

Insights in educational psychology 2021

Edited by

Douglas F. Kauffman, Claudio Longobardi
and Jesus de la Fuente

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 978-2-8325-3373-4
DOI 10.3389/978-2-8325-3373-4

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

Insights in educational psychology 2021

Topic editors

Douglas F. Kauffman — Medical University of the Americas — Nevis, Massachusetts, United States

Claudio Longobardi — University of Turin, Italy

Jesus de la Fuente — University of Navarra, Spain

Citation

Kauffman, D. F., Longobardi, C., de la Fuente, J., eds. (2023). *Insights in educational psychology 2021*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3373-4

Table of contents

- 07 **Developing Future-Ready University Graduates: Nurturing Wellbeing and Life Skills as Well as Academic Talent**
Tzyy Yang Gan, Zuhrah Beevi, Jasmine Low, Peter J. Lee and Deborah Ann Hall
- 18 **The Role of L2 Motivational Self System and Grit in EFL Learners' Willingness to Communicate: A Study of Public School vs. Private English Language Institute Learners**
Faramarz Ebn-Abbasi, Musa Nushi and Nazila Fattahi
- 30 **Identifying Determinants of Dyslexia: An Ultimate Attempt Using Machine Learning**
Sietske Walda, Fred Hasselman and Anna Bosman
- 49 **Determinants of Major Choice and Academic Expectations: Testing a Prediction Model Across Gender**
Sonia Alfonso, António M. Diniz, Angeles Conde and Mar García-Señorán
- 59 **Epistemic Beliefs and Learners' Self-Efficacy as Predictors of Language Learning Strategies: Toward Testing a Model**
Shaghayegh Shirzad, Hamed Barjesteh, Mahmood Dehqan and Mahboubeh Zare
- 74 **Exploring the Relations Among Teachers' Epistemic Theories, Work Engagement, Burnout and the Contemporary Challenges of the Teacher Profession**
Heidi Lammassaari, Lauri Hietajärvi, Katariina Salmela-Aro, Kai Hakkarainen and Kirsti Lonka
- 87 **Aiding Reflective Navigation in a Dynamic Information Landscape: A Challenge for Educational Psychology**
Katarzyna Bobrowicz, Areum Han, Jennifer Hausen and Samuel Greiff
- 94 **Corrigendum: Aiding Reflective Navigation in a Dynamic Information Landscape: A Challenge for Educational Psychology**
Katarzyna Bobrowicz, Areum Han, Jennifer Hausen and Samuel Greiff
- 95 **Investigating the Relationship Among English Language Learning Strategies, Language Achievement, and Attitude**
Anita Habók, Andrea Magyar and Gyöngyvér Molnár
- 104 **Recent Development in University Student Learning Research in Blended Course Designs: Combining Theory-Driven and Data-Driven Approaches**
Feifei Han
- 109 **The Proactive-Reactive Resilience as a Mediation Variable Between the Character Strength and the Flourishing in Undergraduate Students**
Jesús de la Fuente, Begoña Urien, Elkin O. Luis, María Carmen González-Torres, Raquel Artuch-Garde and Alvaro Balaguer

- 122 **Cyberbullying and Cyberhate as Two Interlinked Instances of Cyber-Aggression in Adolescence: A Systematic Review**
Giovanni Fulantelli, Davide Taibi, Lidia Scifo, Veronica Schwarze and Sabrina C. Eimler
- 143 **Understanding *Zhongyong* Using a *Zhongyong* Approach: Re-examining the Non-linear Relationship Between Creativity and the Confucian Doctrine of the Mean**
Ruixiang Gao, Shiqi Huang, Yujie Yao, Xiaoqin Liu, Yujun Zhou, Shijia Zhang, Shaohua Cai, Huang Zuo, Zehui Zhan and Lei Mo
- 158 **Influence of Mentorship and the Working Environment on English as a Foreign Language Teachers' Research Productivity: The Mediation Role of Research Motivation and Self-Efficacy**
Yanping Li and Lawrence Jun Zhang
- 175 **Massive Distance Education: Barriers and Challenges in Shifting to a Complete Online Learning Environment**
Ching-Yi Yeh and Chin-Chung Tsai
- 183 **Music Aptitude, Training, and Cognitive Transfer: A Mini-Review**
Lu Wang
- 190 **Views of Hospital Nurses and Nursing Students on Nursing Engagement—Bridging the Gap Through Communication Courses**
Qing Huang and Jack Pun
- 202 **Learning English as a Foreign Language Writing Skills in Collaborative Settings: A Cognitive Load Perspective**
Dayu Jiang and Slava Kalyuga
- 213 **The Changing Role of Chinese English-as-Foreign-Language Teachers in the Context of Curriculum Reform: Teachers' Understanding of Their New Role**
Man Lei and Jane Medwell
- 227 **English Language Teacher Agency in Response to Curriculum Reform in China: An Ecological Approach**
Lian Wang
- 240 **I can look for it! Modulation of a concurrent Visual Working Memory task in Visual Search in development**
María Quirós-Godoy, Beatriz Gil-Gómez de Liaño and Elena Perez-Hernandez
- 261 **Development of an Online Multilingual Educational Programme for Parents of Dual-Career Athletes: A Participatory Design**
Laura Capranica, Flavia Guidotti, Carlos Gonçalves, Laurence Blondel, Matteo Bovis, Rute Costa, Nadine Debois, Antonio Figueiredo, Ciaran MacDonncha, Viktorija Pecnikar-Oblak, Jean-Luc Patoret, Andrej Pišl, Eoin Rheinisch, Ana Rolo, Gary Ryan, Anne Templett, Antonio Tessitore, Giles Warrington and Mojca Doupona

