanding the different criteria. Although comprehensive, some criteria, including h4, h7, and h8, were considered inappropriate to be applied to quantitative evaluation of Chinese handwriting legibility. h4 has rarely been emphasized in daily Chinese handwriting practice (although the lower grade students frequently use the square frame as reference in actual writing, teachers rarely emphasize the harms of out of grid/line; besides, in Chinese handwriting worksheet, there are not so many horizontal baselines as in alphabetic writing, hence the overshooting or undershooting the baseline in Chinese handwriting is negligible), hence the students performed with a certain randomness. h7 with superfluous/missing strokes generally may not impair the spacing/spatial relationships at the stroke or character level, nor the overall recognition of a specific character, considering that most Chinese characters are multi-stroke. As evidence, Graham et al. (2006) confirmed that in the copying task, h7 was typically used to characterize the construct of the motor program instead of the visual–spatial relationship or formation. Finally, h8 the extent of closure of stroke/radical varied depending on students' handwriting styles and fonts. Especially in higher grades, with the handwriting fluency and styles develop, students pay little attention on the closure issue; even different writing tools can show different visual effects of closure. A closer inspection of the handwriting samples of the Grade 4 students based on these criteria suggesting that it is difficult to summarize regularity then determine standard on these rating supported the experts' feedback. Moreover, the expert panel also provided some recommendations for the wording on the scoring sheet, such as revising the wording of h9 from 'line formation' to 'line quality', considering that h7–h10 collectively depict the stroke formation in the literature. Overall, the experts' opinions were in favor of including the other seven criteria and their feedback clarified the pertinent descriptions and scoring guidelines.

4.2. Inter-rater reliability and construct validity

With h4, h7, and h8 excluded from the evaluation, the raw score of the seven-criteria CHLS ranges from 7 to 35. The score range and percent of each category for Grades 2, 4, and 6 are shown in Table 2. Take Grade 2 as an example, the 'high' category represents more than one standard deviation above the mean of the 209 participants (i.e., 25.77 plus 5.17 = 30.94 and rounded down to a score of 31).

On this basis, the inter-rater reliability was good. For the Grade 2 students, the Cohen's kappa coefficient between the first author and rater 1 was 0.832 ($p < 0.001$), and between the first author and rater 2 was 0.916 ($p < 0.001$); for the Grade 4 students, the Cohen's kappa coefficient between the first author and rater 1 was 0.706 ($p < 0.001$), and between the first author and rater 2 was 0.866 ($p < 0.001$); and for the Grade 6 students, the Cohen's kappa coefficients between the first author and rater 1 was 0.752 ($p < 0.001$), and between the first author and rater 2 was 0.864 ($p < 0.001$).

Construct validity was assessed in two phases. In phase one, 50% of the students' products were randomly selected and subjected to PCA ($N = 341$) using SPSS. The Kaiser–Meyer–Olkin (KMO) value in this

TABLE 2 The mean, standard deviation, score range, and percent of each category for Grades 2, 4, and 6 of the seven-criteria CHLS.

	M	SD	Low	Medium	High
Grade 2	25.756	5.174	7 ≤ CHLS <21 (20%)	21 ≤ CHLS <31 (65.6%)	31 ≤ CHLS ≤35 (14.4%)
Grade 4	26.839	4.323	7 ≤ CHLS <23 (18%)	23 ≤ CHLS <31 (67%)	31 ≤ CHLS ≤35 (15%)
Grade 6	27.566	4.247	7 ≤ CHLS <23 (15.5%)	23 ≤ CHLS <32 (68.5%)	32 ≤ CHLS ≤35 (16%)

TABLE 3 Factor loadings of the seven-criteria CHLS using CFA ($N = 343$).

CHLS criteria	Factor loadings
h1	0.905
h2	0.713
h3	0.708
h5	0.756
h6	0.598
h9	0.784
h10	0.704

study was 0.910; the results of Bartlett's test of sphericity were significant ($\chi^2 = 1266.795$, $df = 21$; $p < 0.001$), indicating that it could be used for factor analysis. Furthermore, a single factor solution (with the eigenvalue being 4.364) was indicated through the examination of the screen plot and eigenvalues; this explained 62.336% of the variance observed. In phase two, the remaining participants' products were subjected to CFA ($N = 343$), which was used to confirm the one-factor structure of Chinese handwriting legibility observed using PCA. The normality of the involved criteria was first checked to confirm that the precondition of maximum likelihood estimation was met. The absolute values of skewness and kurtosis should be <3 and <8 , respectively (Chen et al., 2005). In our study, the skewness ranged from -0.464 to -0.170 and the kurtosis ranged from -0.709 to -0.051 , indicating that the criteria satisfied the standards of normality. On this basis, our CFA results indicated a satisfactory model fit ($\chi^2 = 12.623$, $df = 11$, $\chi^2/df = 1.148$, $p = 0.3187$; CFI = 0.999, TLI = 0.998, RMSEA = 0.021, 90% CI [0.000, 0.062], SRMR = 0.015). The factor loadings were all significant ($p < 0.001$), ranging from 0.598 to 0.905 (See Table 3).

In terms of the whole sample, mean scores for each criterion and the total CHLS for the boys and girls are displayed separately in Table 4. Significant gender differences were detected for each criterion and the total CHLS.

4.3. Internal consistency reliability and measurement invariance across grades

The internal consistency of the CHLS was calculated using Cronbach's alpha. The Cronbach's alpha for all the measurement criteria was 0.902, indicating a high internal consistency. The

TABLE 4 Means, standard deviations, and p values of the boys' and girls' scores for each criterion and the total CHLS.

CHLS criteria	Boys ($N = 367$)		Girls ($N = 317$)		p value
	M	SD	M	SD	
h1	3.59	0.909	3.94	0.795	< 0.001
h2	3.75	0.885	4.04	0.762	< 0.001
h3	3.62	0.876	4.00	0.821	< 0.001
h5	3.68	0.792	3.92	0.763	< 0.001
h6	3.76	0.798	4.04	0.768	< 0.001
h9	3.59	0.898	3.85	0.829	< 0.001
h10	3.82	0.809	4.04	0.728	< 0.001
Total	25.81	4.812	27.83	4.153	< 0.001

item-total correlation varied between 0.742 to 0.858, and the alpha ranged from 0.877 to 0.894 when one of the criteria was deleted, indicating that deleting some criteria would not help improve the overall internal consistency (Table 5).

To assess the measurement invariance, the data normality of each grade was first tested. As shown in Table 6, all criteria for the three groups (Grade 2 [skewness: -0.345 to -0.060 ; kurtosis: -0.962 to -0.533], Grade 4 [skewness: -0.326 to 0.004 ; kurtosis: -1.004 to -0.256], and Grade 6 [skewness: -0.457 to -0.074 ; kurtosis: -0.728 to 0.268]) met the standards of normality for maximum likelihood estimation.

Then, separate CFAs were conducted for the Grades 2, 4, and 6, respectively. The results showed that the model fits were satisfactory (Table 7), and the factor loadings were all significant and almost >0.60 (Table 6), indicating that the one-factor structure was appropriate for all three groups.

On this basis, to detect the cross-grade generalizability of the CHLS, three nested models were compared. In the configural model, all parameters were freely estimated. The metric model fixed the factor loadings among all groups. In the scalar model, both factor loadings and intercepts were constrained.

As shown in Table 8, the fits of the configural and metric models were not significantly different ($p = 0.089$), indicating that the factor loadings were invariant, and the metric invariance was established among the three groups. Scalar invariance was examined by further fixing the intercepts among the groups. A significant change was identified ($p < 0.05$; $\Delta CFI = 0.011 > 0.01$, $\Delta RMSEA = 0.017 > 0.005$; see Chen, 2007; Putnick and Bornstein, 2016), indicating that the scalar invariance of the CHLS was not supported across the three groups.

To identify the specific reason for this non-invariance, we separately assessed the measurement (especially scalar) invariance between two of the three groups. The results (Tables 9–11) showed that scalar invariance was supported between Grades 4 and 6 and between Grades 2 and 6 but not between the Grades 2 and 4. Further, focusing on the partial scalar invariance between Grades 2 and 4 (Table 9), by sequentially releasing criterion intercept constraints and retesting the model, it was found that when h5 and h6 were freely estimated, the partial scalar invariance was supported, indicating the differences of h5 and h6 between Grades 2 and 4.

On this basis, ANOVA-test (and *post hoc* multiple comparison) among the means of each criterion for the Grades 2, 4, and 6 was conducted to further compare students' performance in each criterion. The results showed that when students progressed from Grade 2 to

Grade 4, except for h2 and h3, significant performance improvements were noted for other criteria; whereas when students progressed from Grade 4 to Grade 6, only h2, h3, and h6 presented significant improvements; overall, students significantly progressed on each criterion from Grade 2 to Grade 6 (Table 12).

5. Discussion

The logographic nature and visual-spatial properties of Chinese characters determine the uniqueness and complexity of Chinese handwriting, and its close association with spatial cognition (Chinese-character writing's psycho-geometric theory; Kao, 2000). Numerous cross-cultural comparative studies support this proposition. Combined with the spatial-oriented characteristics of the legibility dimension, the spatial-related nature of Chinese handwriting legibility is further strengthened. In other words, spatial thinking is implicit in legible Chinese character writing, and there is an implicit association between Chinese handwriting legibility and spatial cognition. In this regard, evaluation and diagnosis of Chinese handwriting legibility taking into account a detailed spatial analysis at the analytic and holistic levels, is of great significance for student development and teacher instruction, and the CHLS in the present study was developed in this context.

The seven criteria of the CHLS are supported by the literature and the experts' reviews. Spacing/spatial relationships between strokes/radicals are evaluated based on the extent to which the components are positioned correctly; examples of errors include but

are not limited to overlaps or writing too far apart, collisions and adhesions of components, and dislocation of components. This criterion is an analytic-level measure relative to the spaces between characters, which assess whether and how the characters are evenly separated in the whole script. Alignment focuses on the overall horizontal and vertical layout of the characters. Size uniformity is evaluated at the analytic (focusing on specific strokes/radicals within a character) and holistic (focusing on characters in the whole text) levels. Line quality is another crucial criterion, with errors often originating from poor line formation (e.g., if the curves are angular or straight lines are wavy). Finally, writing in the appropriate direction also facilitates readability (e.g., Parush et al., 2010; Klein et al., 2011). This include no deviations in the orientation of specific strokes, and the character is oriented vertically relative to the horizontal line. These criteria concur well with and complement the claim of Li-Tsang et al. (2022) that the location, proportion, size, and direction of the strokes are all vital for legible Chinese handwriting.

Inter-rater consistency is a critical indicator in weighing the reliability of the CHLS in its application to different raters. According to Barnett et al. (2018)'s method in examining the inter-rater reliability of handwriting scales, the participants' total scores of Chinese handwriting legibility performance of each grade were categorized into the low, medium, and high groups. The inter-rater reliabilities based on these classifications were good but could be further improved. An in-depth analysis uncovered that this was primarily due to the moderate agreement between raters on the 'direction' criterion. Subsequent discussions with the raters indicated that they were somewhat uncertain about assessing strokes/radicals with deviated directions and samples with personalized handwriting styles/fonts. This finding highlights the necessity to further refine and clarify the instructions and provide more examples for future raters.

As suggested by Barnett et al. (2007), a cut-off score is necessary for a test to identify those with poor performance and even difficulties. To achieve this goal, we divided the CHLS total scores into low, medium, and high categories in each grade so as to more accurately identify the sample with poor Chinese handwriting legibility. With a normal distribution, more than 15% of the students in each grade has been identified with potential handwriting difficulties, which aligns with the frequency reported in existing literature (e.g., Feder and Majnemer, 2007). In this sense, the CHLS might be an appropriate screening tool concerning legibility problems; at present we would recommend these levels (Table 2) to identify children with poor

TABLE 5 The item-total correlation of the CHLS measurement criteria and the corresponding alpha if item deleted ($N = 684$).

CHLS criteria	Item-total correlation	Alpha if item deleted
h1	0.858	0.877
h2	0.802	0.886
h3	0.822	0.883
h5	0.760	0.891
h6	0.771	0.890
h9	0.793	0.888
h10	0.742	0.894

TABLE 6 Mean, factor loading, skewness, and kurtosis of the criteria in the three groups.

	Grade 2				Grade 4				Grade 6			
	M	FL	S	K	M	FL	S	K	M	FL	S	K
h1	3.59	0.881	-0.317	-0.534	3.77	0.856	-0.254	-0.535	3.89	0.883	-0.369	-0.473
h2	3.80	0.661	-0.331	-0.695	3.82	0.687	-0.326	-0.302	4.02	0.688	-0.457	-0.341
h3	3.70	0.717	-0.160	-0.856	3.76	0.739	0.070	-1.004	3.92	0.700	-0.416	-0.236
h5	3.62	0.721	-0.321	-0.533	3.85	0.743	-0.016	-0.590	3.88	0.708	-0.200	-0.356
h6	3.69	0.657	-0.060	-0.937	3.91	0.583	-0.159	-0.256	4.06	0.653	-0.231	-0.594
h9	3.56	0.756	-0.186	-0.962	3.74	0.760	0.004	-0.743	3.82	0.806	-0.074	-0.728
h10	3.80	0.692	-0.345	-0.834	3.99	0.719	-0.045	-0.923	3.97	0.703	-0.202	0.268

FL, factor loading; S, skewness; K, kurtosis.

Chinese handwriting legibility, and likely in need of instruction support.

The construct validity depicts how well a scale measures the construct it is intended for. The CHLS was developed to assess the detailed spatial characteristics of Chinese handwriting legibility at both the analytical and holistic level, and PCA of the seven criteria uncovered a one-factor solution explaining a large proportion of the variance. This result supported the significance of these criteria in constructing the overall legibility dimension. Furthermore, the factor loadings of the CHLS criteria were relatively high for h1, h5, and h9; whereas were lower for h2 and h6. The size tendency of factor loadings may reflect on the one hand, the prominent role of analytic criteria in the legibility dimension, and on the other hand, the differences between the analytic versus holistic criteria. The CFA results further supported this one-factor structure. Besides, the significant gender differences observed in this study, that is, boys generally performed lower in legibility than girls, is consistent with existing research (e.g.,

Graham et al., 1998; Cui et al., 2012; Wicki et al., 2014), hence also supplements the validity of the CHLS to some extent.

Combined with the high internal consistency reliability showing that the seven criteria were closely related with the total score, all these reliability and validity results suggest that the CHLS developed in this study is appropriate for measuring the Chinese handwriting legibility performance of Chinese primary students in central China.

On this basis, measurement invariance of the CHLS was further assessed for students across Grades 2, 4, and 6. The results confirmed the metric invariance of the CHLS across the three grades, suggesting that the factor structure of the CHLS, and relative contribution (i.e., the factor loading) of each criterion to the latent construct were consistent for students of these three grades. This invariance, to some extent, supported the generalizability of the CHLS in the Chinese primary school context in central China. Further measurement invariance tests between two of the three grades revealed the scalar invariance between Grades 4 and 6 and between Grades 2 and 6, but significant intercept variations (i.e., h5 and h6) between the Grades 2 and 4. As Cheung and Rensvold (2002) claimed, cross-group differences detected in multi-group CFA may be valuable for understanding 'how different groups view the world' (p. 252). In other words, different groups may hold distinct attitudes, perceptions, or ratings on the criteria concerned. In the context of this study, the cross-group differences in terms of h5 and h6 of the CHLS may reflect the developmental specificity of students' Chinese handwriting legibility in size across lower grades. Although to the researchers' knowledge, there are no existing research in handwriting development explicitly confirmed this finding, children with poor handwriting legibility or dysgraphia, or namely low handwriting ability, usually present the greater variability of stroke/radical/character size (e.g.,

TABLE 7 Model fit statistics for CFAs in the three groups.

Groups	N	χ^2 (df)	χ^2/df	CFI	TLI	RMSEA	SRMR
Grade 2	209	14.343 (11)	1.304	0.996	0.992	0.038	0.019
Grade 4	249	12.624 (11)	1.148	0.998	0.996	0.024	0.018
Grade 6	226	16.106 (11)	1.464	0.994	0.989	0.045	0.018

CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root-mean-square error of approximation; SRMR, standardized root mean squared residual. The same below.

TABLE 8 Measurement invariance results of the CHLS for Grades 2, 4, and 6.

	χ^2 (df)	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	$\Delta RMSEA$
Configural model	42.340 (33)	0.996	0.993	0.035 [0.000, 0.063]	0.018			
Metric model	61.305 (45)	0.994	0.991	0.040 [0.000, 0.063]	0.068	18.965 (12) $p = 0.089$	-0.002	0.005
Scalar model	99.959 (57)	0.983	0.982	0.057 [0.038, 0.076]	0.072	38.654 (12) $p < 0.000^{**}$	0.011	0.017

$^{**}p < 0.01$, two-tailed. The same below.

TABLE 9 Measurement invariance results of CHLS for the Grades 2 and 4.

	χ^2 (df)	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	$\Delta RMSEA$
Configural model	23.015 (22)	0.999	0.999	0.014 [0.000, 0.058]	0.017			
Metric model	32.761 (28)	0.997	0.996	0.027 [0.000, 0.060]	0.053	9.746 (6) $p = 0.136$	0.002	0.013
Scalar model	51.449 (34)	0.990	0.988	0.047 [0.016, 0.072]	0.070	18.689 (6) $p < 0.005^{**}$	0.007	0.02
Partial scalar model (h5, h6 were freed)	37.985 (32)	0.997	0.995	0.029 [0.000, 0.059]	0.056	5.224 (4) $p = 0.265$	0.000	0.002

Hamstra-Bletz and Blöte, 1993; Marr et al., 2001; Volman et al., 2006) support this finding to some extent. Furthermore, a thorough comparison of each criterion across grades can shed more light on the developmental specificity of students' Chinese handwriting legibility. Specifically, the *post hoc* multiple comparison results between Grades 2 and 4 further revealed that students' Chinese handwriting legibility significantly progressed at the analytic level from Grade 2 to Grade 4; whereas there were no *corresponding* significant differences between Grade 4 and Grade 6, i.e., the relevant skills became relatively mature and automatic after Grade 4. This change may be closely related to students' Chinese character learning experiences. In the early grades, due to students' relatively limited capacity, handwriting instruction, learning and assessment usually focus more on the precise spatial analysis and processing of specific strokes, radicals, or other subcomponents, which may contribute to their significant progression in legibility performance at the analytic level. However, when students enter the senior grades, with their accumulation of Chinese handwriting experience, formal handwriting instruction in class reduces, and teachers focus more on the content and fluency of students' work rather than their analytical legibility (Lam et al., 2011). Hence students' Chinese handwriting legibility performance may reach a plateau at this stage. Noteworthy, although the time invested for formal handwriting instruction reduces in Grades 4–6, as Lee et al. (2016) claimed, Chinese students learn new Chinese characters throughout elementary school, and they continuously adhere to all orthographic rules and are consistently reminded to assure the quality of their handwriting output. These constant handwriting requirements might greatly contribute to students' satisfactory legibility in the senior grades, as observed in the present study. Indeed, there were significant improvements in h2, h3 and h6 behind the scalar invariance from Grade 4 to Grade 6. This finding is valuable for it reflects the legibility improvement at the holistic level. Similarly, despite the scalar invariance between Grade 2 and Grade 6, there were significant improvements in handwriting legibility performance at both the analytic and holistic levels. All these changes signify that there exists developmental specificity in these students' Chinese handwriting legibility performance: among the different criteria, handwriting size presents greater variability across the lower grades; and the early ability improvement is more reflected at the analytic level, whereas in the senior grades it centers around the holistic level.