- 270 **The Effects of Teacher Feedback and Automated Feedback on Cognitive and Psychological Aspects of Foreign Language Writing: A Mixed-Methods Research**
Zehua Wang and Feifei Han
- 282 **Moving to remote learning in adult education: Challenges and solutions of limited technological resources and capabilities**
Aydin Yücesan Durgunoğlu, Meltem Cantürk, Uğur Kaya, Betül Yazıcı and Kübranur Şahin
- 288 **Contributions of the psychology of mathematical cognition in early childhood education using apps**
Carlos Mera, Cándida Delgado, Estibaliz Aragón, Inmaculada Menacho, María Del Carmen Canto and José I. Navarro
- 299 **Augmented reality board game with multidimensional scaffolding mechanism: A potential new trend for effective organizational strategic planning training**
Huei-Tse Hou
- 304 **Examining EFL teachers' changing conceptions of research: A case study of a continuing professional development program in mainland China**
Yan Kang and Luxin Yang
- 319 **Examining the differences between information professional groups in perceiving information ethics: An analytic hierarchy process study**
Hsiu-Ping Yueh, Ching-Yin Huang and Weijane Lin
- 332 **Investigating teachers' experience and self-efficacy beliefs across gender in implementing the new standards-based curriculum in Ghana**
Edmond Kwesi Agormedah, Francis Ankomah, James Boadu Frimpong, Frank Quansah, Medina Srem-Sai, John Elvis Hagan Jr. and Thomas Schack
- 344 **Effects of Asian cultural values on parenting style and young children's perceived competence: A cross-sectional study**
Eunice Pui-Yu Yim
- 363 **Syntactic structure and rhetorical combinations of Iranian English research article titles in medicine and applied linguistics: A cross-disciplinary study**
Shadab Moslehi and Reza Kafipour
- 378 **Applying the mixed-blessings model and labeling theory to stigma in inclusive education: An experimental study of student and trainee teachers' perceptions of pupils with ADHD, DLD, and intellectual disability**
Alexander Röhm, Michelle Grengel, Michéle Möhring, Johannes Zensen-Möhring, Cosima Nellen and Matthias R. Hastall

- 392 **A proposed protocol for the registration of evidence-based Educational Psychology programs**
Jesús de la Fuente, Manuel Mariano Vera-Martínez,
Francisco Javier Peralta-Sánchez and José Manuel Martínez-Vicente
- 402 **Knowledge mapping of students' mental health status in the COVID-19 pandemic: A bibliometric study**
Yang Yang, Qingtai Cao, Mingyi Zhao and Quan Zhuang
- 414 **One more step in the study of children's daily stress: The spillover effect as the transfer of tension in family and school environments**
Lidia Infante-Cañete, Lidia Arias-Calero, Agustin Wallace-Ruiz,
Ana María Sánchez-Sánchez and Ángela Muñoz-Sánchez
- 426 **The stability of the twofold multidimensionality of academic self-concept: A study of Chinese secondary school students**
Feifei Han, Kateřina Juklová, Petr Mikoška and Lukáš Novák
- 441 **From entrepreneurship education to entrepreneurial intention: Mindset, motivation, and prior exposure**
Junhua Sun, Jingyi Shi and Junfeng Zhang



Developing Future-Ready University Graduates: Nurturing Wellbeing and Life Skills as Well as Academic Talent

Tzyy Yang Gan^{1†‡}, Zuhrah Beevi^{2†‡}, Jasmine Low^{1†}, Peter J. Lee^{1†} and Deborah Ann Hall^{2*†}

OPEN ACCESS

Edited by:

Jesus de La Fuente,
University of Navarra, Spain

Reviewed by:

Ana-Maria Cazan,
Transilvania University of Braşov,
Romania
Melissa Christine Davis,
Edith Cowan University, Australia

*Correspondence:

Deborah Ann Hall
deborah.hall@hw.ac.uk

†ORCID:

Tzyy Yang Gan
orcid.org/0000-0002-3079-7204
Zuhrah Beevi
orcid.org/0000-0001-6389-6746
Jasmine Low
orcid.org/0000-0003-2048-4818
Peter J. Lee
orcid.org/0000-0001-8403-5320
Deborah Ann Hall
orcid.org/0000-0002-3804-1452

‡These authors share first authorship

Specialty section:

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

Received: 02 December 2021

Accepted: 31 January 2022

Published: 04 March 2022

Citation:

Gan TY, Beevi Z, Low J, Lee PJ
and Hall DA (2022) Developing
Future-Ready University Graduates:
Nurturing Wellbeing and Life Skills as
Well as Academic Talent.
Front. Psychol. 13:827517.
doi: 10.3389/fpsyg.2022.827517

¹ Foundation Programmes, Heriot-Watt University Malaysia, Putrajaya, Malaysia, ² Department of Psychology, Heriot-Watt University Malaysia, Putrajaya, Malaysia

Higher education is starting to embrace its role in promoting student wellbeing and life skills, especially given the concerning levels of poor mental health and uncertainties in the future job market. Yet, many of the published studies evaluating positive educational teaching methods thus far are limited to interventions delivered to small student cohorts and/or imbedded within elective wellbeing courses, and are focussed on developed Western countries. This study addressed this gap by investigating the effectiveness of an institution-wide compulsory course informed by the principles of Seligman's Wellbeing Theory. The course was delivered at a British university in a developing country in Southeast Asia. It purposefully sought to nurture growth-oriented outcomes (including self-awareness, positive emotions, and personal effectiveness) and was taken by an entire cohort of year one undergraduate students. We tested the effectiveness of the curriculum content and staff coaching style in achieving life skills, and evaluated how these perceptions influenced students' subjective wellbeing. A convergent mixed-methods design was used with 350 survey respondents and 11 interviewees. Perceived life skills scores showed a 2.5% improvement at the end of the course. Partial Least Squares Structural Equation Modelling tested the predicted relationships between variables. All relationships were statistically significant, but the influence of course design and educators' style on life skills acquisition (50.8% of the variance) was moderate, while the effect on subjective happiness and life satisfaction (4–5% of the variance) was very weak. Qualitative data indicated that while quantifiable benefits to wellbeing might not be immediate, students did anticipate longer-term benefits for happiness and life satisfaction. This finding suggests that such a novel educational approach is well-received by Asian students and may sow the seeds for future benefit by positively impacting on their skills, behaviours, attitudes, and values. To achieve optimal flourishing at university, we recommend exploring teaching practises that combine positive education with coaching psychology practises.

Keywords: self-reflection and evaluation, leadership, happiness and wellbeing, positive psychology, positive education outcomes

INTRODUCTION

The COVID-19 global pandemic has accelerated a transformation in the skills required by today's workforce, as well as in the need for adaptability to change. Research by the McKinsey Global Institute identified self-leadership and inter-personal skills as two of the four skills categories essential to future-proof those graduates emerging into the job market (Dondi et al., 2021); the others being cognitive and digital skills. McKinsey's definition of self-leadership comprises elements of wellbeing (e.g., optimism, grit, and coping with uncertainty) and life skills (e.g., understanding one's strengths, driving change and achievement orientation). A wake-up call to university educators from this report was the conclusion that higher education is failing to effectively develop graduates with proficiency in self-leadership skills, despite these skills being most closely associated with employment, income, and job satisfaction.

Relevant to the teaching of self-leadership is the branch of psychological theory and practise known as Positive Psychology; the science of wellbeing and flourishing (e.g., Seligman, 2011; Wong and Roy, 2017; Lomas et al., 2020). For the past 10 years or so, a number of positive psychologists have advocated teaching students wellbeing and life skills, in addition to traditional academic skills (c.f. Positive Education; Seligman et al., 2009). The goal of this positive education approach is to improve students' wellbeing *and* learning through activities that explicitly promote components of wellbeing. Seligman's Wellbeing Theory underlies one of the most common approaches by teaching positive emotion, engagement (i.e., being in the flow), positive relationships, meaning (i.e., purpose in life), and achievement (Seligman, 2011). While there are numerous examples of positive education in schools (e.g., Norrish et al., 2013), positive education has had fewer inroads in the higher education setting despite universities being entrusted to prepare generations of youths for a better future. There are a growing number of case studies. For example, at Melbourne University Centre for Positive Psychology, Australia, a 6-week University Wellbeing Programme for undergraduates in a Positive Psychology course led to improvements in mental health and seemed to buffer against stress, relative to a control group (Young et al., 2020). In the United Kingdom, an optional 12-week psychoeducational happiness course for undergraduates at the University of Bristol improved student mental wellbeing, even during the COVID-19 pandemic (Hood et al., 2021). However, of those studies conducted at university, many are limited to evaluating interventions delivered to small student cohorts and/or imbedded within elective wellbeing courses. Hence, their generalisability is somewhat reduced.

A number of specific challenges in implementing positive education principles within the general university curriculum have been recognised (e.g., King et al., 2016). The university environment itself may provide one such tension. Positive education requires a cultural shift away from the university as an institution of knowledge production and teaching delivery, toward one of fostering self-discovery and personal growth. To achieve those goals requires a coaching style that nurtures student potential and talent. Yet, the traditional university system emphasises one-way communication from teacher to student and

independent learning (Simel and Bogner, 2017). Here, leadership from the senior management is needed to push the organisation and its staff in the right direction.

One challenge coming from within the positive education field itself concerns the geographical bias in research outputs toward Western, Educated, Industrialised, Rich and Democratic (WEIRD) countries. The reasons for this are complex (e.g., Wong, 2020). Nevertheless, this paucity of research is especially awkward given that culture is well-known to play a key role in wellbeing and learning. For example, a study of Singaporean secondary school students found that, unlike their Western counterparts, they were equally strongly motivated to study by external factors (e.g., the need to show parents that they are achieving) as well as intrinsic factors (e.g., personal satisfaction and an inherent desire to learn) (Caleon et al., 2015). The authors reasoned that this finding reflects those cultural traditions which put primacy on learning as a means for social mobility and economic survival; a characteristic that is shared by many developing countries in Asia.

Important questions therefore remain whether wellbeing and life skills can be attained in a university culture that embraces positive education across the whole institution and that attempts to offer a coaching style of teaching, and whether a curriculum informed by a WEIRD-centric conceptual framework for wellbeing can have impact in the setting of a developing Asian country. Since 2018, Heriot-Watt University Malaysia has offered students a transformative educational experience through its compulsory year one 'EmPOWER' programme. This coaching programme supports self-discovery and personal growth, accentuates purpose and offers opportunities to achieve a positive impact on the community and society, with achievements being explicitly articulated to future employers through an enhanced transcript.

The Current Study

The purpose of this study was to evaluate this programme using a convergent parallel mixed-methods design, gathering quantitative online survey and qualitative interview data. We tested the hypotheses that (i) the course would improve students' wellbeing in accordance with Wellbeing Theory, (ii) the course would improve students' life skills (namely sense of purpose, self-reflection, self-awareness, leadership, inter-personal communication and positive emotions), in line with the intended positive education learning outcomes, (iii) that the course design and educator style would each positively influence students' learning of those life skills, and (iv) that achievement in life skills would in turn drive greater wellbeing as measured by subjective happiness and life satisfaction. The purpose of the interview data was to give greater insight into how the positive education course influenced students' perceptions of their learning.

MATERIALS AND METHODS

Participants

Invited participants were all year one students ($n = 570$) enrolled in undergraduate programme at Heriot-Watt University Malaysia in the academic year 2019/20. Participants were eligible

to take part in the study if their overall attendance was at least 80%. Survey data was pseudo-anonymised to enable matching responses to attendance records, and was fully anonymised through attribute suppression before data analysis. Although 407 students completed the survey, 57 were excluded because they did not meet the eligibility criterion regarding attendance: leaving 350 survey respondents at T3. From these, 104 did not complete the SHS and SWLS at T2. Analysis of the changes over time handled these missing data using a pairwise deletion method ($n = 246$).

Survey respondents were invited to volunteer for a short follow-up interview at T3 ($n = 11$). A total of 14 interviews were conducted, but data from three participants (participants 02, 06, and 14) were later removed due to technical difficulties with the recording which rendered a fair transcription impossible. No information about gender was collected for the quantitative survey data, since this was not deemed relevant to the hypotheses. For the qualitative interview, the 11 included participants were six male and 5 female students enrolled in a variety of programmes (Psychology, Finance and accounting, Business, Actuarial science, and Engineering).