The complex and dynamic nature of handwriting activities may also contribute to these students' developmental specificity in Chinese handwriting legibility performance. According to Feder and Majnemer (2007), handwriting is a complex human activity interweaving perceptual motor, cognition, and tactile and kinesthetic sensitivities. Kao (2000) further refined that Chinese handwriting involves the dynamic integration of the writer's perception, cognition, and motor components. These definitions imply that there involves inevitably differential skill mobilization behind students' handwriting progression, which can to some extent act as the theoretical basis favoring the transfer of spatial relationship emphasis from the analytic level to the holistic level in handwriting with students' growth in this study. More empirical research is demanded in the future to examine the plausibility of this explanation.

Lastly, it's worthy to note that despite the present study highlights spatial thinking, or specifically spatial cognitive processing needs behind the seven criteria of Chinese handwriting legibility, it is undeniable that other criteria/factors also make a difference in handwriting skill development. For example, as Li-Tsang et al. (2013)

demonstrated, despite being undetectable in written production, reversal of strokes/radicals and sequence errors could also be the possible reasons causing handwriting difficulties. Tseng and Murray (1994) indicated that learning disabilities or behaviors linked to attention disorders may impair the acquisition and maintenance of handwriting proficiency. Besides, through a review of the development of the aural, oral, reading, and writing systems in school children, Berninger (2000) concluded that these communication systems are interrelated with changes in one influencing the development of the others. In this sense, the present study is still the initial work, and a more refined research design with those factors controlled is needed in future work.

6. Limitations and implications

It is essential to keep in mind the limitations of this initial work, and the relevant results should be interpreted with caution. First, only Chinese primary school students from central China were included in this study, making the presented CHLS constrained in its application in the broad Chinese context. Future research should expand the sampling area to examine and improve the instrument's generalization. Second, although the relative importance of different spatial criteria in the Chinese handwriting legibility assessment is reflected, the standard-setting of the CHLS is based on simplified Chinese characters, a specific type of logographic script, and its applicability to other graphic scripts (e.g., traditional Chinese characters, McBride, 2016; Japanese kanji, Sakamoto and Spiers, 2014) remains uninvestigated. Given the possible differences in the representational characteristics of different scripts, we remain cautious in extending our findings to other scripts. Third, as with any individual research, the reliability and validity examinations conducted in this study were limited, and further work is still needed to supplement the CHLS's psychometric properties. For example, based on the expert panel and authors' qualitative evaluation, providing more quantitative evidence on the criteria inclusion/exclusion would afford more rigor to the content validity. This study would also be more comprehensive if children with specified disorders (e.g., Attention Deficit Hyperactive Disorders (ADHD), Developmental Coordination Disorder (DCD), which are commonly known to be connected with handwriting problems; e.g., Mayes and Calhoun, 2006; Barnett et al., 2018) and typical development could be incorporated to examine its discriminant validity. Besides, comparing the CHLS to other measures of handwriting legibility, such as computerized measures and eyeballing analysis by supervising teachers or parents of the relevant criteria, would provide information about the convergent validity of the CHLS (e.g., Li-Tsang et al., 2013). Fourth, the sensitivity of the CHLS to evaluate changes in handwriting legibility performance related to implementing a particular intervention or support remains unresolved. The relevant empirical evidence is needed to further support the application of the CHLS in clinical or instructional contexts. Fifth, although the handwriting produced in the required time is recorded alongside the CHLS instructions and can indicate the production rate, it is not the focus of this study, and a more rigorous and formal evaluation of handwriting speed is

TABLE 10 Measurement invariance results of CHLS for the Grades 4 and 6.

	$\chi^2 (df)$	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2 (\Delta df)$	ΔCFI	$\Delta RMSEA$
Configural model	28.730 (22)	0.996	0.993	0.036 [0.000, 0.069]	0.018			
Metric model	36.911 (28)	0.995	0.992	0.037 [0.000, 0.066]	0.057	8.181 (6) $p = 0.225$	−0.001	0.001
Scalar model	47.761 (34)	0.992	0.990	0.041 [0.000, 0.067]	0.052	10.851 (6) $p = 0.093$	−0.003	−0.002

TABLE 11 Measurement invariance results of CHLS for the Grades 2 and 6.

	$\chi^2 (df)$	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2 (\Delta df)$	ΔCFI	$\Delta RMSEA$
Configural model	26.497 (22)	0.997	0.995	0.031 [0.000, 0.067]	0.017			
Metric model	34.945 (28)	0.996	0.994	0.034 [0.000, 0.066]	0.058	8.447 (6) $p = 0.207$	−0.001	0.003
Scalar model	46.123 (34)	0.993	0.991	0.040 [0.000, 0.068]	0.064	11.178 (6) $p = 0.083$	−0.003	−0.006

TABLE 12 Means, standard deviations, and results of the ANOVA-test (and *post hoc* multiple comparison) for the Grades 2, 4, and 6 for each criterion of the CHLS.

CHLS	Grade 2 (N = 209)		Grade 4 (N = 249)		Grade 6 (N = 226)		F	p value	p value (Grade 2 vs. 4)	p value (Grade 4 vs. 6)	p value (Grade 2 vs. 6)
	M	SD	M	SD	M	SD					
h1	3.59	0.916	3.77	0.848	3.89	0.844	6.358	0.002**	0.033*	0.126	0.000**
h2	3.80	0.912	3.82	0.804	4.02	0.802	4.681	0.010*	0.804	0.010*	0.007**
h3	3.70	0.930	3.76	0.868	3.92	0.804	3.656	0.026*	0.428	0.050*	0.009**
h5	3.62	0.870	3.85	0.728	3.88	0.745	7.581	0.001**	0.001**	0.640	0.000**
h6	3.69	0.932	3.91	0.698	4.06	0.719	12.111	0.000**	0.003**	0.038*	0.000**
h9	3.56	0.980	3.74	0.828	3.82	0.808	5.349	0.005**	0.025*	0.294	0.001**
h10	3.80	0.960	3.99	0.727	3.97	0.625	4.010	0.019*	0.010*	0.839	0.020*

* $p < 0.05$, ** $p < 0.01$, two-tailed.

recommended in any subsequent assessment. Last but not the least, relying only on school records and teachers' feedback, but no rigorous and specific tests were conducted to identify and exclude ADHD and other implicit disorders from the sample, might potentially introduce some confounding effects on the present students' poor performance in Chinese handwriting legibility. Future work should strictly scrutinize the sample so as to avoid this phenomenon.

Notwithstanding these limitations, the CHLS is the first handwriting evaluation scale focusing on the detailed spatial aspects of the legibility dimension in the Chinese primary school context. Theoretically, the establishment of the CHLS can extend the Chinese handwriting theory and shed light on a deeper understanding of Chinese characters and Chinese handwriting, which may lay the foundation for further uncovering the underlying mechanism of their contribution to and linkage with spatial cognition. Practically, the CHLS may be a useful tool that can be used to evaluate the Chinese handwriting legibility performance of students with and without handwriting

difficulties. It has potential to favor analyzing and understanding the characteristics of students with handwriting difficulties from the perspective of spatial cognition, thus affording teachers and clinicians guidance on handwriting support and intervention accordingly.

7. Conclusion

Given the close relationship between Chinese handwriting and spatial cognition, and the spatial-oriented characteristics of the legibility dimension, we developed and validated a tailored assessment scale for Chinese handwriting legibility based on analyzing its spatial characteristics at both the analytic and holistic levels. With data from 684 Grades 2, 4, and 6 Chinese primary school students from central China, the CHLS showed good inter-rater reliability, and satisfactory construct validity and internal consistency reliability. The results of the measurement invariance confirmed the generalizability (in terms of factor structures and loadings) of the CHLS in these primary school

students in various grades. Our results also highlight the developmental specificity of these students' Chinese handwriting legibility performance with their growth.

What this paper adds

This paper demonstrates the close relationship between Chinese handwriting and spatial cognition and claims spatial-oriented characteristics of the legibility dimension; in this context, we developed and validated the Chinese Handwriting Legibility Scale (CHLS), a new tool focusing on a detailed analysis of spatial characteristics of Chinese handwriting legibility performance. Based on previous research on (Chinese) handwriting legibility and relevant scales and tools, the CHLS is established with spatial-related criteria at both the analytic and holistic levels. With 684 Chinese primary school students from Grades 2, 4, and 6 in central China as participants, the CHLS presented good inter-rater reliability, satisfactory construct validity and internal consistency reliability. This study also furthers the field with measurement invariance examination. The result, on the one hand, confirms the generalizability (in terms of factor structures and loadings) of the CHLS in Chinese primary school students in central China in various grades, on the other hand, highlights the developmental specificity of these students' Chinese handwriting legibility performance with their growth.

Data availability statement

The datasets presented in this article are not readily available because this dataset is one part of the first author's thesis study and cannot be shared because it involves confidential information of participants. Requests to access the datasets should be directed to HL, luhong2018@connect.hku.hk.

Ethics statement

The students' parents/legal guardians and schoolteachers and the school principal were informed of the study purpose and provided informed consents supporting the children's participation in the study. The students signed the assent form before the tests were administered. This study was ethically approved by the Human Research Ethics Committee of the University of Hong Kong.

References

- Amundson, S. J. (1995). *Evaluation tool of Children's handwriting*. Homer, AK: OT Kids.
- Asselborn, T., Gargot, T., Kidziński, L., Johal, W., Cohen, D., Jolly, C., et al. (2018). Automated human-level diagnosis of dysgraphia using a consumer tablet. *NPJ Dig. Med.* 1:42. doi: 10.1038/s41746-018-0049-x
- Barnett, A., Henderson, S., Scheib, B., and Schulz, J. (2007). *Detailed assessment of speed of handwriting*. Oxford: Harcourt Assessment.
- Barnett, A. L., Prunty, M., and Rosenblum, S. (2018). Development of the handwriting legibility scale (HLS): a preliminary examination of reliability and validity. *Res. Dev. Disabil.* 72, 240–247. doi: 10.1016/j.ridd.2017.11.013
- Barnett, A., Rosenblum, S., and Prunty, M. (2013). *Development of the handwriting legibility scale (HLS): An examination of reliability and validity*. Nara: International Graphonomics Society.
- Berninger, V. W. (2000). Development of language by hand and its connections with language by ear, mouth, and eye. *Top. Lang. Disord.* 20, 65–84. doi: 10.1097/00011363-200020040-00007
- Berninger, V. W., and Swanson, H. L. (1994). "Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writers" in *Advances in cognition and educational practice: Vol. 2 Children's writing: Toward a process theory of the development of skilled writing*. eds. J. S. Carlson and E. C. Butterfield (JAI: Greenwich, CT), 57–81.
- Berninger, V., Vaughan, K., Abbott, R., Begay, K., Byrd, K., Curtin, G., et al. (2002). Teaching spelling and composition alone and together: implications for the simple view of writing. *J. Educ. Psychol.* 94, 291–304. doi: 10.1037/0022-0663.94.2.291
- Bo, J., Colbert, A., Lee, C. M., Schaffert, J., Oswald, K., and Neill, R. (2014). Examining the relationship between motor assessments and handwriting consistency in children

Author contributions

HL and FL contributed to the conception and design of the study. HL organized the database and wrote the first draft of the manuscript. HL and XC performed the statistical analysis. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

Publication made possible in part by support from the HKU Libraries Open Access Author Fund sponsored by the HKU libraries.

Acknowledgments

We would like to express our gratitude to the students who participated in this study as well as to the experts and raters who assisted in developing the CHLS. We also want to thank the reviewers for their efforts and insightful comments to improve the quality and rigor of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1050894/full#supplementary-material>

with and without probable developmental coordination disorder. *Res. Dev. Disabil.* 35, 2035–2043. doi: 10.1016/j.ridd.2014.04.027

Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: comparative approaches to testing for the factorial validity of a measuring instrument. *Int. J. Test.* 1, 55–86. doi: 10.1207/S15327574IJT0101_4

Cascio, W. F., Alexander, R. A., and Barrett, G. V. (1988). Setting cutoff scores: legal, psychometric, and professional issues and guidelines. *Pers. Psychol.* 41, 1–24. doi: 10.1111/j.1744-6570.1988.tb00629.x

Chang, S. H., and Yu, N. Y. (2005). Evaluation and classification of types of Chinese handwriting deficits in elementary schoolchildren. *Percept. Mot. Skills* 101, 631–647. doi: 10.2466/pms.101.2.631-647

Chang, S. H., and Yu, N. Y. (2012). *Chinese handwriting evaluation form: Manual*. Taipei, Taiwan: Psychological Publishing.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14, 464–504. doi: 10.1080/10705510701301834

Chen, F. F., Sousa, K. H., and West, S. G. (2005). Teacher's corner: testing measurement invariance of second-order factor models. *Struct. Equ. Model.* 12, 471–492. doi: 10.1207/s15328007sem1203_7

Cheng-Lai, A., Li-Tsang, C. W., Chan, A. H., and Lo, A. G. (2013). Writing to dictation and handwriting performance among Chinese children with dyslexia: relationships with orthographic knowledge and perceptual-motor skills. *Res. Dev. Disabil.* 34, 3372–3383. doi: 10.1016/j.ridd.2013.06.039

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Chow, S. M., Choy, S. W., and Mui, S. K. (2003). Assessing handwriting speed of children biliterate in English and Chinese. *Percept. Mot. Skills* 96, 685–694. doi: 10.2466/pms.2003.96.2.685

Cui, Y., Zhu, Y., Laukkanen, H., and Rabin, J. (2012). Evaluation of visual-motor integration skills in preschool and elementary school-aged Chinese children. *J. Behav. Optim.* 23, 123–128.

Demetriou, A., Kui, Z. X., Spanoudis, G., Christou, C., Kyriakides, L., and Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or universal? *Intelligence* 33, 109–141. doi: 10.1016/j.intell.2004.10.003

Ellis, A. (1982). “Spelling and writing (and reading and speaking)” in *Normality and pathology in cognitive functions*. ed. A. Ellis (London: Academic Press), 113–148.

Erez, N., Yochman, A., and Parush, S. (1996). *The Hebrew handwriting evaluation*. Jerusalem: The Hebrew University of Jerusalem.

Feder, K. P., and Majnemer, A. (2007). Handwriting development, competency, and intervention. *Dev. Med. Child Neurol.* 49, 312–317. doi: 10.1111/j.1469-8749.2007.00312.x

Flaherty, M., and Connolly, M. (1995). Space perception, coordination and a knowledge of kanji in Japanese and non-Japanese. *Psychologia: an international J. Psychol. Orient.* 38, 229–237.

Gargot, T., Asselborn, T., Pellerin, H., Zammouri, I., Anzalone, M. S., Casteran, L., et al. (2020). Acquisition of handwriting in children with and without dysgraphia: a computational approach. *PLoS One* 15:e0237575. doi: 10.1371/journal.pone.0237575

Gilboa, Y., Josman, N., Fattal-Valevski, A., Toledano-Alhadeif, H., and Rosenblum, S. (2010). The handwriting performance of children with NF1. *Res. Dev. Disabil.* 31, 929–935. doi: 10.1016/j.ridd.2010.03.005

Graham, S., Berninger, V., Weintraub, N., and Schafer, W. (1998). Development of handwriting speed and legibility in grades 1–9. *J. Educ. Res.* 92, 42–52. doi: 10.1080/00220679809597574

Graham, S., Struck, M., Santoro, J., and Berninger, V. W. (2006). Dimensions of good and poor handwriting legibility in first and second graders: motor programs, visual-spatial arrangement, and letter formation parameter setting. *Dev. Neuropsychol.* 29, 43–60. doi: 10.1207/s15326942dn2901_4

Graham, S., and Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educ. Psychol. Rev.* 8, 7–87. doi: 10.1007/BF01761831

Hamstra-Bletz, L., and Blöte, A. W. (1993). A longitudinal study on dysgraphic handwriting in primary school. *J. Learn. Disabil.* 26, 689–699. doi: 10.1177/002221949302601007

Hamstra-Bletz, L., De Bie, J., and Den Brinker, B. P. L. M. (1987). *Beknopte beoordelingsmethode voor kinderhandschriften: Experimentele versie [Concise evaluation-scale for children's handwriting: Experimental version]*. Lisse: Swets & Zeitlinger.

Havaci, N., Azad, A., Alizade, Z. M., and Ebadi, A. (2017). Development and validity of the Persian handwriting assessment tool for primary school-aged children. *Crescent Med. J.* 19, 1–9. doi: 10.5812/ircmj.40508

Hellinckx, T., Roeyers, H., and Van Waelvelde, H. (2013). Predictors of handwriting in children with autism spectrum disorder. *Res. Autism Spectr. Disord.* 7, 176–186. doi: 10.1016/j.rasd.2012.08.009

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Kaiser, M. L., Albaret, J. M., and Doudin, P. A. (2009). Relationship between visual-motor integration, eye-hand coordination, and quality of handwriting. *J. Occupat. Ther. Schools Early Intervent.* 2, 87–95. doi: 10.1080/19411240903146228

Kandel, S., Lassus-Sangosse, D., Grosjacques, G., and Perret, C. (2017). The impact of developmental dyslexia and dysgraphia on movement production during word writing. *Cogn. Neuropsychol.* 34, 219–251. doi: 10.1080/02643294.2017.1389706

Kao, H. S. R. (2000). *Chinese calligraphic therapy*. Hong Kong: University of Hong Kong Press.

Kazi, S., Demetriou, A., Spanoudis, G., Zhang, X. K., and Wang, Y. (2012). Mind-culture interactions: how writing molds mental fluidity in early development. *Intelligence* 40, 622–637. doi: 10.1016/j.intell.2012.07.001

Klein, S., Guiltner, V., Sollereder, P., and Cui, Y. (2011). Relationships between fine-motor, visual-motor, and visual perception scores and handwriting legibility and speed. *Phys. Occup. Ther. Pediatr.* 31, 103–114. doi: 10.3109/01942638.2010.541753

Lai, M. Y. (2008). *An exploratory study into Chinese and English speaking children's visual perception and their spatial and geometric conceptions in Piagetian tasks*. HKU Theses Online (HKUTO).

Lam, S. S., Au, R. K., Leung, H. W., and Li-Tsang, C. W. (2011). Chinese handwriting performance of primary school children with dyslexia. *Res. Dev. Disabil.* 32, 1745–1756. doi: 10.1016/j.ridd.2011.03.001

Law, N., Ki, W. W., Chung, A. L. S., Ko, P. Y., and Lam, H. C. (1998). “Children's stroke sequence errors in writing Chinese characters” in *Cognitive processing of the Chinese and the Japanese languages* (Dordrecht: Springer), 113–138.