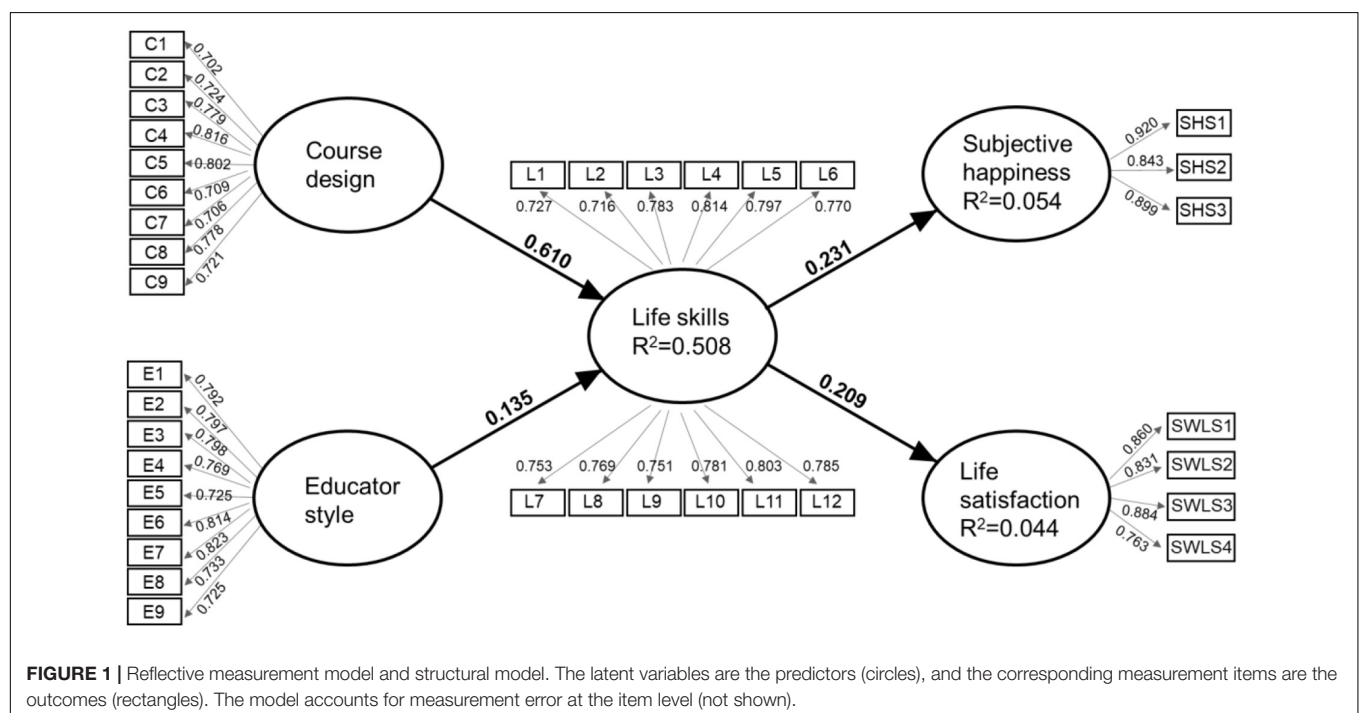
Instruments

Prior to data analysis, a structural model was constructed from a review of the published literature, indicating the latent and measured variables, and their causal relationships (Figure 1). The central part of the structural model is founded on the premise that the effectiveness by which a student achieves *life skills* as measured by the positive education learning outcomes, are determined by an interplay between factors, which can be classified primarily as the *course design* and *educator style* (McKone, 1999; Eom and Ashill, 2016). In the context of

the higher education setting, we predict that the degree to which a student perceives that (s)he has achieved these life skills may in turn influence their emotional state and sense of optimal wellbeing (Seligman et al., 2009) which we operationalise within the structural model as *happiness* and *life satisfaction*, respectively. The full measurement model is shown in Figure 1. It is a reflective measurement model since the indicators of each construct are considered to be caused by the construct itself and the items are interchangeable.

The five latent variables are shown in Figure 1 (measurement scales are reported in Supplementary Table 1). Course design was measured using a set of nine tailor-made statements which asked about the content (C1-3), learning objectives (C4), structure (C5), level of enjoyment (C6), and assessment (C7-9). Educator style was measured using another set of nine tailor-made statements which asked about educator's personal characteristics and attitudes (E1, E4-5), feedback (E2-3), knowledge (E6-7), and inter-personal skills (E8-9). Life skills were measured using a set of bespoke statements that were directly informed by the course learning objectives. Statements covered each of the three major taxonomic categories of student learning; knowledge (L1), skills (L2-3, L6, L9-10, and L12), and attitudes (L8, L11), plus behaviour (L4-5, L7). All three sets of questions on course design, educator style and life skills had a forced four-point Likert scale (*Strongly disagree* to *Strongly agree*). There was no safe 'neutral' option which is sometimes seen as an easy option to take when a respondent is unsure, and so may not reflect a truly neutral point of view (Allen and Seaman, 2007).

Global subjective happiness was measured using the unidimensional "Subjective Happiness Scale" (SHS, Lyubomirsky and Lepper, 1999). Developed and validated in the United States, the SHS considers happiness from the respondent's own



perspective and therefore accounts for differences in cultural expectations between Western and Asian societies. The original scale has four statements, but we removed “Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. To what extent does this characterisation describe you?” because the counterpart statement about happiness (item 3) was deemed sufficient. The response format was a seven-point Likert scale, with higher scores reflecting greater happiness. We defined life satisfaction as the cognitive assessment of one’s life as a whole. Individuals tend to use their own benchmarks and criteria in making this assessment and so it is more meaningful to assess global judgements of life satisfaction, rather than satisfaction with specific life domains (Diener and Diener, 1995; Proctor et al., 2009). Students’ life satisfaction was measured using the unidimensional “Satisfaction With Life Scale” (SWLS; Diener et al., 1985). Developed and validated in the United States, it has been validated in Asia (Bai et al., 2011). The original scale has five statements, three referring to satisfaction with the present, and two to satisfaction with the past. We removed “If I could live my life over, I would change almost nothing” since this was deemed to have inadequate construct validity for our target youth sample. Responses were indicated using a seven-point Likert scale (*Strongly disagree* to *Strongly agree*), and the global score has previously been shown to be sensitive in discriminating the full range of life satisfaction from extremely dissatisfied to extremely satisfied (Pavot and Diener, 1993). We are justified in removing individual items from the SHS and SWLS since in the context of the reflective measurement model, all items are expected to be correlated (i.e., possess internal consistency reliability), and so dropping an item from the measurement model should not alter the meaning of the items (Jarvis et al., 2003).

Procedure

The positive education lectures and workshops were conceived by the senior management team, compulsory for students and formed part of a credit-bearing “Self-Empowerment and Social Responsibility” course that constituted level 1 of the EmPOWER programme. This 24-week course was designed with expert input from educators and a practising consultant in positive psychology. The curriculum was informed by Seligman’s Wellbeing Theory (Seligman, 2011), incorporating the principles of positive psychology by blending development in skills that promote thriving (e.g., self-awareness, intrinsic motivation, and leadership; Benson and Scales, 2009), with an emphasis on developing by mapping personal strengths and opportunities (Sprangel et al., 2011), character strengths and virtues (Peterson and Seligman, 2004). The curriculum also included discovering purpose and creating a plan to mobilise that purpose into positive impact (Benson and Scales, 2009). Each week had 2 h of contact time, comprising conventional lecture-style presentations plus workshops to develop a personal impact statement comprising three personalised statements “I am a . . . My purpose is . . . I will . . .” Additional personal study time was required to complete an impactful community project through team-working and homework exercises that required self-reflective practise.

The study was approved by the Social Sciences Ethics Committee Heriot-Watt University on September 2019 (ref: 2019-120), and all participants gave online (survey) and verbal and written (interview) informed consent in April 2020. Quantitative data were gathered through an online survey using Qualtrics software (Seattle, United States). Course design, educator style, and life skills were measured at T2 and T3. Two of the life skills (items L11: positive thoughts, and L12: critical thinking) were assessed only at the end of the course (T3) because they had not yet been adequately covered at the mid-way point. SHS and SWLS were assessed at three time points: T1 - prior to the start of the course, T2 - at the mid-way point (12 weeks), and T3 - at the end of the course (24 weeks).

Qualitative data were gathered through a semi-structured interview, conducted by ZB, one of the co-authors and a health psychologist, who was not involved in the delivery of the course. Questions probed how the life skills were achieved. The interview followed the approach of Interpretative Phenomenological Analysis which enables multiple participants who experience similar events to tell their stories without any distortions and then seeks to make sense of the ‘lived experiences’ of the research participants (Alase, 2017).

Design and Analysis

Change over time in SHS and SWLS were examined using a one-way ANOVA, and change over time in life skills were examined using a paired *t* test.

For the quantitative survey data, the hypothesised structural model in **Figure 1** was analysed using Partial Least Squares Structural Equation Modelling (PLS-SEM) implemented with Smart PLS version 3 (SmartPLS GmbH, Germany). PLS contains a two-step procedure as recommended by Henseler et al. (2009), which involves the evaluation of the (outer) measurement model followed by evaluation of the (inner) structural model.

The assessment of the measurement model (**Figure 1**) used the T3 data ($n = 350$) and involved examining the reliabilities of the individual items (indicator reliability), internal consistency [Cronbach’s alpha and composite reliability], convergent validity (Average Variance Extracted (AVE)) and discriminant validity (Fornell-Larcker criterion, heterotrait-monotrait (HTMT) ratio of correlations criterion, cross-loading; Fornell and Larcker, 1981; Hair et al., 2017).

As a measure of indicator reliability, the factor loadings of the measurement items should be higher than 0.70 for inclusion and should be deleted if less than 0.40 and 0.70 (Hulland, 1999; Hair et al., 2017). Regarding internal consistency in the scale items, Cronbach’s alpha (α) is a measure of internal consistency (i.e., how closely related the observed items were in representing the underlying construct), while Composite Reliability (CR) is an indicator of the shared variance among the observed items used as an indicator of a latent construct. For good reliability, α should be greater than 0.70 and CR should be 0.70–0.95 (Hair et al., 2017).

Convergent validity assesses whether the actual results confirm expected relationships between items and underlying constructs. The Average Variance Extracted (AVE) measures the level of variance captured by a construct vs. the level due to

measurement error, and a value of 0.50 is acceptable while 0.7 is very good (Fornell and Larcker, 1981). Discriminant validity assesses the extent to which the underlying constructs in the model are distinct from one another (Hulland, 1999). To satisfy this requirement, items should load more onto their intended construct than on another. This requirement was tested by three indices (Hair et al., 2017). Under the Fornell-Larcker criterion, the square root of each construct's AVE should have a greater value than the variance of the other constructs. The HTMT value should be lower than 0.85 (Kline, 2011), although some authors suggest 0.90 (Teo et al., 2008). And each measurement item should correlate weakly with all other constructs, except for the one to which it is theoretically associated, which should be higher than 0.70 (Hair et al., 2017).