Lee, T. I., Howe, T. H., Chen, H. L., and Wang, T. N. (2016). Predicting handwriting legibility in Taiwanese elementary school children. *Am. J. Occup. Ther.* 70, 7006220020p1–7006220020p9. doi: 10.5014/ajot.2016.016865

Li, C., and Nuttall, R. (2001). Writing Chinese and mathematics achievement: a study with Chinese-American undergraduates. *Math. Educ. Res. J.* 13, 15–27. doi: 10.1007/BF03217096

Li, C., Nuttall, R. L., and Zhu, W. (1999). Writing Chinese characters and success on mental rotation test. *Percept. Mot. Skills* 88, 1261–1270. doi: 10.2466/pms.1999.88.3c.1261

Li, Y., O'Boyle, M. W., Liu, L., Zeng, X., Zhang, J., Zhu, J., et al. (2014). The influence of native acquisition of Chinese on mental rotation strategy preference: an EEG investigation. *Psychol. Rec.* 64, 321–328. doi: 10.1007/s40732-014-0028-9

Likhanov, M. V., Ismatullina, V. I., Fenin, A. Y., Wei, W., Rimpler, K., Maslennikova, E. P., et al. (2018). The factorial structure of spatial abilities in Russian and Chinese students. *Psychol. Russia* 11, 96–114. doi: 10.11621/pir.2018.0407

Linda, F. L., Thanapalan, K. C., and Chan, C. C. (2014). Visual-perceptual-kinesthetic inputs on influencing writing performances in children with handwriting difficulties. *Res. Dev. Disabil.* 35, 340–347. doi: 10.1016/j.ridd.2013.11.013

Li-Tsang, C. W., Au, R. K., Chan, M. H., Chan, L. W., Lau, G. M., Lo, T. K., et al. (2011). Handwriting characteristics among secondary students with and without physical disabilities: a study with a computerized tool. *Res. Dev. Disabil.* 32, 207–216. doi: 10.1016/j.ridd.2010.09.015

Li-Tsang, C. W., Li, T. M., Yang, C. N., Leung, H. W., and Zhang, E. W. (2022). Evaluating Chinese handwriting performance of primary school students using the mart handwriting analysis and recognition platform (SHARP). *medRxiv*. doi: 10.1101/2022.02.19.22270984

Li-Tsang, C. W., Wong, A. S., Leung, H. W., Cheng, J. S., Chiu, B. H., Linda, F. L., et al. (2013). Validation of the Chinese handwriting analysis system (CHAS) for primary school students in Hong Kong. *Res. Dev. Disabil.* 34, 2872–2883. doi: 10.1016/j.ridd.2013.05.048

Maeland, A. F. (1992). Handwriting and Perceptual-Motor Skills in Clumsy, Dysgraphic, and ‘Normal’ Children. *Perceptual and Motor Skills* 75, 1207–1217. doi: 10.2466/pms.1992.75.3f.1207

Marr, D., Windsor, M. M., and Cermak, S. (2001). Handwriting readiness: locatives and visuomotor skills in the kindergarten year. *Early Childhood Res. Pract.* 3, 1–17.

Mayes, S. D., and Calhoun, S. L. (2006). Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learn. Individ. Differ.* 16, 145–157. doi: 10.1016/j.lindif.2005.07.004

McBride, C. A. (2016). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educ. Psychol. Rev.* 28, 523–549. doi: 10.1007/s10648-015-9318-2

McCutchen, D. (2000). Knowledge acquisition, processing efficiency, and working memory: implications for a theory of writing. *Educ. Psychol.* 35, 13–23. doi: 10.1207/S15326985EP3501_3

Mekyska, J., Faundez-Zanuy, M., Mzourek, Z., Galaz, Z., Smekal, Z., and Rosenblum, S. (2016). Identification and rating of developmental dysgraphia by handwriting analysis. *IEEE Trans. Human Machine Syst.* 47, 235–248. doi: 10.1109/THMS.2016.2586605

Paglierini, E., Scocchia, L., Vernice, M., Zoppello, M., Balottin, U., Bouamama, S., et al. (2017). Children's first handwriting productions show a rhythmic structure. *Sci. Rep.* 7:5516. doi: 10.1038/s41598-017-05105-6

Parush, S., Lifshitz, N., Yochman, A., and Weintraub, N. (2010). Relationships between handwriting components and underlying perceptual-motor functions among students during copying and dictation tasks. *OTJR* 30, 39–48. doi: 10.3928/15394492-20091214-06

- Poon, W. Y., and Hong, P. M. (2003). *A study of the Chinese characters recommended for the subject of Chinese language in primary schools*. Kowloon Tong: Hong Kong Baptist University Language Centre.
- Purcell, J. J., Turkeltaub, P. E., Eden, G. F., and Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Front. Psychol.* 2:Article 239. doi: 10.3389/fpsyg.2011.00239
- Putnick, D. L., and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev. Rev.* 41, 71–90. doi: 10.1016/j.dr.2016.06.004
- Reisman, J. (2004). *Minnesota handwriting assessment*. San Antonio, TX: Psychological.
- Rosenblum, S., Aloni, T., and Josman, N. (2010). Relationships between handwriting performance and organizational abilities among children with and without dysgraphia: a preliminary study. *Res. Dev. Disabil.* 31, 502–509. doi: 10.1016/j.ridd.2009.10.016
- Rosenblum, S., Chevion, D., and Weiss, P. L. (2006). Using data visualization and signal processing to characterize the handwriting process. *Pediatr. Rehabil.* 9, 404–417. doi: 10.1080/13638490600667964
- Rosenblum, S., and Livneh-Zirinski, M. (2008). Handwriting process and product characteristics of children diagnosed with developmental coordination disorder. *Hum. Mov. Sci.* 27, 200–214. doi: 10.1016/j.humov.2008.02.011
- Roux, S., McKeef, T. J., Grosjacques, G., Afonso, O., and Kandel, S. (2013). The interaction between central and peripheral processes in handwriting production. *Cognition* 127, 235–241. doi: 10.1016/j.cognition.2012.12.009
- Sakamoto, M., and Spiers, M. V. (2014). Sex and cultural differences in spatial performance between Japanese and north Americans. *Arch. Sex. Behav.* 43, 483–491. doi: 10.1007/s10508-013-0232-8
- Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online* 43, 483–491. doi: 10.1007/s10508-013-0232-8
- Tan, L. H., Feng, C.-M., Fox, P. T., and Gao, J.-H. (2001). An fMRI study with written Chinese. *Brain Imag.* 12, 83–88. doi: 10.1097/00001756-200101220-00024
- Tan, L. H., Laird, A. R., Li, K., and Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Hum. Brain Mapp.* 25, 83–91. doi: 10.1002/hbm.20134
- Tomchek, S. D., and Schneck, C. M. (2006). “Evaluation of handwriting” in *Hand function in the child: Foundations for remediation*. eds. A. Henderson and C. Pehoski (MO Mosby Elsevier: St. Louis), 293–320.
- Tseng, M. H. (1993). Factorial validity of the Tseng handwriting problem checklist. *J. Occup. Ther. Assoc.* 11, 13–27.
- Tseng, M. H., and Chow, S. M. (2000). Perceptual-motor function of school-age children with slow handwriting speed. *Am. J. Occup. Ther.* 54, 83–88. doi: 10.5014/ajot.54.1.83
- Tseng, M. H., and Murray, E. A. (1994). Differences in perceptual-motor measures in children with good and poor handwriting. *Occup. Ther. J. Res.* 14, 19–36. doi: 10.1177/153944929401400102
- Van Hartingsveldt, M. J., Cup, E. H., Hendriks, J. C., de Vries, L., de Groot, I. J., and Nijhuis-van der Sanden, M. W. (2015). Predictive validity of kindergarten assessments on handwriting readiness. *Res. Dev. Disabil.* 36, 114–124. doi: 10.1016/j.ridd.2014.08.014
- Volman, M. J. M., van Schendel, B. M., and Jongmans, M. J. (2006). Handwriting difficulties in primary school children: a search for underlying mechanisms. *Am. J. Occup. Ther.* 60, 451–460. doi: 10.5014/ajot.60.4.451
- Weil, M. J., and Cunningham Amundson, S. J. (1994). Relationship between visuomotor and handwriting skills of children in kindergarten. *Am. J. Occup. Ther.* 48, 982–988. doi: 10.5014/ajot.48.11.982
- Wicki, W., Lichtsteiner, S. H., Geiger, A. S., and Müller, M. (2014). Handwriting fluency in children. *Swiss J. Psychol.* 73, 87–96. doi: 10.1024/1421-0185/a000127
- Zhang, Q., and Feng, C. (2017). The interaction between central and peripheral processing in Chinese handwritten production: evidence from the effect of lexicality and radical complexity. *Front. Psychol.* 8:334. doi: 10.3389/fpsyg.2017.0033



OPEN ACCESS

EDITED BY

Pedro García Guirao,
WSB Universities, Poland

REVIEWED BY

Judit García-Martin,
University of Salamanca, Spain
Moisés Selfa Sastre,
Universitat de Lleida, Spain
José Manuel de Amo Sánchez-Fortún,
University of Almería, Spain
Inmaculada Clotilde Santos Díaz,
University of Malaga, Spain

*CORRESPONDENCE

Jingjun Chen
✉ jingjunchenpsy@163.com

RECEIVED 30 January 2023

ACCEPTED 19 April 2023

PUBLISHED 11 August 2023

CITATION

Luo R, Peng Y and Chen J (2023) Internet use predicts Chinese character spelling performance of junior high school students: multiple mediating roles of pinyin input proficiency and net-speak experience. *Front. Psychol.* 14:1153763. doi: 10.3389/fpsyg.2023.1153763

COPYRIGHT

© 2023 Luo, Peng and Chen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Internet use predicts Chinese character spelling performance of junior high school students: multiple mediating roles of pinyin input proficiency and net-speak experience

Rong Luo, Yifan Peng and Jingjun Chen*

Department of Psychology, Hunan University of Science and Technology, Xiangtan, China

To examine the complex relationship between Internet use experience and character spelling performance among Chinese junior high school students, the study explored the multiple mediating roles of Pinyin input proficiency and net-speak experience. A total of 447 Chinese junior high school students aged 12–15 years old completed the Internet Use Experience and Pinyin Input Proficiency Assessment, the Net-speak Experience Questionnaire and the Chinese Spelling Test. The results showed that: (1) All investigated variables were significantly correlated with each other, but there was no direct relationship between Internet use and Chinese spelling performance. (2) Pinyin input proficiency and net-speak experience play a chain mediating role in the relationship between Internet use and Chinese character spelling performance. Teens' Internet use experience indirectly and positively predicted Chinese character spelling performance through the mediation of Pinyin input method use and net-speak experience. The implication of this study is that Chinese children should be guided to engage in Internet activities that require Pinyin typing and use net-speak creatively in order to promote the traditional Chinese character spelling skills when instructing teenagers to engage in Internet activities.

KEYWORDS

internet use, pinyin input proficiency, net-speak experience, multiple mediating, Chinese character spelling

Introduction

According to *The 48th China Statistical Report on Internet Development* released by China Internet Network Information Center, as of June 2021, the number of Internet users in China reached 1.011 billion, the number of mobile phone users reached 1.007 billion, and the number of instant message users reached 983 million, accounting for 97.3% of the total number of Internet users; the proportion of teenager Internet users using mobile phones reaches 90%, among which the number of Internet users aged 6 to 19 reached 185 million, accounting for 15.7% of the total number of Internet users, with the largest group of students of middle school, high school and other secondary education levels. It can be seen that, with the rapid development of modern information processing technology and network communication technology, the number of young Internet users is on the rise. With the popularization of the Internet, a series

of changes have taken place in the way of reading and writing among young people. In terms of reading, paper materials are gradually being replaced by electronic documents (or web pages); in terms of writing, handwriting is being replaced by keyboard input, and digital writing (such as Blog etc.) is becoming increasingly popular; in terms of information communication, Internet expression and communication (such as WeChat, QQ, etc.) is becoming widespread. As most of these activities are related to language activities (communication, reading, writing, etc.), does the change in the medium of reading and writing have an impact on the psychological processes of reading and writing, and therefore on children's traditional literacy development?

So far, a large number of studies have focused on the relationship between Internet use and reading, but few studies have focused on the relationship between Internet use and Chinese character spelling (Xu et al., 2020; Chen et al., 2021). For Chinese children, the relationship between Internet use and writing is complicated. According to the 2020 National Research Report on the Internet Use of Minors, 82.9% of Chinese underage Internet users have their own Internet access devices, with mobile phones being the most popular electronic devices. The data showed that the proportions of junior high school students studying online, chatting and using social networking sites are 90.7, 74.2 and 45.7% respectively, which are higher than the overall level of Internet users. When Chinese children engage in Internet activities, they will use Pinyin input method to type Chinese characters when it comes to doing online homework, sending messages, etc., which is completely different from English typing. Studies (Xu et al., 2020; Chen et al., 2021) show that, the use of Pinyin input method is beneficial to Chinese character spelling. In addition, Chinese children can also use net-speak in Internet communication. However, the form of Chinese net-speak is completely different from that of English. The latter is more of an abbreviated form of standard English, while the former is more of codified characters or Chinese standard characters that extend new meanings on the Internet. That is, old words are endowed with new implications (Chen et al., 2020). Studies have also shown that the use of net-speak seems to be beneficial to Chinese character processing (Chen et al., 2020), which implies that the use of Pinyin input method and net-speak style in Internet activities by Chinese children may be beneficial to their mastery of standardized Chinese character spelling.

A survey has shown that 97% of primary and middle school students use Pinyin input method in their Internet activities (Chen et al., 2016). Another survey also found that middle school students are most familiar with net-speak (Jin and Cao, 2010). Will the extensive use of Pinyin input and the familiarity of net-speak have a beneficial effect on the Chinese character spelling of middle school students? In addition, is it true that the more children use the Internet, the more proficient they are in Pinyin input and the more familiar they are with the net-speak, which in turn positively affects their standard writing level? This study hypothesizes a multiple mediation model to answer the above questions.

Literature review

Internet use and literacy

The Internet is the world's largest open network of computers of all sizes connected to each other. The Internet plays a different role in

the learning lives of young people, and its emergence has made a huge difference to the environment in which they grow up. Recent estimates show that one third of the world's young people under the age of 18 use the Internet, and in developed countries, 75% of young people play video games every day. The latest results from the US ESA show that young people spend more than 11 h a day on modern electronic media such as computers, the internet and video games, with 71% of young people under the age of 18 playing video games regularly. The under-18 group accounts for 24% of video game players (ESA, 2022).

Teenagers' internet use has a serious impact on their mental health, social, educational and emotional development (Kirmayer et al., 2013). Surveys have shown that the growing incidence of teenagers still being addicted to the Internet and that excessive Internet use makes teenagers depressed, anxious (Cataldo et al., 2021; Saadati et al., 2021; Lozano-Blasco et al., 2022). A range of mental problems such as loneliness and corresponding interpersonal troubles (Hao et al., 2022). However, as a new learning tool, information medium and cultural carrier, among other things, most young people are still currently profiting reasonably well from the Internet (de Sousa et al., 2022). Internet use has led to improved memory and attention span and more active thinking skills among teenagers (Casares and Ramirez, 2019). In addition, early experiences with online technologies may have a beneficial impact on children's later academic performance and digital skills as well as on their cognitive development (Hurwitz and Schmitt, 2020).

In terms of the relationship between Internet use and teenager literacy, researchers have focused more on the relationship between Internet use and reading, writing and spelling, and in general, the researches that have been conducted showed mixed results.

Internet use and reading

In terms of reading, researchers (Umbara et al., 2015) found a significant correlation between Internet use and reading comprehension levels ($r=0.2$). Petko et al. (2017) found that using ICT to finish homework (browsing the Internet to finish homework, using email to communicate with other students about homework, and doing homework on the computer) was positively associated with reading scores. Consistent with this finding, it was found that students who used ICT resources more frequently for leisure activities (e.g., playing online games, chatting online, reading news on the Internet) tended to perform better on reading tests (Hu et al., 2018). A survey on Mexican university students (Márquez et al., 2020) also found that students who used the Internet frequently not only read more digital texts, but also paper materials. Whereas a study on university students in Thailand (Anthonay et al., 2020) found no significant relationship between Internet use and reading attitudes, some studies have found negative effects: Casey et al. (2012) found that instant message was negatively associated with reading performance. In recent years, researchers (Borgonovi and Pokropek, 2021) have analyzed the relationship between Internet use and reading performance among students in 28 countries and found an inverted U-shaped association between different levels of Internet use and reading performance: students with low and high levels of Internet use have lower reading performance than students with moderate levels of Internet use. Overall, there was a positive effect of appropriate levels of Internet use on students' reading performance.

Internet use and writing

In terms of writing, researches have focused more on the impact of ICT use on writing. A study by the US Internet Research Center (Lenhart et al., 2008) on the relationship between teens' technology use and writing noted that although children often used the Internet to write, they did not consider email, instant messaging and short messages to be writing. In addition, both children and their parents felt that the use of technology had both positive and negative effects on writing. The positive one is that it was easier to write better because they can revise and edit easily, present ideas clearly, communicate well and they tend to be more creative. The negative effects include taking short cuts and not putting effort into writing, using wrong spelling and grammar, writing too fast and being careless, having a short attention span. Weiser et al. (2019) found that the reading, writing and spelling performance of students with reading learning disabilities was improved by giving teachers technology support. Among the technology supports were how to record and upload lessons and search for resources on the platform that would be beneficial to both teacher instruction and student engagement and learning. Fernández Batanero et al. (2021) reviewed literature from 2010 to 2020 that examined the use of ICT in writing and reading and found that although science and technology in this area was still underdeveloped, digital technology may contribute to the development of literacy skills.

Internet use and spelling

In terms of spelling, as digital writing is mostly to be found in children's Internet communication activities, the previous studies have focused more on the effects of using information communication (e.g., text message and instant message) in Internet use, with related studies showing inconsistent results. For example, some studies have shown no significant relationship between the frequency of Internet information technology-mediated communication and traditional spelling skills (Verheijen, 2013). Similarly, Zebroff and Kaufman (2017) also indicated that text messaging practices did not appear to be significantly associated with spelling among teenagers aged 15–17. While Vella (2014) found a negative effect of chatting on spelling performance among Maltese Secondary School students, generally the higher the frequency of chatting is, the lower the literacy score is. Chen et al. (2021) noted that, if just considering the frequency or time of IM/texting, there was a significant negative correlation between literacy skills and IM/texting. However, the use of textism, an abbreviated language style, in online communication appears to improve literacy skills, as Plester et al. (2009) found the positive correlation between the use of text message abbreviations and spelling ability among 88 British 10–12-year-old children. A study by Kemp and Bushnell (2011) of 86 Year 5–6 children in Australia also showed that text-messaging practices improved spelling skills. Mills (2014) also indicated that texting could be a literacy practice based on instructional purpose. In conclusion, the use of the Internet seems to be an "oxymoron" for English spelling, with the negative effects of excessive use on the one hand, and the positive effects of textism in the process on the other. This may be one of the main reasons why some studies have not found a direct relationship between Internet use and English spelling, and it is likely that Internet use is only a distal variable affecting spelling levels, while frequency of IM or texting and language style are intermediate variables.