In step two of the PLS approach, we evaluated the structural model's predictive ability and the relationships between the constructs. The coefficient of determination (R^2), path coefficients (β), and t -statistic values, effect size (f^2) and the predictive relevance (Q^2) are the key standards for such evaluation (Hair et al., 2017). R^2 measures the amount of variance explained in the endogenous constructs (i.e., life skills, happiness and life satisfaction). Following Hair et al. (2017), $R^2 = 0.75$ is considered substantial, $R^2 = 0.50$ is regarded as moderate, and $R^2 = 0.26$ is considered weak.

The path coefficients (β) represent the strength of the hypothesised relationships between the constructs. A bootstrapping technique with 5,000 resamples was conducted to estimate the beta (β) and corresponding t values as recommended by Chin et al. (2003). The greater the beta coefficient (β), the stronger the effect of an exogenous construct on the endogenous construct. Path coefficients with a value close to 1 represents a strong positive relationship, and conversely a value closer to -1 represents a strong negative relationship. The overall effect size (f^2) is a measure of the degree of impact of the path relationship. Following Hair et al. (2017), $f^2 = 0.75$ is considered large, $f^2 = 0.15$ is regarded as medium, and $f^2 = 0.02$ is considered small. The predictive relevance for the structural model was evaluated using Q^2 (Tenenhaus et al., 2005) and so it can be considered an indicator of the quality of the structural model. Interpretation of Q^2 followed that of Hair et al. (2017), with a value > 0 indicating adequate predictive relevance, and a value < 0 indicating poor predictive relevance.

For the qualitative interview data, we employed a Framework Method to analyse the anonymised interview transcripts. The Framework Method is most commonly used for the thematic analysis of semi-structured interview and it is appropriate for application to reasonably homogenous data covering similar topics (Gale et al., 2013). Our approach was partly deductive because coding proceeded through the lens of predefined categories based on the structural model: (i) how the course content influenced perceived life skills, (ii) how the instructor style of delivery influenced students' learning, (iii) happiness, and (iv) life satisfaction. Coders followed the Framework Method steps (i.e., Transcription; Familiarisation with the interview; Open coding; Developing an analytical framework by structuring the codes under each category; Applying the analytical framework to the remaining transcripts; Charting the

data into the framework matrix; Interpreting the data), and used Quirkos version 2.4.2 (Quirkos, Scotland). Two coders (ZB and DH) who were not part of the EmPOWER teaching team independently completed the open coding. Open coding was first conducted on three transcripts to confirm that important aspects of the data were not missed by restricting to these four categories. Data interpretation interrogated the theoretical concepts (either prior concepts or ones emerging from the data) and mapped connections between categories to explore relationships or causality.

RESULTS

Data from the 246 included respondents did not support hypothesis (i) that the course would improve students' wellbeing in accordance with Wellbeing Theory (Seligman, 2011). There was no detectable change in wellbeing as measured by the SHS [$F(2,490) = 1.317$, $p = 0.269$] and SWLS [$F(2,490) = 0.363$, $p = 0.696$] (Table 1 and Supplementary Table 2).

As the course progressed from T2 to T3, there was a small, but statistically significant increase in life skills scores [$t(245) = -3.741$, $p < 0.001$] (Table 1 and Supplementary Table 2). Thus, hypothesis (ii) was supported.

Directional influences between variables (i.e., hypotheses iii and iv) were evaluated using Structural Equation Modelling. The (outer) measurement model defined in Figure 1 met all *a priori* evaluation criteria regarding the reliability and validity of the items in the constructs (Table 2). The factor loadings of the measurement items ranged between 0.702 and 0.920 which meets the threshold for indicator reliability and confirms the relative importance of each observed item to the underlying construct. The α values for all construct factors were 0.871–0.938 and CR values were 0.902–0.946, indicating the scales had acceptable reliability. AVE values were acceptable for course design, educator style and life skills (i.e., 0.562–0.602) and very good for happiness and life satisfaction (i.e., 0.699–0.788), and so convergent validity was confirmed for this hypothesised model.

The model also met all *a priori* evaluation criteria regarding the discriminant validity of the latent constructs. From Table 3, it can be seen that the square root of each construct's AVE for each construct was greater than the correlation involving the constructs thus meeting the criterion of Fornell and Larcker (1981), and from Table 4 the results also passed the HTMT criterion test with values not exceeding 0.85. Moreover, from the cross-loading evaluation all the loading indicators on the assigned construct were > 0.70 and higher than the loadings on the other constructs (Supplementary Table 3).