Relationship between pinyin input, net-speak, and Chinese character spelling

Chinese Pinyin is a tool to assist in the pronunciation of Chinese characters. It is the official Latinization scheme for the pronunciation of Chinese characters promulgated by the People's Republic of China and is mainly used for the pronunciation of Mandarin Chinese characters as a kind of Mandarin phonetic symbol. Input method is a way of encoding various language symbols into a computer, mobile phone or other device, and is a writing tool that has been created by the development of human writing, breaking the limits of ink. Unlike English, which has the advantage of 26 Latin letters, Chinese characters are a combination of phonetics, morphology and semantics, and are basically typed using a combination of phonetics, morphology and semantics linked to specific keys, and then combined according to different Chinese characters. In other words, the Pinyin input method is used to input Chinese characters according to the Pinyin regulations.

Handwritten Chinese characters evolve from semantic activation, then orthography and speech automatically, and then glyphs are produced by means of the Peripheral Motor Process. Therefore, children need both phonetics and written morphology to write Chinese characters. Pinyin input involves both the process of speech production and the recognition of Chinese characters (Chen et al., 2017) by first typing Pinyin based on semantics and then selecting the desired character from alternative homophones. Numerous studies have found that pinyin input experience facilitates both phonological and orthographic processing, and that high proficiency in Pinyin input is beneficial for character recognition and further improves literacy performance. Zhu et al. (2009) found that Pinyin input experience facilitated phonological and orthographic processing through a study on the relationship between Pinyin input and literacy performance among university students, indicating that Pinyin input proficiency is positively related to individual literacy ability. Chen et al. (2016) invited primary school students in grade 6 to separately use two methods (handwriting and Pinyin input) to learn new characters and review newly-learned characters, and found no significant differences in the effects of the two methods. Further findings indicated that Pinyin input experience indeed strengthened the link between semantics and phonology, and did not weaken the link between semantics and orthography. Xu et al. (2020) found that junior high school students' use of Pinyin as a digital writing tool during Internet use could "shelter" its direct negative effect on Chinese character handwriting ability through the positive mediation of Pinyin input proficiency. Chen et al. (2021) also found that Pinyin input experience played a positive mediating role in the prediction of handwriting Chinese character performance by instant messaging (IM). This suggests that the use of Pinyin input method is beneficial to teens' Chinese character writing. The reason for this, according to the researcher (Chen et al., 2017), is that pinyin input involves first typing Pinyin based on semantics and then selecting the required Chinese character from alternative homophones, whereby the Pinyin input process involves the process of speech production as well as the process of Chinese character recognition, which means that it reinforces both Chinese phonology and orthography.

In terms of teens' net-speak usage style, English net-speak is mostly used in abbreviated form, called textism (Chen et al., 2020). As textisms are informal (unconventional spelling and grammatical shortcuts) and differ significantly from standardized English words or phrases in terms

of orthography and spelling rules, there is concern that the long-term use of textisms may have a negative impact on children's normative literacy skills. A number of studies have examined the relationship between children's use of textisms and normative literacy, but most of them have demonstrated that children who use textisms more often score higher on literacy-related assessments (Waldron et al., 2015; Blom et al., 2017); these literacy skills include: word reading (Plester et al., 2009; Coe and Oakhill, 2011), phonological awareness (Plester et al., 2009; Wood et al., 2014) and spelling (Plester et al., 2009; Bushnell et al., 2011; Wood et al., 2014). Most textisms appear to be homophonous variant of standard English vocabularies, which implies that, for net-speak experienced speakers, English net-speak is closely related to standard words in the individual's psychological vocabularies, so that high proficiency in English net-speak implies high proficiency in standard English. For Chinese net-speak, most of its forms are homographs, and the study (Chen et al., 2020) also found that the more experienced college students are in net-speak, the better their recognition and judgment of normative vocabulary are. In summary, studies using both English and Chinese as materials suggest that net-speak experience can benefit teens' literacy skills.

Internet chatting activities ranked second in teens' Internet use in China (CNNIC, 2021), which means that teens have a lot of opportunities to use net-speak in Internet communication, and as researchers have pointed out, most teens are not satisfied with normative words and use net-speak more frequently in their Internet communication (Grace et al., 2012; Drouin and Driver, 2014; Wood et al., 2014). Clearly, the more frequently the Internet is used and the more experienced the teens are in using it, especially with the frequent use of communication mediated by Internet information technology, the more experienced they are in net-speak. For Chinese teens, Chinese characters are the only ideographic system in the world that is still widely used and cannot be typed directly on a computer keyboard, and 90% of Chinese teens use Pinyin input methods to input Chinese characters in various Internet activities (Chen et al., 2021). Children overwhelmingly use Pinyin input method for both standardized Chinese words and Chinese net-speak. It is reasonable to believe that the more children use Pinyin input method in online communication, the more familiar they will become with net-speak and the more experienced they will be with net-speak.

Present study

The research objectives of this paper: Firstly, to identify the mechanisms by which teenagers' Internet use affects the traditional literacy skills. Secondly, to explore novel pathways for teenagers' use of the Internet for Chinese character learning.

Given that previous studies have suggested that the relationship between Internet use and English spelling seems to depend on certain intermediate variables (e.g., text message frequency or net-speak style), that Internet chatting is the most widespread non-educationally relevant activity for Chinese children, and that the study on the relationship between Internet use and character spelling among Chinese children is insufficient in the previous literature, this study hypothesizes, based on the previous analyzes, that Internet use positively affects character spelling performance through Pinyin input and net-speak experience. The present study hypothesized multiple mediating mechanisms through which Internet use positively affects spelling performance through Pinyin input and net-speak experience.

A survey method was used to collect data on junior high school students' Internet use, Pinyin input proficiency, net-speak experience and to test word spelling performance. Then a multi-mediation model was established to test the above hypothesis.

Method

Participants

Four hundred fifty-two teens aged between 12 and 15 were selected as participants by cluster sampling. Data were collected from 3 different junior middle schools in Hunan Province of China. 49.1% (222) are males and 49.8% (225) are females (gender data of 5 participants are missing). The numbers of students from Grade 7 to 9, account for 33.6, 31.2, and 35.2%, respectively, (shown in Table 1).

Instruments

Internet use experience

Two self-assessment questions were used to investigate teens' experience of using the Internet. One was "How long have you been online?" with five options ranging from 0 to 1 year, 1–4 years, 5–8 years, 9–12 years, and more than 13 years; the second was "If you were asked to rate how often you usually go online, how often do you think you do?," with options ranging from "rarely" to "very frequently" on a 5-point scale (shown in Table 2).

Pinyin input proficiency

Teens needed to self-assess their proficiency in Pinyin input (Chen et al., 2020). The question was: if the maximum score of proficiency and ability about using Pinyin method to input Chinese character is rated 10 points, what is your assessment of yourself? Participants needed to rate their Pinyin input level on a 10-point scale from 1 to 10 (shown in Table 2).

The net-speak experience questionnaire

The Net-Speak Experience Questionnaire (See Supplementary materials) developed by Chen et al. (2021) was used. The 50 net-speak words used in the questionnaire came from <http://wangci.net/> and are currently the most commonly used and most popular Chinese words. After an exploratory factor analysis following

TABLE 1 Demographic profile of respondents ($n=447$).

Variable	Measure	Frequency	Percentage (%)
Gender	Male	222	49.1
	Female	225	49.8
Grade	Seven	150	33.6
	Eight	141	31.2
	Nine	156	35.2

TABLE 2 Key variables and questionnaire items.

Variables	Source	Items	Measures
Internet use experience	Chen et al. (2020)	“How long have you been online?” “If you were asked to rate how often you usually go online, how often do you think you do?”	0–1 year, 1–4 years, 5–8 years, 9–12 years, and more than 13 years; “rarely” to “very frequently” on a 5-point scale.
Pinyin input proficiency	Chen et al. (2020)	if the maximum score of proficiency and ability about using Pinyin method to input Chinese character is rated 10 points, what is your assessment of yourself?	Pinyin input level on a 10-point scale from 1 to 10.
The net-speak experience questionnaire	Chen et al. (2021)	A selection of 23 internet language words were presented, such as “hold住”	Six options ranging from “I have never heard of it” to “very frequently used.”
Word spelling test	Chen et al. (2021)	In each word, one word was missing, “kè zhàn 客 ()”; in a sentence, the word to be spelt was missing, “His face yáng yì () had a heartfelt smile on it.”	The scoring system is one-word-one-point. The correct answers for the two examples are: “kè zhàn 客 (栈)” and, “His face yáng yì (洋溢) had a heartfelt smile on it”

TABLE 3 The means and standard deviations of various factors by gender and grade (*M*, *SD*).

	<i>M</i> (<i>SD</i>)		<i>M</i> (<i>SD</i>)		<i>M</i> (<i>SD</i>)	
Total <i>n</i> = 44	Grade 7		Grade 8		Grade 9	
	Male <i>n</i> = 67	Female <i>n</i> = 83	Male <i>n</i> = 72	Female <i>n</i> = 69	Male <i>n</i> = 83	Female <i>n</i> = 73
Internet use experience	4.75 (1.69)	3.99 (1.61)	5.19 (1.71)	4.07 (1.59)	5.59 (4.31)	4.23 (1.63)
Spelling performance	26.07 (16.25)	28.93 (15.37)	27.70 (17.60)	34.91 (15.04)	33.42 (17.25)	41.49 (14.85)
Pinyin input proficiency	5.73 (2.00)	5.81 (2.11)	6.42 (2.57)	6.52 (1.86)	6.38 (2.06)	6.66 (1.94)
Net-speak Experience	78.89 (24.93)	70.19 (25.30)	82.05 (25.67)	77.44 (21.70)	79.35 (23.39)	84.05 (24.14)

the initial test, 23 net-speak words were screened as material for the formal survey. For each word, a six-point rating scale was adopted and the participants were asked to choose from six options ranging from “I have never heard of it” to “very frequently used.” The Cronbach’s α of the questionnaire was 0.967 (shown in Table 2).

Word spelling test

Forty Chinese characters and 30 words marked with the correct Pinyin are used to test writing levels (See [Supplementary materials](#)). In every two-character word, one character was missing and participants were asked to fill in the blank such as “kè zhàn 客 ()”. Words are presented in sentences. In a complete sentence, the word to be spelt was missing, e.g., “他的脸上 yáng yì () 着会心的笑容.” The scoring system is one-word-one-point. The correlation coefficient was $r = 0.453$, $p < 0.001$, using the teens’ language scores as a calibration standard for examining word writing tests. It means that the criterion validity of writing Chinese character test is sufficient (shown in Table 2).

Results

Gender and grade differences

Gender and grade level may influence Internet use and Chinese character spelling performance. Boys and girls have different preferences for Internet use, with girls communicating more frequently online (Zhou, 2014; Zhu et al., 2017). Girls often express intimacy

through “words,” while boys build friendships through non-verbal means such as “games,” so girls’ use of social services on the Internet is more prominent (Lei and Liu, 2005). These findings seem to suggest that girls have more access to keyboarding and may perform better in Chinese character spelling. Lei and Liu (2005) surveyed 339 students in middle school and found that adolescent students’ preference for Internet social services increased with grade level, with a qualitative change in adolescents’ preference for Internet social services occurring in the Grade 2 of junior middle school and continuing to increase thereafter. Grade level reflects age, with the higher the grade level, the longer the exposure to the Internet, the more experience with digital writing activities, and the more practice with literacy skills.

MANOVA (multivariate analysis of variance) was adopted to analyze the effects of gender and grade on individual variables. The means and standard deviations are shown in Table 3.

The results of the multivariate ANOVA showed significant main effects for gender ($Wilks' \Lambda = 0.919$, $F = 9.70$, $p < 0.001$, $\eta^2 p = 0.081$) and grade ($Wilks' \Lambda = 0.915$, $F = 4.95$, $p < 0.001$, $\eta^2 p = 0.043$), but both the interaction effects were not significant. Gender differences in Internet use ($F = 17.41$, $p < 0.001$, $\eta^2 p = 0.038$) and Chinese character spelling performance ($F = 15.64$, $p < 0.001$, $\eta^2 p = 0.034$) were highly significant, with boys using the Internet significantly more than girls and girls outperforming boys in Chinese character spelling. Grade differences were also significant for Internet use ($F = 3.89$, $p = 0.021$, $\eta^2 p = 0.017$) and Chinese character spelling ($F = 14.80$, $p < 0.001$, $\eta^2 p = 0.063$), with an increasing trend from first to third year. We also found no significant gender differences in children’s ratings of Pinyin input proficiency and Net-speak experience, but there were significant grade differences ($F = 5.89$,

$p = 0.003$, $\eta^2 p = 0.026$; $F = 3.06$, $p = 0.048$, $\eta^2 p = 0.014$), which also showed an increasing trend with grade.

Partial correlation among internet use, pinyin input proficiency, net-speak experience and spelling performance

Based on the above results, we adopted partial correlation to examine the relationships among variables by controlling gender and grade. The results are shown in Table 4.

The results showed that all investigated variables were significantly correlated with each other, satisfying the prerequisites for a mediation analysis. In conjunction with the aforementioned hypothesis of a mediating process between Internet use and Chinese character spelling performance, a multiple mediated path model was developed and validated in this study.

TABLE 4 Partial correlation coefficients of variables among junior middle school students (*r*).

Variable	1	2	3	4
Internet experience	1			
Spelling performance	0.12**	1		
Pinyin input proficiency	0.23***	0.23***	1	
Net-speak Familiarity	0.34***	0.22***	0.31***	1

$N = 447$, ** $p = 0.012$, *** $p < 0.001$.

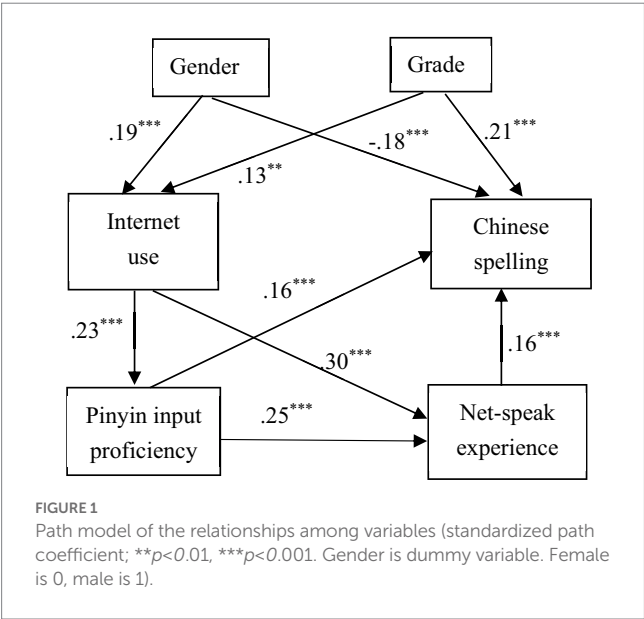


TABLE 5 Fitting indexes of path model.

χ^2	df	χ^2/df	GFI	CFI	NFI	IFI	RMSEA	SRMR
12.345	6	2.058	0.991	0.971	0.947	0.972	0.048	0.038

A path model of internet use predicting spelling performance

Amos software was used to fit for the path model and the model of the path analysis is shown in Figure 1.

The results of path analysis show that Internet experience has an indirect positive effect on the Chinese character spelling. Bootstrap was used to test the significance of the mediate effect. The number of bootstrap samples is 2000. The results show that standardized indirect effect is.098 and bias-corrected 95% confidence interval is [0.053, 0.157] which does not include 0. It means that the mediate effect of Pinyin input proficiency and net-speak experience is significant between teens' Internet use and Chinese character spelling performance. The fitting indexes are shown in Table 5.

Generally speaking, GFI, CFI, NFI and IFI should be higher than 0.9. RMSEA and SRMR should be lower than 0.05. The results of fitting indexes show that the path model is acceptable.

Discussion

The study found that teens' Internet use indirectly, significantly and positively predicted Chinese character spelling performance through the mediation of Pinyin input and net-speak experience. There was no direct relationship between Internet use and Chinese character spelling performance, which is consistent with the findings of some studies using English as materials (Verheijen, 2013; Zebroff and Kaufman, 2017). Furthermore, this finding demonstrates that Internet use is a distal variable predicting Chinese character spelling performance, acting on writing through Pinyin input and net-speak experience.

For 90% of Chinese children, Pinyin input method is used whenever Chinese character typing is involved in the use of the Internet and its ancillary devices, whether it is for writing, web searching, email, or chatting (Chen et al., 2020). Unlike English input where letters are typed directly on the keyboard, Chinese character cannot be typed directly in the character form, but need to be typed in the Pinyin form of the character, which matched the character form by the built-in character database of the input method software, and then the correct character is selected from a number of homophones and anagrams. For example, if the pinyin for “汉字” is “hàn zì” and you want to input the word “汉字,” you can type the letters “h-a-n-z-i” in order, the input method software may show “汉子”“汉字”“憨子”“汗渍” and other options. In this process, the typist activates both the phonetic sound of the Chinese character and receives feedback reinforcement of the character's form. Experimental studies have shown (Chen et al., 2017) that Pinyin input experience indeed strengthened the link between semantic and phonology, and did not weaken the link between semantics and orthography. Learning new Chinese characters by using the Pinyin input method likewise contributed to the mastery of Chinese handwriting (Chen et al., 2016). The

present study also confirms that Pinyin input experience facilitates the spelling of Chinese characters.

Chatting is one of the most popular online activities for children in China. The net-speak (textism) used by children in chatting often has a different style from standard language. English net-speak is generally abbreviated, using different forms of orthographic or phonological abbreviations (Waldron et al., 2015; Blom et al., 2017; Chen et al., 2020), such as contractions (e.g., msg for message), phonological abbreviations (e.g., thru for through), initialisms (e.g., omg for oh my God), shortenings (e.g., goin for going), single letters (e.g., u for you), combined letters (e.g., 2 day for today), and accent stylizations (e.g., gonna for going to). Plester et al. (2009) have pointed out that people experienced in net-speak themselves may have a high level of language sensitivity and ability, which is the reason why they are interested in net-speak words and use them creatively. In contrast, the form of Chinese net-speak is generally homograph, where traditional words take on new semantic meanings in the Internet context (e.g., a waistcoat originally meant a coat, but its extension in the Internet context means ID on the Internet). Proficiency in the use of net-speak is precisely based on proficiency in the standard language, and in turn, the use of net-speak reinforces the mastery of its standardized form. Research has found that Chinese net-speak experience facilitates the recognition and processing of Chinese standard words (Chen et al., 2020). In summary, studies of both Chinese and English language materials have found that the more frequently the net-speak is used, the more beneficial it is for the recognition of standard words (Plester et al., 2009; Kemp and Bushnell, 2011; Blom et al., 2017). The present study also confirms this finding.

In addition, this study also confirms that Internet use predicts Pinyin input proficiency and net-speak experience. This suggests that the more Chinese children use the Internet, the more proficient they become in Pinyin input, due to the fact that Pinyin input is a common typing software in children's use of the Internet. Therefore, the more frequently the Internet is used, the more proficient teenagers are in using Pinyin input. In turn, most of the activities that teenagers do when using Pinyin input are online chatting, which may be the reason why Pinyin input predicts net-speak experience.

Conclusion and implications

To sum up, this study found no direct relationship between Chinese teens' Internet use and Chinese character spelling, but rather indirectly predicted Chinese character spelling performance through Pinyin input proficiency and net-speak experience. This implies that for Chinese children, Internet use that requires the use of Pinyin input is helpful in mastering Chinese characters, and the use of net-speak in online social interaction is beneficial for Chinese character spelling. However, this does not mean that online chatting is helpful for spelling, and research has found that without the compensatory effect of Pinyin typing, the frequency of chatting activities such as IM is actually detrimental to Chinese character handwriting (Chen et al., 2021).

This suggests that when guiding young people to engage in online activities, if the aim is to promote their Chinese spelling skills, we need to encourage them to engage in some activities that require Pinyin typing, such as computer homework, writing blogs, emails, forum comments, etc., rather than some unrelated activities such as surfing,

playing games, watching videos, online shopping and other game-like activities. As for common chatting activities, on the one hand we need to limit their excessive frequency, and on the other hand, to promote text input chatting rather than voice chatting, and finally encourage creative use of net-speak in chatting.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by School of Education, Hunan University of Science and Technology. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

RL and JC wrote the first draft of the manuscript. YP performed material preparation and collection and analysis of the data. JC designed the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the MOE (Ministry of Education) in China Project of Humanity and Social Science under Grant Number 16YJCZH067.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1153763/full#supplementary-material>

References

- Anthoney, J., Kuttappan, J., and Cheewaparakobkit, P. (2020). Internet use and attitude towards reading among the students of an international university in Thailand. *Sikkha J. Educ.* 7, 1–11.
- Blom, E., van Dijk, C., Vasić, N., Van Witteloostuijn, M., and Avrutin, S. (2017). Texts and use of texting by children with typical language development and specific language impairment. *Comput. Hum. Behav.* 66, 42–51. doi: 10.1016/j.chb.2016.09.010
- Borgonovi, F., and Pokropek, M. (2021). The evolution of the association between ICT use and reading achievement in 28 countries. *Comput. Educ. Open* 2:100047. doi: 10.1016/j.cao.2021.100047
- Bushnell, C., Kemp, N., and Martin, F. H. (2011). Text-messaging practices and links to general spelling skill: a study of Australian children. *Aust. J. Educ. Dev. Psychol.* 11, 27–38.
- Casares, M. P., and Ramirez, E. A. (2019). Emotional intelligence assessment in primary and secondary education students. *Rev. Educ. Univ. Granada* 26, 53–68. doi: 10.30827/reugra.v26i0.118
- Casey, A., Layte, R., Lyons, S., and Silles, M. (2012). Home computer use and academic performance of nine-year-olds. *Oxf. Rev. Educ.* 38, 617–634. doi: 10.1080/03054985.2012.731207
- Cataldo, I., Lepri, B., Neoh, M. J. Y., and Esposito, G. (2021). Social media usage and development of psychiatric disorders in childhood and adolescence: a review. *Front. Psych.* 11:508595. doi: 10.3389/fpsyg.2020.508595
- Chen, J. J., Huang, S., and Luo, R. (2020). Does net-speak experience interfere with the processing of standard words? Evidence from net-speak word recognition and semantic decisions. *Front. Psychol.* 11:1932. doi: 10.3389/fpsyg.2020.01932
- Chen, J. J., Lei, X., Cheng, X. R., and Liu, H. S. (2016). Chinese character practice: comparison between children using hand-writing and pinyin keyboarding. *Acta Psychol. Sin.* 48, 1258–1269. doi: 10.3724/SP.J.1041.2016.01258
- Chen, J. J., Luo, R., and Liu, H. S. (2017). The effect of pinyin input experience on the link between semantic and phonology of Chinese character in digital writing. *J. Psycholinguist. Res.* 46, 923–934. doi: 10.1007/s10936-016-9470-y
- Chen, J. J., Xia, Y., Wan, L., and Luo, R. (2021). The effect of instant message on Chinese character handwriting among teens: pinyin input experience as a compensating role. *Asia Pac. Educ. Res.* 30, 177–185. doi: 10.1007/s40299-020-00524-y
- CNNIC. (2021). *The 48th China statistical report on internet development*. China Internet Network Information Center. Available at: <http://www.cnnic.net.cn/n4/2022/0401/c88-1132.html>.
- Coe, J. E., and Oakhill, J. V. (2011). 'txtN is ez fu no h2 rd': The relation between reading ability and text-messaging behaviour. *J. Comput. Assist. Learn.* 27, 4–17. doi: 10.1111/j.1365-2729.2010.00404.x
- de Sousa, D., Fogel, A., Azevedo, J., and Padrão, P. (2022). The effectiveness of web-based interventions to promote health behaviour change in adolescents: a systematic review. *Nutrients* 14:1258. doi: 10.3390/nu14061258
- Drouin, M., and Driver, B. (2014). Texting, textese and literacy abilities: a naturalistic study. *J. Res. Read.* 37, 250–267. doi: 10.1111/j.1467-9817.2012.01532.x
- ESA. (2022). *2022 essential facts about the video game industry*. Entertainment software association. Available at: <http://www.theesa.com>.
- Fernández Batanero, J. M., Montenegro Rueda, M., Fernández Cerero, J., and Román Gravan, P. (2021). Impact of ICT on writing and reading skills: a systematic review (2010–2020). *Texto Livre* 14:e34055. doi: 10.35699/1983-3652.2021.34055
- Grace, A., Kemp, N., Martin, F. H., and Parrila, R. (2012). Undergraduates' use of text messaging language: effects of country and collection method. *Writ. Syst. Res.* 4, 167–184. doi: 10.1080/17586801.2012.712875
- Hao, Q., Peng, W., Wang, J., Tu, Y., Li, H., and Zhu, T. (2022). The correlation between internet addiction and interpersonal relationship among teenagers and college students based on person's correlation coefficient: a systematic review and Meta-analysis. *Front. Psych.* 13:818494. doi: 10.3389/fpsyg.2022.818494
- Hu, X., Gong, Y., Lai, C., and Leung, F. K. S. (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: a multilevel analysis. *Comput. Educ.* 125, 1–13. doi: 10.1016/j.compedu.2018.05.021
- Hurwitz, L. B., and Schmitt, K. L. (2020). Can children benefit from early internet exposure? Short-and long-term links between internet use, digital skill, and academic performance. *Comput. Educ.* 146:103750. doi: 10.1016/j.compedu.2019.103750
- Jin, Y., and Cao, J. (2010). An empirical survey of internet language use and communication population. *Mod. Commun.* 11, 111–115. doi: 10.19997/j.cnki.xdc.2010.11.026
- Kemp, N., and Bushnell, C. (2011). Children's text messaging: abbreviations, input methods and links with literacy. *J. Comput. Assist. Learn.* 27, 18–27. doi: 10.1111/j.1365-2729.2010.00400.x
- Kirmayer, L. J., Raikhel, E., and Rahimi, S. (2013). Cultures of the internet: identity, community and mental health. *Transcult. Psychiatry* 50, 165–191. doi: 10.1177/1363461513490626
- Lei, L., and Liu, M. S. (2005). The relationship between adolescents' personality traits and preferences for internet social service use. *Acta Psychol. Sin.* 37, 797–802.
- Lenhart, A., Smith, A., Macgill, A. R., and Arafeh, S. (2008). *Writing, technology and teens*. Pew Internet & American Life Project. Available at: <http://apo.org.au/node/16746>
- Lozano-Blasco, R. L., Latorre-Martínez, M., and Cortés-Pascual, A. (2022). Screen addicts: a meta-analysis of internet addiction in adolescence. *Child Youth Serv. Rev.* 135:106373. doi: 10.1016/j.chilyouth.2022.106373
- Márquez, I. L. C., Flores Morales, C. R., Ordóñez Parada, A. I., and Sánchez Acosta, L. R. (2020). Digital natives: internet and its relationship with reading in university students. *Apertura* 12, 94–107. doi: 10.32870/Ap.v12n2.1876
- Mills, M. (2014). Analysis of research on the effects of texting and tweeting on literacy. In: Searson, M., Ochoa, M., (Eds.), *Proceedings of SITE 2014--Society for Information Technology & teacher education international conference*. Jacksonville, Florida, United States: Association for the Advancement of computing in education (AACE). pp. 1996–1997. Available at: <https://www.learnlib.org/primary/p/131080/>.
- Petko, D., Cantieni, A., and Prasse, D. (2017). Perceived quality of educational technology matters: a secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *J. Educ. Comput. Res.* 54, 1070–1091. doi: 10.1177/0735633116649373
- Plester, B., Wood, C., and Joshi, P. (2009). Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes. *Br. J. Dev. Psychol.* 27, 145–161. doi: 10.1348/026151008x320507
- Saadati, H. M., Mirzaei, H., Okhovat, B., and Khodamoradi, F. (2021). Association between internet addiction and loneliness across the world: a meta-analysis and systematic review. *SSM Popul. Health* 16:100948. doi: 10.1016/j.ssmph.2021.100948
- Umbara, R., Vianty, M., and Inderawati, R. (2015). The correlation among reading attitude, the use of internet and reading comprehension. *J. Engl. Lit. Educ.* 2, 77–88. doi: 10.36706/jele.v2i2.2316
- Vella, L. (2014). "Language used in internet chatting and its effect on spelling" in *Perspectives on Maltese linguistics*. eds. A. Borg, S. Caruana and A. Vella (München: Akademie Verlag), 319–340.
- Verheijen, L. (2013). The effects of text messaging and instant messaging on literacy. *Engl. Stud.* 94, 582–602. doi: 10.1080/0013838X.2013.795737
- Waldron, S., Kemp, N., Plester, B., and Wood, C. (2015). "Texting behavior and language skills in children and adults" in *The Wiley handbook of psychology, technology and society*. eds. L. D. Rosen, N. A. Cheever and L. M. Carrier (Hoboken, New Jersey: Wiley Blackwell), 232–249.
- Weiser, B., Buss, C., Sheils, A. P., Gallegos, E., and Murray, L. R. (2019). Expert reading coaching via technology: investigating the reading, writing, and spelling outcomes of students in grades K–8 experiencing significant reading learning disabilities. *Ann. Dyslexia* 69, 54–79. doi: 10.1007/s11881-018-00175-1
- Wood, C., Kemp, N., Waldron, S., and Hart, L. (2014). Grammatical understanding, literacy and text messaging in school children and undergraduate students: a concurrent analysis. *Comput. Educ.* 70, 281–290. doi: 10.1016/j.compedu.2013.09.003
- Xu, L., Chen, J., and Liu, H. (2020). The effect of junior middle school students' digital writing on Chinese character writing ability: pinyin input proficiency as mediating role. *Educ. Res. Exp.* 4, 86–90.
- Zebroff, D., and Kaufman, D. (2017). Texting, reading, and other daily habits associated with adolescents' literacy levels. *Educ. Inf. Technol.* 22, 2197–2216. doi: 10.1007/s10639-016-9544-3
- Zhou, X. (2014). An analysis of gender differences in internet use among college students. *Youth Quest* 1, 62–68. doi: 10.13583/j.cnki.issn1004-3780.2014.01.015
- Zhu, Z., Liu, L., Ding, G., and Peng, D. (2009). The influence of pinyin typewriting experience on orthographic and phonological processing of Chinese characters. *Acta Psychol. Sin.* 41, 785–792. doi: 10.3724/SP.J.1041.2009.00785
- Zhu, L., Tu, W., and Song, Q. (2017). An empirical study of contemporary college students' online behavior-based on differences in gender, place of birth and only child. *Glob. Educ.* 46, 117–128.



OPEN ACCESS

EDITED BY

Elena Jiménez-Pérez,
University of Malaga, Spain

REVIEWED BY

Mojtaba Rajabi,
Gonbad Kavous University, Iran
Laura García,
Universidad Europea del Atlántico, Spain

*CORRESPONDENCE

Lawrence Jun Zhang
✉ lj.zhang@auckland.ac.nz

RECEIVED 12 December 2022

ACCEPTED 02 June 2023

PUBLISHED 14 August 2023

CITATION

Xu TS and Zhang LJ (2023) Effects of cognitive task complexity and online planning on second language learners' argumentative writing. *Front. Psychol.* 14:1121994. doi: 10.3389/fpsyg.2023.1121994

COPYRIGHT

© 2023 Xu and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Effects of cognitive task complexity and online planning on second language learners' argumentative writing

Ting Sophia Xu^{1,2} and Lawrence Jun Zhang^{2*}

¹School of Literature and Media, Dongguan University of Technology, Dongguan, China, ²Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

Based on Kellogg's writing model, Skehan's Limited Attentional Capacity Model (LACM), and Robinson's Cognition Hypothesis, our study investigated the effects of cognitive task complexity on syntactic complexity, lexical complexity, accuracy, fluency, and functional adequacy in Chinese L2 students' argumentative writing, when students were under an online planning condition. Sixty-eight participants from a Chinese university were recruited to complete two writing tasks with task complexity varied in terms of [\pm argument elements]. The findings showed that increasing task complexity led to decreased subordination in terms of clauses per T-unit and dependent clauses per clause, increased phrasal elaboration in terms of coordinate phrases per clause, and no changes in mean length of T-unit, T-units per sentence, mean length of clause, and complex nominals per clause. Neither significant differences in accuracy nor fluency were found as a function of increasing task complexity. Detrimental effects on functional adequacy in content, organization, and overall scores were identified with the increases in task complexity. The trade-offs between syntactic and lexical complexity and between syntactic complexity and functional adequacy support the basic principle of Skehan's LACM that human's information processing capacity is limited and Kellogg's claim that learners have a limited central executive capacity in writing. Implications of the results of this research are discussed.

KEYWORDS

cognitive task complexity, online planning, EFL writing, argumentative writing, Limited Attentional Capacity Model, Cognition Hypothesis

Introduction

The cognitive processes of writing have been explored since the 1980s. Some scholars (e.g., Flower and Hayes, 1980; Bereiter and Scardamalia, 1987) proposed the writing models and identified the writing process in L1 writing, among which Kellogg's (1996) writing model is well-known for the initial prediction of the cognitive demands placed by writing processes concerning working memory. Three processes of formulation, execution, and monitoring have been distinguished during L1 writing in Kellogg's (1996) writing models, laying the foundation for understanding L2 writing processes. However, Kellogg's model neither considers the influences of task characteristics on writing processes nor predicts how learners' attentional resources are devoted to the writing processes as a function of varying task-generated cognitive demands.

In the context of the introduction of task-based language pedagogy, Skehan (1996a, 1998) and Robinson (2001, 2005, 2011a,b) respectively proposed two cognitive-interactionist models, the Limited Attentional Capacity Model and Cognition Hypothesis, to describe learners' information processing and predict how task design factors influence task complexity, learners' attentional allocation, and then learners' complexity, accuracy, and fluency performance during oral task completion. Inspired by the two models of the Limited Attentional Capacity Model and Cognition Hypothesis and task-based research studies in speaking, an increasing number of empirical studies related to task complexity were sparked in L2 writing (Kuiken et al., 2005; Kuiken and Vedder, 2007a, 2008; Ong and Zhang, 2010; Johnson et al., 2012; Ong and Zhang, 2013; Cho, 2015; Rahimi, 2019; Rahimi and Zhang, 2019; Zhan et al., 2021; Xu T. et al., 2022; Xu Z. et al., 2022). However, the findings, with regard to the effects of task complexity in L2 writing, have been far from conclusive. The various operationalizations of task characteristics might be one of the potential reasons for the mixed findings, highlighting the need for replication studies with consistent task complexity manipulation (Johnson, 2017).

Another possible reason for the contradictory results is that the theoretical underpinnings of the Limited Attentional Capacity Model and Cognition Hypothesis, that were proposed for oral tasks, may not be accurately appropriate for writing scope (Johnson et al., 2012). Thus, there is a pressing need to understand learners' information processing stages under the writing schemata by tying the cognitive task complexity models of Cognition Hypothesis and the Limited Attentional Capacity Model together with Kellogg's writing models. Some scholars have pioneered such connections in L2 task-based research. For example, Kormos (2011) posited that task complexity can be inherent in the formulation stage of the writing process and place cognitive demands in some writing processes at a time or in an independent stage of the writing process depending on task complexity manipulations. More research is warranted to study task complexity with reference to the cognitive writing process, which helps understand how different task features and implementation conditions vary the amount of attention available and the focus of learners' attentional resources when a message is expressed during composing.

Additionally, in the studies concerning the resource-dispersing features of task complexity, pre-task planning was the single most commonly investigated variable, with other planning types, such as online planning, scarcely being focused. Given pre-task planning and online planning impact different aspects of English as a foreign language (EFL) writing processes (Ellis and Yuan, 2004), online planning might play different roles when compared with pre-task planning in L2 writing when cognitive task complexity is changed. The involvement of different planning types, such as online planning, could add to our knowledge of where and how learners attend their attentional resources to different writing processes during task completion in L2 writing.

Informed by the gaps identified in the research literature, this study attempted to investigate the effects of cognitive task complexity on L2 learners' argumentative writing in terms of syntactic complexity, lexical complexity, accuracy, fluency, and functional adequacy, when learners were under an online planning condition. Our study chose Kellogg's (1996) writing model, Robinson's (2001, 2005, 2011a,b) Cognition Hypothesis, and Skehan's (1996a, 1998) Limited Attentional Capacity Model as the theoretical underpinnings to explain the complex nature

of writing processes in the task-based L2 writing research. Our study is anticipated to contribute to the knowledge of how different task features and implementation conditions influence L2 learners' cognitive resources allocation and then, their writing performance.

Literature review

Limited Attentional Capacity Model

The Limited Attentional Capacity Model was proposed based on the working memory theories (Gathercole and Baddeley, 1993; Carter, 1998), in its claim that learners "have a limited information processing capacity and must therefore prioritize where they allocate their attention" (Skehan, 2013, p. 189). When learners complete a cognitively demanding task and reach their attentional limits, trade-off effects will occur among the three dimensions of complexity, accuracy, and fluency (Skehan, 1996a; Skehan and Foster, 1999, 2001). Skehan (1996a, 1998) has suggested three principles to analyze the cognitive complexity of tasks: code complexity, cognitive complexity, and communicative stress. Code complexity is related to the linguistic demands of a task. A complex task is likely to require more advanced structures or greater densities of advanced structures. Cognitive complexity, including processing and familiarity, is concerned with the content or meaning of task performance. Processing refers to "the amount of on-line computation that is required while doing a task," and is closely related to "the extent to which the learner has to actively think through task content" (Skehan, 1996a, p. 52). Familiarity refers to the extent to which participants possess, and can readily use task-related schematic information or knowledge to complete the task at hand. The greater the familiarity, the lower the task complexity, whereas the greater the processing demand, the higher the task complexity. Communicative stress is about the performance conditions, resulting from differences in time pressure, modality, scale, and participant variables.

These three dimensions determine task complexity, consequently influencing learners' attentional allocation during task completion. For example, learners may become less willing to take risk of elaborating the interlanguage system and reduce their writing fluency when prioritizing accuracy in language performance. Learners' focus on complexity may increase "the chances that new forms will be incorporated into interlanguage systems, promote risk-taking, and require attention being devoted to the new forms of language" with low priority being attached to fluency and accuracy (Skehan, 1996a, p. 50). Learners' priority of fluency, such as emphasizing language accessibility, would be at the cost of the development of interlanguage system and the control of interlanguage system. Moreover, Skehan (2009) speculated that there might be competition for limited working memory between a high level of cognitive processing and the local linguistic form. For a task that demands more working memory to convey the content, learners' attentional resources will be less available for language forms, leading to poor language performance in complexity and accuracy, which is especially true for EFL learners.

Cognition Hypothesis

Robinson's (2001, 2005, 2011a,b) Cognition Hypothesis states that learners' attentional resources could be attached to language form to

process the input more deeply and elaborately with increased cognitive demands of tasks, as learners have multiple attentional resource pools. Robinson (2001, 2005, 2011a,b) claimed that cognitive complexity could be operationalized across the three triads of task complexity, task conditions, and task difficulty that interact to influence learners' learning and performance. Task complexity refers to the characteristics that elicit cognitive demands on learners. Two types of cognitive task features, resource-directing and -dispersing variables, comprise task complexity. The resource-directing variables, like few or many elements, [\pm Here-and-Now], past or present events, and fewer or more reasoning demands, place cognitive and conceptual demands on participants, directing learners' attention to particular aspects of the language code system and facilitating "the development and acquisition of new L2 form-concept mappings" (Robinson, 2007, p. 18). Resource-dispersing variables, like prior knowledge, planning time, and the number of tasks to complete, produce performative and procedural demands on learners' cognitive resources, promoting learners' fast real-time access to "an already established interlanguage system" (Robinson, 2007, p. 18). Increasing task complexity with resource-directing variables "has the potential to direct learners' attentional and memory resources to the way the L2 structures and codes concepts" (Robinson, 2011a, p. 15), thus negatively affecting fluency, but positively influencing accuracy and complexity. Increasing task complexity with the manipulation of resource-dispersing variables can promote learners' consolidation and automatic access to their existing interlanguage resources rather than the development and control of their interlanguage system.

Task condition includes participation and participant variables. Task difficulty refers to learners' perceptions of task-generated cognitive demands. Across task complexity, task conditions, and task difficulty, task complexity should be the sole basis for sequencing decisions, whereas task difficulty and task conditions can be used to inform the online decisions on how to implement tasks (Robinson, 2001, 2005, 2011a,b).

Kellogg's writing model

In Kellogg's (1996) writing model (see Figure 1), three different components have been distinguished in the writing process, with each

comprising two sub-processes: formulation (planning and translating), execution (programming and executing), and monitoring (reading and editing).

The formulation system plans the content and translates ideas into sentences; during planning, writers use their working memory to establish goals and retrieve ideas and knowledge to write. Translation is "the amalgam of linguistic processes needed to convert an idea into a written message" (Kellogg, 1996, p. 60); the lexical, syntactic, and rhetorical items are selected to encode ideas into words during translating. In the execution system, while programming refers to adopting the appropriate motor system (e.g., typing, handwriting, or dictating) to translate the output, executing occurs when the production is created into words and sentences based on the translating process. Monitoring refers to reading and editing the produced version to ensure the writer's intention is adequately expressed. During the editing, writers can either edit the localized errors, like spelling and diction, or the global problems, such as the paragraph and text organization issues.

Kellogg (1996) highlighted the significance of working memory capacity in writing. The central executive, a critical working memory component, is involved in all the sub-processes except executing. According to Kellogg (1996), learners have a limited capacity of central execution and have to compromise among three writing processes when under the pressure of completing cognitively demanding tasks, which is consistent with Skehan's Limited Attentional Capacity Model.

Empirical studies of task complexity

Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity Model, although originally proposed for L2 speaking, have been used by a large number of empirical studies to examine the impacts of task complexity in L2 writing. Availability of content support (Kormos, 2011; Révész et al., 2017; Abrams, 2019), the number of elements to be considered (Kuiken et al., 2005; Kuiken and Vedder, 2007a, 2008; Cho, 2015; Rahimi, 2019; Rahimi and Zhang, 2019; Zhang and Zhang, 2021; Xu et al., 2023), and [\pm Here-and-Now] (Ishikawa, 2007) are the resource-directing variables of task complexity that researchers have

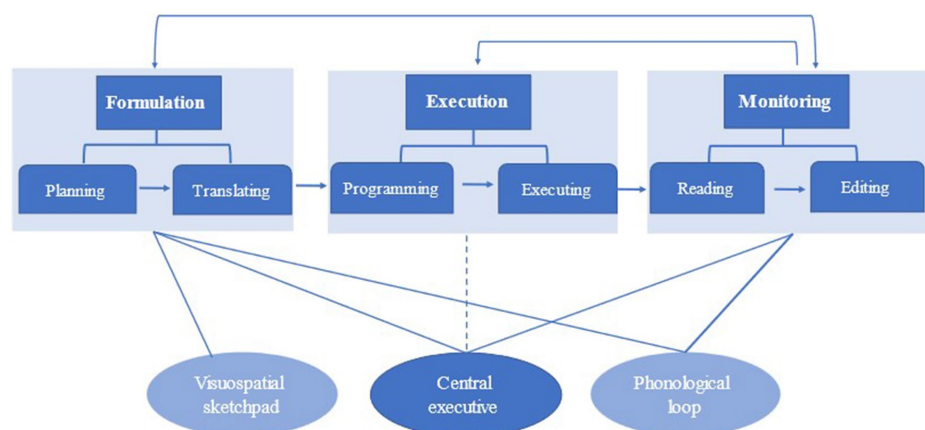


FIGURE 1

Kellogg's (1996) model of writing process, from Kellogg (2013). Reproduced by permission of Taylor and Frances Group, LLC, a division of Informa plc.

given a lot of attention to. In what follows, the studies, related to our research design, on the task complexity manipulated by changing the number of elements were reviewed.

Kuiken and his colleagues (Kuiken et al., 2005; Kuiken and Vedder, 2007b, 2008, 2012) investigated the effects of task complexity on L2 narrative letter writing in a set of studies (Kuiken et al., 2005; Kuiken and Vedder, 2007b, 2008, 2012). The complex tasks were to make a choice of a Bed and Breakfast from five options based on six criteria (six elements), while the simple tasks were about choosing a holiday resort according to three requirements (three elements). Participants' syntactic complexity, lexical variation, and accuracy were assessed. The four studies found that increases in task complexity positively affected learners' accuracy performance, but no significant effects on syntactic complexity were identified. Mixed findings of lexical performance were found in these studies.

To extend Kuiken and Vedder's (2007b, 2008, 2012) studies, Frear and Bitchener (2015) made low, medium, and high task complexity. The low complex task asked participants, based on their own resources, to introduce New Zealand to a friend in the letter, which was expected to have a limited cognitive demand. The medium and complex task required participants to choose a restaurant for visiting friend(s) based on different amount of information about each restaurant and the friend's requirements. Enhanced lexical complexity, but reduced ratio of adverbial dependent clauses to T-units were found as a function of increasing task complexity, indicating that participants may prioritize their limited cognitive resources to "the easier lexical means of meeting the pragmatic requirements of the tasks" over "the generation of grammar" (Frear and Bitchener, 2015, p. 53).

Later, Cho (2015) manipulated task complexity regarding more or fewer argument elements and varying reasoning demands within the topic of "choosing the best roommates" in a Korean EFL context. Participants' argumentative writing was measured by syntactic complexity, accuracy, and fluency in both global and task-specific indices. The results showed that increased task complexity did not influence accuracy and syntactic complexity, but positively affected fluency. The findings contradict the prediction of Cognition Hypothesis, but support Skehan's Limited Attentional Capacity Model. Cho (2015) argued that participants' prioritization of fluency, according to Skehan (1996a), may interfere with attention to the dimensions of accuracy and complexity.

Likewise, Rahimi (2019) studied how cognitive task complexity impacted argumentative writing in an Iran EFL learning context, but with a different topic of project funding allocation. In addition to syntactic complexity, lexical complexity, accuracy, and fluency, text quality in terms of content, organization and overall quality was also measured. The results demonstrated that the increased complexity of the task led to enhancements in syntactic complexity, lexical complexity, content, organization, and quality of writing, which partially supports the predictions of Cognition Hypothesis. A significant negative impact on accuracy was, however, found. The trade-off between complexity and accuracy is aligned with Skehan's Limited Attentional Capacity Model.

The relationship between planning and task complexity in L2 writing has been discussed in many studies (Ong and Zhang, 2010, 2013; Johnson et al., 2012). Planning, according to when it takes place, can be operationalized as pre-task planning and online planning (Ellis, 2005). Different planning types play various roles in learners' writing

process. For example, rehearsal, a type of pre-task planning, may assist writers in attending in both formulation and monitoring, leading to all-around improvements. Strategic planning, also a type of pre-task planning, may help learners mainly focus on conceptualization, which promotes better message complexity and increased fluency. Unpressured online planning may enable learners to devote attention to formulation and monitoring, contributing to a more accurate performance (Ellis, 2005). In the existing studies of task complexity, pre-task planning, rehearsal or strategic, has been increasingly studied, while there is a lack of research on the role of online planning in task complexity in L2 writing.

Ong and Zhang (2010, 2013) conducted a set of two studies to explore the effects of task complexity on L2 learners' argumentative writing production. Task complexity was manipulated by three independent variables: varying planning conditions (20 min extended pre-task planning, 10 min pre-task planning, 30 min free writing, and a control group), \pm ideas and/or macrostructure support (topic, ideas and macro-structure available, topic and ideas available, and topic available), and \pm draft given during revision.¹ Based on Robinson's theory, Ong and Zhang supposed that task complexity decreased from free writing, to pre-task planning and then to extended pre-task planning conditions with planning time before task increased from 0 to 20 min. Participants' lexical complexity, fluency, and writing quality were analyzed. The results showed increased task complexity contributed to significantly greater fluency, lexical complexity, and writing quality. In other words, the findings suggested that the provision of pre-task planning time might inhibit writers' writing performance, contrary to previous studies in the field of L1 writing (Kellogg, 1987, 1990) or prior findings of EFL oral production (Skehan and Foster, 1997).

Of note, free writing cannot be defined as online planning, although participants were required to start writing immediately and write continuously for 30 min (Ong and Zhang, 2010, 2013). The concept of free writing in Ong and Zhang's (2010, 2013) studies was adopted from Elbow's (1973, 1981) and Galbraith and Torrance's (2004) research. During free writing, learners were encouraged to do no planning before writing and write immediately and continuously what comes to their minds without considering how well the text is expressed or organized. In other words, continuous writing rather than online planning is what free writing emphasized, which was confirmed by Ong and Zhang's (2010, 2013) conclusion that writers in the extended pre-task and pre-task planning conditions were more engaged in online planning than those in the free-writing condition.

Motivated by the contradictory findings, Johnson et al. (2012) investigated the effect of different sub-processes of pre-task planning on argumentative writing in terms of grammatical complexity, lexical complexity, and fluency. Five groups, with each group under a sub-pre-planning condition, were set: Idea generation planning group asked learners to list ideas on the topic; organization planning group to make an outline of the main ideas and the supporting information; goal setting planning group to list rhetorical goals; goal setting and organization planning group to list goals and make an outline; and control group to complete word match to cover the planning time. Neither lexical complexity nor grammatical

¹ Revision conditions, unrelated to planning, are not reviewed in this part.

complexity was found due to the provision of pre-task planning. Some inconsequential effects on fluency were identified with organization planners making shorter sentences than their counterparts in the control group. These findings are in contrast not only to the negative findings that pre-task planning reduced L2 writing fluency, lexical complexity, and writing quality (Ong and Zhang, 2010, 2013), but also Skehan's Limited Attentional Capacity Model and Robinson's Cognition Hypothesis. Johnson et al. (2012) speculated that Limited Attentional Capacity Model and Cognition Hypothesis, which were proposed for oral tasks, may not be accurately appropriate for writing scope, and the distinction between writing and speaking may make the inconsistent findings in the studies on pre-task planning in the L2 writing field.

Later, Rahimi and Zhang (2018) explored how cognitive task complexity and pre-task planning together affect EFL writing. The pre-task planning group (provision of 10-min pre-task planning) and control group (absence of 10-min pre-task planning) completed two tasks with writing task complexity manipulated in terms of [\pm element]. The results showed that pre-task planning had no effects on accuracy or lexical complexity, but positive effects on fluency, syntactic complexity, content, and quality of writing regardless of task complexity. Increases in task complexity positively affected syntactic complexity, lexical complexity, content, and quality of writing, but negatively influenced accuracy and fluency regardless of planning conditions. Rahimi and Zhang (2018) concluded that simultaneously increasing task complexity and providing pre-task planning improved participants' syntactic complexity, content, organization, and quality of writing, which supports Robinson's Cognition Hypothesis, but is inconsistent with the previous findings (Ong and Zhang, 2010, 2013; Johnson et al., 2012).

Ellis and Yuan's (2004) study was one of the very few that investigated online planning in L2 writing, but not under the umbrella of cognitive task complexity. The effects of pre-task planning, unpressured online planning, and no planning on EFL narrative written performance were compared in their study. The results showed that pre-task planning led to greater fluency, while unpressured online planning resulted in better accuracy. In contrast, the no-planning condition negatively influenced fluency, complexity, and accuracy. The results suggest that different planning types affected participants' writing processes differently. Pre-task planning facilitated formulation process, while unpressured online planning assisted monitoring process. In contrast, writers with no-planning time had to deal with formulation, execution, and monitoring under pressure. The results, however, did not control the total time in each group, with pre-task planning group spending 27 min, online planning 21 min on average, and no-planning writers 17 min, which may act as confounding variables and consequently affect learners' task performance.

To sum up, with L2 writing being the focus of research, neither the results of studies on task complexity manipulated by [\pm few elements] (Kuiken et al., 2005; Kuiken and Vedder, 2007a, 2008; Cho, 2015; Rahimi, 2019; Rahimi and Zhang, 2019; Xu et al., 2023), nor on pre-task planning (Ong and Zhang, 2010, 2013; Johnson et al., 2012) have achieved consistency. The mixed findings indicated that the issue of how L2 learners allocate their attention when facing different levels of task complexity or under a planning condition during writing still remains to be resolved. The knowledge of the role of online planning, different from pre-task planning, in task complexity and L2 writing is very limited. These gaps support the need to explore how cognitive

task complexity, together with online planning, affects L2 learners' cognitive resources allocation and then writing performance.

The present study was guided by the research question as follows: What are the effects of cognitive task complexity on syntactic complexity, lexical complexity, accuracy, fluency, and functional adequacy, when L2 learners are provided online planning?

Methods

Participants

The participants were 68 L2 English learners from a university in China, and all of them regarded Chinese as their first language. Their ages ranged from 19 to 21, with a mean age of 20. Sixty-one of the participants were female, and 7 were male. The participants had an average of 11.23 years of English learning, and none had overseas learning experiences. All the participants had English as their major, and all courses were taught in English in their undergraduate program. At the time of the study, they were in their second-year study and had almost finished the required courses designed to develop their English competence in reading, writing, listening, and speaking, which means they were reaching an intermediate level of English proficiency, according to the requirements of Ministry of Education.

Task

The two argumentative writing tasks (see Supplementary Appendix), a simple one and a complex one, were based on Xu Z. et al. (2022), Xu et al. (2023) tasks, which adapted topics and prompts from Cho's (2015) study. The choice of "cooking" in Cho's (2015) task prompts was replaced by an alternative option of "playing football" to suit the Chinese context. The tasks were presented in Chinese to ensure all the participants could easily understand the instructions. Xu Z. et al. (2022), Xu et al. (2023) have independently validated the task complexity manipulations by using dual-task methodology, expert judgments, and post-task questionnaire and proven the efficacy of the tasks that the complex task was more cognitively demanding than the simple version. The topic of the tasks is to choose the best roommates. The complexity of the tasks was manipulated in terms of more or fewer argument elements and different reasoning demands. The simple task asked learners to decide on one pair of the best roommates from four options based on different information about hobbies, personalities, studying styles, and sleeping patterns. In the complex task version, participants were instructed to make decisions of the best roommates, choosing two pairs from six options based on varying information of hobbies, personalities, studying styles, sleeping patterns, favorite subjects, and individual sanitary habits. The complex task, with one more pair of the best roommates to decide and more information on each roommate candidate, cost more cognitive demands than the simple one (Cho, 2015; Xu Z. et al., 2022, Xu et al., 2023).

Procedure

Before the study, a pilot study was conducted to decide the writing time, where the slowest participants spent 35 and 37 min completing

TABLE 1 Counterbalance of the two writing tasks.

Participants	First time (week one)	Second time (week five)
Half ($n = 35$)	Task 1	Task 2
Half ($n = 33$)	Task 2	Task 1

The tag of “simple task” and “complex task” was labeled as Task 1 and Task 2.

the simple and complex tasks, respectively, when no time limits were set. Forty minutes of writing time was finally decided, as it, on the one hand, was enough for participants to complete the tasks in no rush. On the other hand, it could prevent students from feeling no challenge and being less willing to effectively extend any goal of complexity, accuracy, or fluency in the absence of time limits.

The two writing tasks were completed in a counterbalanced way to avoid potential practice and fatigue effects: half the participants took the simple task first, and half took the complex one first. There was a one-month interval between the two tasks. The timeline and counterbalance of the two tasks are presented in Table 1. During writing, participants were instructed to begin writing immediately and plan while they were writing. The researcher and a research assistant monitored and reminded the participants who did not start writing within 5 min. It was assumed that participants would have ample time for online planning while they were writing based on the pilot study.

Participants were given a free option to take part in a retrospective interview after the completion of the tasks, five of whom voluntarily participated in the interview that was conducted on the day after the second writing task. Interview questions concerned students’ planning and writing focus during the writing and planning process. Participants were interviewed individually in a comfortable place (an empty classroom). All the contents or questions were explained in Chinese. Students were informed that their responses were recorded and they had the rights to skip any questions and have the recorder switched off at any time.

Data coding and analysis

Analysis of writing performance

Syntactic complexity

To tap distinct and complementary dimensions of complexity, syntactic complexity was measured in four sub-constructs: overall complexity, subordination, phrasal elaboration, and coordination (Norris and Ortega, 2009). The overall complexity was assessed by the mean length of T-unit (MLT) that has consistently been used as a length-based measure (Frear and Bitchener, 2015; Ruiz-Funes, 2015; Yang et al., 2015; Yoon, 2017). Learners develop their ability to use phrasal-level complexification later than hypotaxis, such as subordination (Halliday and Mattiesen, 1999). Subordination, measured by clauses per T-unit (C/T) and dependent clauses per clause (DC/C), was considered a useful complexification index at an intermediate level. Phrasal-level elaboration, as a pervasive means to measure participants’ syntactic complexity, is an index of an advanced level in the field of writing (Norris and Ortega, 2009; Biber et al., 2011; Housen et al., 2019). Mean length of clause (MLC) was used for phrase-level measures, coordinate phrases per clause (CP/C) for phrasal coordination, and complex nominals per clause (CN/C) for

noun phrase complexity. Complexity via coordination, measured by T-unit per sentence (T/S), was included as participants in the present study were L2 writers, whose English proficiency was lower than native speakers (Norris and Ortega, 2009). These seven measurements were analyzed by using the EFL Syntactic Complexity Analyzer (Lu, 2010).

Lexical complexity

Lexical complexity was measured by lexical density, the ratio of lexical words, and lexical diversity, the range of vocabulary used. Lexical words are “nouns, adjectives, verbs (excluding modal verbs, auxiliary verbs, ‘be’ and ‘have’), and adverbs with an adjective base, including those that can function as both an adjective and adverb (e.g., ‘fast’) and those formed by attaching the *-ly* suffix to an adjectival root (e.g., ‘particularly’)” (Lu, 2012, p. 192). To take text length into account, $WT/\sqrt{2W^2}$ was used to measure lexical diversity (Ong and Zhang, 2010). The number of words, word types, and lexical words was counted automatically by the EFL Lexical Complexity Analyzer (Lu, 2012).

Accuracy

Three general accuracy measurements, the number of errors per T-unit (Err/T), the number of errors per 100 words (EP100), and the ratio of error-free clauses (EFC/C), were employed to examine learners’ overall ability in using the second language. Errors were coded and tailed based on Polio and Shea’s (2014) guidelines. An error was considered as any digression in lexical choice, syntax, and morphology but not in capitalization or punctuation (Ellis and Yuan, 2004).

Fluency. Fluency was determined by the average number of words produced per minute to capture learners’ ability to automatically access their linguistic resources.

Functional adequacy

Functional adequacy is about the “successful task fulfillment” (Kuiken and Vedder, 2017, p. 596). With a focus on whether the pragmatic goals have been achieved, functional adequacy is independent from the linguistic dimension, such as CAF (Pallotti, 2009; Kuiken and Vedder, 2017). Students’ functional adequacy in terms of content, organization, and overall score was scored based on a composition rating scheme adopted by Jacobs et al.’s (1981) and Kuiken and Vedder’s (2017) criteria. More information about the rating criteria can be found in the Xu et al.’s (2023) study.

Reliability of data coding and rating

The first author rated all the collected data for functional adequacy. A randomly selected 60% of the data were rated by a doctoral student majoring in applied linguistics, with the intraclass correlation coefficient for inter-rater reliability reaching 0.835 for the total score, 0.776 for the content, and 0.751 for the organization. As for the intra-rater reliability, the first author re-rated all the data 2 months apart, and the intraclass correlation coefficient was high,

2 WT=word type, W=the number of words.

0.881 for the total score, 0.837 for the content, and 0.790 for the organization.

With regard to errors, the first author coded all the data, and an over five-year experienced English writing lecturer coded 20% of the data separately. The identified errors marked by her and those identified by the first author were compared. The intraclass correlation coefficient reached 0.973.

Statistical analysis

Paired-samples *t*-tests were used to compare the writing performance between the simple and complex tasks when the normality of the measures was met. The alpha level of $p < 0.05$ was set for all measures. Cohen's *d* was reported as an indicator of the effect size (Cohen, 1992).

Analysis of interview

Information on the students' focus while writing, as well as planning, was gathered through the retrospective interviews. Based on participants' responses and previous research (Manchón and De Larios, 2007), their focus during planning and writing was defined into three broad categories: Content, organization, and language. The categories cover the answers in which the words "content, organization, or language" were explicitly mentioned (e.g., my main focus is content) and those that implicitly indicated a certain focus (e.g., "I made the decision for the roommates first"). It is acknowledged that a single statement may contain more than one point of focus and that the three categories are sometimes interwoven in students' statements. In such cases, the statement was coded according to the category represented in the statement. A single statement was, therefore, coded as more than one category. For example: "I will make an outline to figure out the logic, structure, and content." In this case, the focus point was coded as "organization" and "content," as the words "logic and structure" implicitly indicated the category "organization" and the word "content" was explicitly mentioned.

Students' focus during online planning and writing were analyzed quantitatively and qualitatively. Quantitatively, the interviewee's statements were coded, the number of the focus points provided in each statement was tallied, and the percentage of the three categories was calculated. The qualitative analysis mainly identified what students were doing during online planning and writing, thus providing explanations for the quantitative results.

Results

Effects on syntactic complexity

Table 2 displays the descriptive results for the seven syntactic complexity measures in the two writing tasks.

A series of paired-samples *t*-tests was conducted to investigate the effects of task complexity on syntactic complexity when participants were under the online planning condition. The results showed that no significant difference was found in MLT between the simple ($M = 12.42$, $SD = 2.47$) and complex tasks ($M = 11.82$, $SD = 2.52$), t

TABLE 2 Descriptive statistics for syntactic complexity.

Syntactic complexity	N	Simple task		Complex task	
		M	SD	M	SD
MLT	68	12.42	2.47	11.82	2.52
C/T	68	1.52	0.26	1.41	0.23
DC/C	68	0.33	0.10	0.28	0.10
T/S	68	1.10	0.09	1.11	0.10
MLC	68	8.17	0.88	8.38	1.14
CP/C	68	0.31	0.11	0.36	0.13
CN/C	68	0.85	0.19	0.84	0.23

MLT, mean length of T-unit; C/T, clauses per T-unit; DC/C, dependent clauses per clause; T/S, T-units per sentence; MLC, mean length of clause; CP/C, coordinate phrases per clause; CN/C, complex nominals per clause.

TABLE 3 Descriptive statistics for lexical complexity.

Lexical complexity	N	Simple task		Complex task	
		M	SD	M	SD
LD	68	0.509	0.03	0.512	0.03
WT/ $\sqrt{2}W$	68	5.56	0.52	5.40	0.53

LD, lexical density; WT, word type; W, the number of words.

($67 = 1.909$, $p = 0.061$, $d = 0.238$; in T/S between the simple ($M = 1.10$, $SD = 0.09$) and the complex tasks ($M = 1.11$, $SD = 0.10$), $t(67) = -0.528$, $p = 0.599$, $d = -0.08$; in MLC with the simple task ($M = 8.17$, $SD = 0.88$) versus the complexity task ($M = 8.38$, $SD = 1.14$), $t(67) = -1.535$, $p = 0.130$, $d = -0.206$; or in CN/C between the simple ($M = 0.85$, $SD = 0.19$) and complex tasks ($M = 0.84$, $SD = 0.23$), $t(67) = 0.427$, $p = 0.671$, $d = 0.054$. The effect sizes ranged as small.

Online planners used significantly more C/T in the simple task ($M = 1.52$, $SD = 0.26$) than in the complex task ($M = 1.41$, $SD = 0.23$), $t(67) = 3.252$, $p = 0.002$, $d = 0.458$, with the effect size close to medium. Likewise, students' DC/C performance for the simple task ($M = 0.33$, $SD = 0.10$) was significantly greater than for the complex task ($M = 0.28$, $SD = 0.10$), $t(67) = 3.228$, $p = 0.002$, $d = 0.446$, with the effect size close to medium. In contrast, online planners produced significantly more CP/C in the complex task ($M = 0.36$, $SD = 0.13$) than in the simple one ($M = 0.31$, $SD = 0.11$), $t(67) = -2.919$, $p = 0.005$, $d = -0.427$, and the effect size was close to medium. These findings suggest that increasing task complexity led to a significant reduction in participants' C/T and DC/C, but a marked increase in CP/C, with no significant influence on MLT, T/S, MLC, and CN/C.

Effects on lexical complexity

Table 3 displays the descriptive results for the online planners' lexical complexity performance in the two writing tasks.

Paired-samples *t*-tests revealed no significant differences in lexical density between the simple ($M = 0.509$, $SD = 0.03$) and complex tasks ($M = 0.512$, $SD = 0.03$), $t(67) = -0.740$, $p = 0.462$, $d = -0.097$, when participants were under the online planning condition. The effect size was small. However, online planners' WT/ $\sqrt{2}W$ production in the complex task ($M = 5.40$, $SD = 0.53$) was statistically lower than their performance in the simple task ($M = 5.56$, $SD = 0.52$), $t(67) = 2.960$,

TABLE 4 Descriptive statistics for accuracy.

Syntactic complexity	N	Simple task		Complex task	
		M	SD	M	SD
Err/T	68	0.74	0.29	0.73	0.37
EP100	68	0.060	0.02	0.061	0.03
EFC/C	68	0.63	0.12	0.63	0.15

Err/T, the number of errors per T-unit; EP100, the number of errors per 100 words; EFC/C, ratio of error-free clauses.

TABLE 5 Descriptive statistics for functional adequacy.

Syntactic complexity	N	Simple task		Complex task	
		M	SD	M	SD
Content	68	18.55	1.44	18.03	1.38
Organization	68	13.99	1.28	13.60	1.30
Overall score	68	64.38	4.05	63.14	4.48

$p = 0.004$, $d = 0.320$, with a small effect size. These results suggest that students' lexical diversity in terms of $WT/\sqrt{2}W$ was inhibited by increased task complexity.

Effects on accuracy

The descriptive results for the three accuracy measures in the simple and complex tasks were presented in Table 4.

No significant differences were detected in either Err/T between the simple ($M = 0.74$, $SD = 0.29$) and complex tasks ($M = 0.73$, $SD = 0.37$), $t(67) = 0.449$, $p = 0.665$, $d = 0.060$, in EP100 with the simple task ($M = 0.060$, $SD = 0.02$) versus the complex task ($M = 0.061$, $SD = 0.03$), $t(67) = -0.318$, $p = 0.752$, $d = -0.038$, or in EFC/C between the simple task ($M = 0.63$, $SD = 0.12$) and the complex task ($M = 0.63$, $SD = 0.15$), $t(67) = 0.105$, $p = 0.917$, $d = 0.014$. These findings indicated that increases in task complexity had negligible effects on online planners' accuracy performance.

Effects on fluency

A paired-samples t -test was applied to examine the effects of task complexity on fluency. The findings showed that students' fluency was not markedly different between their simple ($M = 5.88$, $SD = 0.92$) and their complex task completion ($M = 5.89$, $SD = 1.04$), $t(67) = -0.018$, $p = 0.985$, $d = -0.002$. The effect size was small.

Effects on functional adequacy

Table 5 presents the descriptive results for functional adequacy measures in the two writing tasks when participants were under the online planning condition.

Paired-samples t -tests revealed that the content of the online planners was better in the simple task ($M = 18.55$, $SD = 1.44$) than the complex one ($M = 18.03$, $SD = 1.38$), $t(67) = 3.341$, $p = 0.001$, $d = 0.370$, with a small effect size. The same pattern was found in organization

with scores in the simple task ($M = 13.99$, $SD = 1.28$) higher than those in the complex task ($M = 13.60$, $SD = 1.30$), $t(67) = 2.224$, $p = 0.030$, $d = 0.302$; the effect size was small. Likely, online planners' overall scores were outperformed in the simple task ($M = 64.38$, $SD = 4.05$) than the complex task ($M = 63.14$, $SD = 4.48$), $t(67) = 2.459$, $p = 0.017$, $d = 0.290$; effect size was small. These findings suggest that increasing task complexity had a detrimental effect on students' functional adequacy in content, organization, and overall score.

Interview

Four of the five interviewees reported they first spent a little time reading the instructions, quickly understanding the task requirements, and then started writing. Only one interviewee acknowledged that she made a general plan in her mind (maybe 5 min) after understanding the prompts and before beginning to write. All of them stated they had enough time to do online planning during writing, while none reported they had edited the essays after finishing writing. From the responses, it is apparent that the pre-task planning was limited, and online planning occurred during writing; this was the intention of the research design of this study.

With regard to the question of what they focused on during online planning, they mentioned content four times (30.77%), organization once (7.69%), and language eight times (61.54%). Online planners' main focus was language, followed by content, and a small proportion of time spent on organization.

Two participants (#1 and #3) gave a very general explanation, such as "language expression is my main focus," with very few about how they planned the language during online planning and writing. One participant (#4) stated that she carefully considered word choices and sentence structures during her online planning and writing. Another two (#2 and #5) reported they had a specific focus on language complexity, acknowledging that they tried to use more complex words and sentence patterns, as the following excerpts illustrate:

"I tried to search for some complex sentence patterns and sophisticated vocabulary, such as the words with more than three syllables, in my mind."—#2

"During writing, I mainly focused on language expression. I tried to use some complex sentence constructions and native expressions."—#5

It was interesting that neither the explicit words, "accuracy" and "grammar," nor the implicit expressions of accuracy were detected in the interview data, although most attention was paid to the language category.

Discussion

The results showed that increased task complexity led to a decrease in C/T and DC/C, and an increase in CP/C, although no significant influences on MLT, T/S, MLC, and CN/C. Ostensibly, the decline in subordination in terms of C/T and DC/C is contrary to Robinson's (2001, 2005, 2007) Cognition Hypothesis, since increasing

task complexity negatively affected the number of students' subordinate clauses. However, it should be taken into account that learners develop their ability to use phrasal-level complexification later than subordination (Halliday and Mattiesen, 1999). The increases in phrasal-level elaboration (i.e., CP/C) but decreases in subordination (i.e., C/T and DC/C) indicate that online planners in the complex task produced more advanced language, with "lower levels of subordination" but "more complex phrases" (Norris and Ortega, 2009, pp. 562–563) than in the simple task.

The favorable results of syntactic complexity in the complex task, essentially, echo Robinson's (2001, 2005, 2007) Cognition Hypothesis. As assumed by Robinson, learners' attention and effort are expected to be channeled to the language code system to meet the increasing conceptual demands, consequently facilitating interlanguage development, when task complexity is augmented with more elements and more reasoning demands involved. It is plausible that students, with the help of online planning and increased task complexity, might be encouraged to use new structures, or to extend their interlanguage system, to the next level of syntactic development in L2 writing to express their meanings. Evidence can be found in the interview data that online planners focused more on language, such as considering word choice and sentence structures and selecting complex syntactic frames to encode their ideas.

The findings of syntactic complexity also corroborate the results for phrase-level complexity reported in the previous studies (Biber et al., 2011; Lu, 2011; Bulté and Housen, 2014; Mazgutova and Kormos, 2015; Yoon, 2017), proving the sensitivity and validity of the phrase-level elaboration as an indicator of syntactic development in EFL writing. The results of our study highlight the necessity to use syntactic constructs in different dimensions to capture a multi-dimensional and dynamic picture of L2 learners' writing development (Norris and Ortega, 2009; Yang et al., 2015; Yoon, 2017).

Increasing task complexity reduced participants' lexical complexity production in terms of $WT/\sqrt{2}W$. This might be because the increased cognitive demands augmented participants' pressure on working memory to complete the task. Cognitive task complexity is supposed to be inherent in the formulation stage of L2 writing process (Kormos, 2011), and therefore competition for working memory within the *translating* sub-process (i.e., between selecting lexical units and building syntactic frames) may occur when task complexity is increased. In view of the findings with regard to syntactic complexity and lexical complexity, it seemed that fewer attentional resources were available for online planners to retrieve advanced lexical units with more working memory attended to the syntactic structures, when they completed the complex task. As a result, they were more likely to employ familiar or simple words to express messages in the complex task, compared to the simple task, resulting in a decline in lexical diversity but an increase in syntactic complexity, which was evidenced in participants' essays. For example, in a participant's essays, the word "personalities" has been repeatedly used in the complex task, while synonyms, like "traits," "personalities," and "characteristics" have been employed to convey meanings in the simple task. Our findings also extend Frear and Bitchener's (2015) hypothesis that there may be a trade-off between syntactic and lexical means of expression when learners experience pressure which is "brought to bear on limited attentional resources by cognitive task complexity" and pressure "of producing language that was not fully automatized" (p. 53).

No significant changes in any measures of accuracy were found between the simple and complex tasks. The results indicate that online planners devoted similar attentional resources to monitor linguistic accuracy when completing the simple and complex task versions. No emphases on accuracy in the interview might provide some evidence for the unnoticed differences in the production of accuracy between the simple and complex tasks. This finding is contrary to Skehan's LACM, as no trade-off between accuracy and complexity (either syntactic or lexical complexity) was identified.

Increasing task complexity resulted in a significant decrease in functional adequacy performance in terms of content, organization, and overall score. There appears to be a trade-off between syntactic complexity and functional adequacy, which lends support to the LACM, that learners' attentional resources are limited in capacity, although LACM does not make explicit predictions regarding the effect of task complexity on functional adequacy. Some scholars have suggested trade-off effects between learners' linguistic outcomes and higher-order dimensions of EFL production. For instance, Kuiken and Vedder (2008) posited that learners' favorable linguistic outcomes might be achieved with low priority being attended to the higher-order language processes. Pallotti (2009) speculated that learners might produce structures syntactically complex or accurate at the cost of lacking pragmatic adequacy. Similarly, Kormos (2011) also expressed a concern about the potential competition for the attentional resources of L2 writers, which may occur between syntactic encoding and text global organization during the writing process and between linguistic accuracy and discourse structure during monitoring process.

Kellogg's (1996) writing model seems to provide some insights into such trade-off effects. The trade-off between learners' linguistic output and higher-order performance may result from the pressure on working memory between the sub-processes of formulation (i.e., between *planning* and *translating*) that was triggered by some task implementation conditions, such as the increased task complexity and/or the provision of online planning. In our study, online planning seemed to help learners put the most attention (61.54%) to linguistic production during writing, with less attention devoted to content (30.77%) and organization (7.69%), as reflected in the responses in the interview. Increased task complexity may have greatly augmented online planners' pressure on working memory which has a limited capacity (Skehan, 1996b). When the competition between *planning* and *translating* became increasingly obvious with the increases in task complexity, online planners prioritized *translating* with inability to engage in planning the content and organization. In other words, online planners in our study were probably pressured to focus more on form at the expense of achieving semantic and pragmatic goals when the writing task was cognitively taxing.

A significant correlation was found between $WT/\sqrt{2}W$ and the overall score in the present study ($\rho=0.346$, $p<0.001$), which corroborates Lu's (2012) findings, when analyzing large-scale data from a corpus of Chinese learners, that Chinese students' lexical diversity is significantly correlated to the raters' judgment of the task quality. In light of this, the decline in the overall quality of the complex task might be related to the decreasing lexical diversity.

Increases in task complexity did not yield any significant differences in fluency. As previously discussed, online planners are likely to prioritize restructuring, which gets attention to the new forms of language, with accuracy and fluency being secondary, and adequacy possibly the last. This may explain why increasing the complexity of a

TABLE 6 Result patterns for the effects of increasing task complexity on EFL writing performance.

Syntactic	Accuracy	Lexical	Adequacy	Fluency
+	=	–	–	=

+ means that increased task complexity had a positive effect, = means a non-significant effect, – means a negative effect.

task has not led to any significant changes in learners' accuracy or fluency, but a significant decline in functional adequacy.

To conclude, the favorable findings in syntactic complexity, with increases in phrasal-level elaboration (i.e., CP/C) and decreases in subordination, resulted from the increased task complexity support Robinson's Cognition Hypothesis. However, given that writing is a cognitive activity with recursive processes (Kellogg, 1996; Weigle, 2002), the relationship between task complexity and writing should be considered with a dynamic and comprehensive view rather than a segmentary way. When we take the results of syntactic complexity, lexical complexity, accuracy, fluency, and functional adequacy as a whole (see Table 6), trade-offs between syntactic and lexical complexity and between syntactic complexity and functional adequacy were identified with the increased task complexity. The overall result pattern lends support to Skehan's (2013) LACM that "humans have a limited information processing capacity and must therefore prioritize where they allocate their attention" (p. 189). The results indicate that tensions may exist between syntactic and lexical complexity, and between form and pragmatic goals. It can be inferred that the tension in the formulation stage might be raised as a function of increasing task complexity, which probably leads to the competition for working memory between planning and translating, or within the translating process (i.e., between selecting lexical units and building syntactic frames). This interpretation, however, should be treated with caution, as think-aloud was not used to record the real-time data during online planning and writing.

Conclusion

This study aimed to explore the effects of cognitive task complexity on Chinese L2 students' argumentative writing in terms of syntactic complexity, lexical complexity, accuracy, fluency, and functional adequacy, when students were asked to start writing immediately and do online planning while writing. Increasing task complexity led to a trade-off between syntactic and lexical complexity, and between syntactic complexity and adequacy. No trade-off effect between complexity (either syntactic complexity or lexical complexity) and accuracy was found. The results indicate that tensions may exist between syntactic and lexical complexity, and between the local linguistic output (i.e., syntactic complexity) and the higher-order process (i.e., functional adequacy) when L2 students were completing the complex task under the online planning condition.

Theoretically, the overall trade-off result pattern lends support to the basic principle of Skehan's (2013) LACM that "humans have a limited information processing capacity and must therefore prioritize where they allocate their attention" (p. 189). Furthermore, this study contributes to the enrichment of empirical evidence and theory on cognitive task complexity by introducing functional adequacy, independent from CAF, into the research, which leads to an

understanding of the relationship between higher-order process and local linguistic output during the argumentative writing process.

When task complexity was increased for online planners, tensions between planning and translating sub-processes, and between lexical and syntactic encoding seemed to be triggered, resulting in the trade-offs between syntactic complexity and functional adequacy, and between lexical and syntactic complexity. These findings echo Kellogg's (1996) writing model that learners' central executive capacity is limited and also advanced our knowledge of how learners under the online planning condition compromised their working memory among the writing processes and sub-writing processes when under the pressure of completing the cognitively demanding argumentative tasks. Informed by Kormos (2011) speculation that task complexity is inherent in the formulation stage of writing, this study connected Kellogg's writing model with the cognitive-interactionist models, Robinson's Cognition Hypothesis and Skehan's LACM, to explore the complex nature of writing processes in the task-based L2 writing. The results provided empirical insights into understanding and analyzing cognitive task complexity in L2 argumentative writing with reference to the writing processes of formulation, execution, and monitoring.

Methodologically, our study highlighted the necessity of using multi-dimensional linguistic measures and functional adequacy to describe the task and proficiency-related variation. For example, indices measuring syntactic complexity from overall complexity, subordination, phrasal elaboration, and coordination proved to be very useful in capturing the effects of task complexity on syntactic complexity. If phrasal elaboration was not tested in the present study, the results of syntactic complexity would be interpreted in a contradictory way with the decline in subordination in terms of C/T and DC/C. Likewise, the involvement of functional adequacy helps researchers capture the trade-off between the form (i.e., syntactic complexity) and pragmatic goals. Also, the multi-dimensional writing measures could allow teachers to comprehensively evaluate L2 learners' writing.

Pedagogically, this study provided some guidance on making task-grading and sequencing decisions for EFL teachers when they schedule a syllabus for a given course considering the complexity of EFL argumentative writing (Zhang, 2021). Increasing task complexity could trigger pressure in the formulation stage of writing, and learners had the potential to devote more attention to language coding during online planning, which may assist the translating process primarily. Teachers may tailor task complexity to meet the learners' needs analysis, encouraging students to use new structures and extend the interlanguage system with the help of providing online planning strategies. The tension between the local linguistic output and the higher-order process found in the argumentative writing in our study should also attract teachers' attention. Teachers may design and sequence the argumentative tasks based on learners' proficiency to minimize the negative effects of the increased cognitive demands, promoting a balanced writing development in complexity, accuracy, fluency, and functional adequacy (Skehan, 1998; Robinson, 2010).

Limitations need to be mentioned for studies in the related field to consider in the future. First, participants recruited in our study all regarded Chinese as their first language and were from a single Chinese university, which limit the generalizability of the findings. Also, to ensure that all participants were, to a large extent, under online planning during writing, only 68 L2 learners were included. Further studies are encouraged to include a larger sample size of

participants from diverse L2 learning contexts, so that an understanding of the effects of cognitive task complexity and online planning on L2 learners' argumentative writing can be generalized and broadened.

Second, the research attempted to find out the effects of task complexity on L2 writing when participants were provided online planning. The result pattern under the online planning condition was found, but no control group was included in the study, making it impossible to compare students' performance under different planning conditions as a function of increasing task complexity. Further studies could involve different planning types, like pre-task planning and a control group. In this way, interactions between task complexity and planning conditions and the impacts of different planning types on writing processes could be explored, which helps advance the knowledge of how learners allocate their cognitive resources during the writing process.

Third, only the argumentative genre is included in the current study. In this way, findings may not apply to other types of writing beyond argumentative writing. The inclusion of various writing genres is warranted to explore the applicability and generalization of our findings to other types of writing.

Finally, qualitative data of what participants were doing and focusing on during online planning and writing were collected after the second task at one point in time. The retrospective interview cannot track what learners are actually doing during writing and online planning, as writing and planning are dynamic processes. Further research using the think-aloud approach is recommended to track all the real-time data (Rose et al., 2019; Zhang and Zhang, 2020), which helps deepen the understanding of how task complexity and online planning influence learners' attentional allocation during task completion.

The effects of cognitive task complexity on L2 writing need further exploration, with mixed findings identified in the existing studies. This research could be replicated in future research, as there are a limited number of studies either using multi-dimensional writing measurements or focusing on the relationship between task complexity and online planning in L2 writing. This study could also be extended by further cognitive task complexity related research, involving other variables, such as different proficiency levels, task types, and task modalities.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

References

- Abrams, Z. I. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task-complexity. *System* 81, 110–121. doi: 10.1016/j.system.2019.01.009
- Bereiter, C., and Scardamalia, M. (1987). *The psychology of written composition*. Mahwah, NJ: Lawrence Erlbaum.
- Biber, D., Gray, B., and Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Q.* 45, 5–35. doi: 10.5054/tq.2011.244483
- Bulté, B., and Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *J. Second. Lang. Writ.* 26, 42–65. doi: 10.1016/j.jslw.2014.09.005
- Carter, R. (1998). *Mapping the Mind*. London, England: Weidenfeld & Nicolson.
- Cho, H. (2015). Effects of task complexity on English argumentative writing. *English Teach.* 70, 107–131. doi: 10.15858/engtea.70.2.201506.107
- Cohen, J. (1992). Quantitative methods in psychology: a power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Elbow, P. (1981). *Writing with power*. New York: Oxford University Press.
- Elbow, P. (1973). *Writing without teachers*. London: Oxford University Press.
- Ellis, R. (2005). *Planning and Task Performance in a Second Language*. Benjamins, Amsterdam.

Ethics statement

The studies involving human participants were reviewed and approved by The University of Auckland Human Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TX and LZ contributed equally to the study in terms of research design, data collection, analysis, interpreting and writing it up, with LZ finalizing the manuscript for submission as the corresponding author. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a joint scholarship for doctoral study offered to TX by The University of Auckland of New Zealand and the China Scholarship Council of the Ministry of Education of China.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1121994/full#supplementary-material>

- Ellis, R., and Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Stud. Second. Lang. Acquis.* 26, 59–84. doi: 10.1017/S0272263104261034
- Flower, L., and Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College composition and communication*, 31, 21–32.
- Frear, M. W., and Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *J. Second. Lang. Writ.* 30, 45–57. doi: 10.1016/j.jslw.2015.08.009
- Galbraith, D., and Torrance, M. (2004). Revision in the context of different drafting strategies. In *Studies in writing: Vol 13. Revision: Cognitive and instructional processes*. eds. L. Allal, L. Chanquoy and P. Largy (Netherlands: Kluwer Academic Publishers), 63–85.
- Gathercole, S. E., and Baddeley, A. D. (1993). *Working Memory and Language*. Erlbaum. Mahwah, NJ
- Halliday, M. A., and Mattiesen, C. (1999). *Construing experience through meaning: A language-based approach to cognition*. London, England: Continuum International Publishing.
- Housen, A., De Clercq, B., Kuiken, F., and Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second. Lang. Res.* 35, 3–21. doi: 10.1177/0267658318809765
- Ishikawa, T. (2007). “The effect of manipulating task complexity along the [+Here-and-now] dimension on L2 written narrative discourse” in *Investigating Tasks in Formal Language Learning*. ed. M. D. P. G. Mayo (Bristol, England: Multilingual Matters), 136–156.
- Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V. F., and Hughey, J. (1981). *Testing EFL Composition: A Practical Approach*. Rowley, MA: Newbury House.
- Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: a research synthesis and meta-analysis. *J. Second. Lang. Writ.* 37, 13–38. doi: 10.1016/j.jslw.2017.06.001
- Johnson, M. D., Mercado, L., and Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *J. Second. Lang. Writ.* 21, 264–282. doi: 10.1016/j.jslw.2012.05.011
- Kellogg, R. T. (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes. *Mem. Cogn.* 15, 256–266. doi: 10.3758/BF03197724
- Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *Am. J. Psychol.* 103, 327–342. doi: 10.2307/1423213
- Kellogg, R. T. (1996). “A model of working memory in writing” in *The Science of Writing: Theories, Methods, Individual Differences, and Applications*. eds. C. M. Levy and S. Ransdell (Mahwah, NJ: Lawrence Erlbaum), 57–71.
- Kellogg, R. T. (2013). “A model of working memory in writing,” in *The Science of Writing: Theories, Methods, Individual Differences, and Applications*. eds. C. M. Levy and S. Ransdell (Routledge), 57–71.
- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *J. Second. Lang. Writ.* 20, 148–161. doi: 10.1016/j.jslw.2011.02.001
- Kuiken, F., Mos, M., and Vedder, I. (2005). Cognitive task complexity and second language writing performance. *EUROSLA Yearb.* 5, 195–222. doi: 10.1075/eurosla.5.10kui
- Kuiken, F., and Vedder, I. (2007b). Task complexity and measures of linguistic performance in L2 writing. *Int. Rev. Appl. Linguist. Teach.* 45, 261–284. doi: 10.1515/iral.2007.012
- Kuiken, F., and Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *J. Second. Lang. Writ.* 17, 48–60. doi: 10.1016/j.jslw.2007.08.003
- Kuiken, F., and Vedder, I. (2017). Functional adequacy in L2 writing: towards a new rating scale. *Lang. Test.* 34, 321–336. doi: 10.1177/0265532216663991
- Kuiken, F., and Vedder, I. (2012). “Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking” in *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. eds. A. Housen, F. Kuiken and I. Vedder (Amsterdam, Netherlands: John Benjamins), 143–170.
- Kuiken, F., and Vedder, I. (2007a). “Cognitive task complexity and linguistic performance in French L2 writing” in *Investigating Tasks in Formal Language Learning*. ed. M. D. P. G. Mayo (Bristol, England: Multilingual Matters), 117–135.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *Int. J. Corpus Linguist.* 15, 474–496. doi: 10.1075/ijcl.15.4.02lu
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers’ language development. *TESOL Q.* 45, 36–62. doi: 10.5054/tq.2011.240859
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners’ oral narratives. *Mod. Lang. J.* 96, 190–208. doi: 10.1111/j.1540-4781.2011.01232.1.x
- Manchón, R. M., and De Larios, J. R. (2007). On the temporal nature of planning in L1 and L2 composing. *Lang. Learn.* 57, 549–593. doi: 10.1111/j.1467-9922.2007.00428.x
- Mazgutova, D., and Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *J. Second. Lang. Writ.* 29, 3–15. doi: 10.1016/j.jslw.2015.06.004
- Norris, J. M., and Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Appl. Linguis.* 30, 555–578. doi: 10.1093/applin/amp044
- Ong, J., and Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students’ argumentative writing. *J. Second. Lang. Writ.* 19, 218–233. doi: 10.1016/j.jslw.2010.10.003
- Ong, J., and Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers’ text quality. *TESOL Q.* 47, 375–398. doi: 10.1002/tesq.55
- Pallotti, G. (2009). CAF: defining, refining and differentiating constructs. *Appl. Linguis.* 30, 590–601. doi: 10.1093/applin/amp045
- Polio, C., and Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *J. Second. Lang. Writ.* 26, 10–27. doi: 10.1016/j.jslw.2014.09.003
- Rahimi, M. (2019). Effects of increasing the degree of reasoning and the number of elements on L2 argumentative writing. *Lang. Teach. Res.* 23, 633–654. doi: 10.1177/1362168818761465
- Rahimi, M., and Zhang, L. J. (2018). Effects of task complexity and planning conditions on L2 argumentative writing production. *Discourse Process.* 55, 726–742. doi: 10.1080/0163853X.2017.1336042
- Rahimi, M., and Zhang, L. J. (2019). Writing task complexity, students’ motivational beliefs, anxiety and their writing production in English as a second language. *Read. Writ.* 32, 761–786. doi: 10.1080/0163853X.2017.1336042
- Révész, A., Kourtali, N. E., and Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Lang. Learn.* 67, 208–241. doi: 10.1111/lang.12205
- Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Appl. Linguis.* 22, 27–57. doi: 10.1093/applin/22.1.27
- Robinson, P. (2005). Cognitive complexity and task sequencing: studies in a componential framework for second language task design. *Int. Rev. Appl. Linguist. Lang. Teach.* 43, 1–32. doi: 10.1515/iral.2005.43.1.1
- Robinson, P. (2011b). Task based language learning: a review of issues. *Lang. Learn.* 61, 1–36. doi: 10.1111/j.1467-9922.2011.00641.x
- Robinson, P. (2007). “Criteria for classifying and sequencing pedagogic tasks” in *Investigating Tasks in Formal Language Learning*. ed. M. D. P. G. Mayo (Bristol, England: Multilingual Matters), 7–26.
- Robinson, P. (2010). “Situating and distributing cognition across task demand: the SSARC model of pedagogic task sequencing” in *Cognitive Processing in Second Language Acquisition: Inside the Learner’s Mind*. eds. M. Pütz and L. Scola (Amsterdam, Netherlands: Benjamins), 243–269.
- Robinson, P. (2011a). “Second language task complexity, the cognition hypothesis, language learning, and performance” in *Second Language task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. ed. P. Robinson (Amsterdam, Netherlands: Benjamins), 3–38.
- Rose, H., McKinley, J., and Baffoe-Djan, J. B. (2019). *Data collection Research Methods in Applied Linguistics*. London, England: Bloomsbury Publishing.
- Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: the effects of task factors and learner variables. *J. Second. Lang. Writ.* 28, 1–19. doi: 10.1016/j.jslw.2015.02.001
- Skehan, P. (1996a). A framework for the implementation of task-based instruction. *Appl. Linguis.* 17, 38–62. doi: 10.1093/applin/17.1.38
- Skehan, P. (1998). *A cognitive Approach to Language Learning*. Oxford, England: Oxford University Press.
- Skehan, P. (2009). Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. *Appl. Linguis.* 30, 510–532. doi: 10.1093/applin/amp047
- Skehan, P., and Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Lang. Teach. Res.* 1, 185–211. doi: 10.1177/136216889700100302
- Skehan, P., and Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Lang. Learn.* 49, 93–120. doi: 10.1111/1467-9922.00071
- Skehan, P., and Foster, P. (2001). “Cognition and tasks” in *Cognition and second language instruction*. ed. P. Robinson (Cambridge, England: Cambridge University Press), 183–205.
- Skehan, P. (1996b). “Second language acquisition research and task-based instruction” in *Challenge and Change in Language Teaching*. eds. J. Willis and D. Willis (London: Heinemann), 17–30.
- Skehan, P. (2013). “Tasks and language performance assessment” in *Research Pedagogic Tasks: Second Language Learning, Teaching, and Testing*. eds. M. Bygate, P. Skehan and M. Swain (London, England: Routledge), 167–185.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge, England: Cambridge University Press.
- Xu, T. S., Zhang, L. J., and Gaffney, J. S. (2022). Examining the relative effects of task complexity and cognitive demands on students’ writing in a second language. *Stud. Second. Lang. Acquis.* 44, 483–506. doi: 10.1017/S0272263121000310

- Xu, T. S., Zhang, L. J., and Gaffney, J. S. (2023). A multidimensional approach to assessing the effects of task complexity on L2 students' argumentative writing. *Assess. Writ.* 55, 100690–100616. doi: 10.1016/j.asw.2022.100690
- Xu, Z. Q., Zhang, L. J., and Parr, J. M. (2022). Incorporating peer feedback in writing instruction: examining its effects on Chinese English-as-a-foreign-language (EFL) learners' writing performance. *Int. Rev. Appl. Linguist. Lang. Teach.* 1–25. doi: 10.1515/iral-2021-0078
- Yang, W., Lu, X., and Weigle, S. C. (2015). Different topics, different discourse: relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *J. Second. Lang. Writ.* 28, 53–67. doi: 10.1016/j.jslw.2015.02.002
- Yoon, H. J. (2017). Linguistic complexity in L2 writing revisited: issues of topic, proficiency, and construct multidimensionality. *System* 66, 130–141. doi: 10.1016/j.system.2017.03.007
- Zhang, L. J. (2021). "Crossing literacy borders through writing: Transformational apprenticeship and repositioning of EAL learners" in *Crossing Borders, Writing Texts, Being Evaluated: Cultural and Disciplinary Norms in Academic Writing*. eds. A. Golden, L. A. Kulbrandstad and L. J. Zhang (Bristol, England: Multilingual Matters), 147–163.
- Zhang, L. J., and Zhang, D. (2020). "Think-aloud protocols" in *The Routledge Handbook of Research Methods in Applied Linguistics*. eds. J. McKinley and H. Rose (New York, NY: Routledge), 301–312.
- Zhang, T. T., and Zhang, L. J. (2021). Taking stock of a genre-based pedagogy: sustaining the development of EFL students' knowledge of the elements in argumentation and writing improvement. *Sustainability* 13, 1–25. doi: 10.3390/su132111616
- Zhan, J., Sun, Q., and Zhang, L. J. (2021). Effects of manipulating writing task complexity on learners' performance in completing vocabulary and syntactic tasks. *Lang. Teach. Res.* 1–22. doi: 10.1177/13621688211024360

Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