Based on these findings, the proposed structural model was accepted, with confirmation of adequate reliability, convergent validity, and discriminant validity. The next step was to evaluate the (inner) structural model outcomes comprising the latent constructs and path coefficients (Figure 1).

Hypothesis (iii) predicted that the course design and educator style would each positively influence students' learning of life skills, and this was supported by our data. Course design contributed in a meaningful way to learning life skills, while the

TABLE 1 | Scores on survey data assessed over the duration of the course.

Measure	T1 M (SD)	T2 M (SD)	T3 M (SD)	Statistic	p value
Life skills	–	28.19 (4.40)	29.17 (4.28)	–3.741	< 0.001
L11 (Positive thoughts)	–	–	2.92 (0.50)		
L12 (Critical thinking)	–	–	2.96 (0.54)		
Subjective happiness	14.31 (3.40)	14.18 (3.21)	14.47 (3.29)	1.317	0.269
Satisfaction with life	18.73 (4.73)	18.59 (4.49)	18.52 (4.53)	0.363	0.696

T1, pre-intervention; T2, mid-way point; T3, post-intervention. Subjective happiness scores range from 3 to 21. Satisfaction with life scores from 4 to 28. Life skills scores range from 10 to 40.

educator style played only a small positive role. The R^2 value for life skills is 0.508 meaning that the two exogenous constructs (course design and educator style) explain 50.8% of the variance in this endogenous construct (**Figure 1**). This is a moderate effect

(Hair et al., 2017). The Q^2 value for the life skills construct was equal to 0.290, which was higher than the threshold limit, indicating that the path model had sufficient predictive relevance for this construct (Hair et al., 2017). From the path coefficients, both course design and educator style significantly and positively influenced life skills, $\beta = 0.610$ and 0.135 , respectively, $p < 0.05$ (**Table 5**). Considering the effect size for the influence of course design on life skills ($f^2 = 0.372$, **Table 5**), the result exceeded a ‘medium’ effect (Hair et al., 2017). In contrast, for educator style, f^2 was 0.018 which is considered small.

Hypothesis (iv) predicted that achievement in life skills would in turn drive greater wellbeing as measured by subjective happiness and life satisfaction. Statistically speaking, the hypothesis was supported, but the size of the influence was very small. The smaller R^2 values for happiness (0.054) and life satisfaction (0.044) suggest that life skills explain very little of the variance in these endogenous constructs (i.e., 5.4 and 4.4%, respectively) (**Figure 1**). While the Q^2 values for the happiness and life satisfaction constructs also reached sufficient predictive relevance (0.036 and 0.027, respectively), they were only just greater than 0, indicating barely adequate predictive relevance (Hair et al., 2017). From the path coefficients, life skills significantly and positively influenced students’ self-reported happiness and life satisfaction, $\beta = 0.230$ and 0.209 respectively, $p < 0.001$ (**Table 5**). However, the corresponding effect sizes were small (Hair et al., 2017).

Understanding the Lived Experience of the Learning Process

The interview findings shed light on the paths influencing perceived life skills acquisition. Four themes emerged about the course design. It was students’ first opportunity to engage in life skills training, it gave them a chance to demonstrate their ability to articulate what they had learned, the course instilled a genuine sense of excitement to learn, and it promoted a greater self-awareness and sense of purpose. Three themes emerged about educator style. Learning activities were engaging and informative and students were encouraged to apply their knowledge, but there were some criticisms about quality of the teaching staff. No themes on happiness were observed, but there was some evidence that the course contributed to life satisfaction as it encouraged students to engage in cognitions and behaviours that would promote success.

TABLE 2 | Indicator reliability and internal consistency.

Main constructs	Items	Loadings	α	CR	AVE
Course design	C1	0.702	0.903	0.920	0.562
	C2	0.724			
	C3	0.779			
	C4	0.816			
	C5	0.802			
	C6	0.709			
	C7	0.706			
	C8	0.778			
	C9	0.721			
Educator style	E1	0.792	0.917	0.931	0.602
	E2	0.797			
	E3	0.798			
	E4	0.769			
	E5	0.725			
	E6	0.814			
	E7	0.823			
	E8	0.733			
	E9	0.725			
Life skills	L1	0.727	0.938	0.946	0.595
	L2	0.716			
	L3	0.783			
	L4	0.814			
	L5	0.797			
	L6	0.770			
	L7	0.753			
	L8	0.769			
	L9	0.751			
Happiness	L10	0.781	0.871	0.918	0.788
	L11	0.803			
	L12	0.785			
	SHS1	0.920			
Life satisfaction	SHS2	0.843	0.856	0.902	0.699
	SHS3	0.899			
	SWLS1	0.860			
	SWLS2	0.831			
	SWLS3	0.884			
	SWLS4	0.763			

For details of the items see **Supplementary Table 1. A**, Cronbach’s α ; CR, composite reliability; AVE, average variance extracted.

TABLE 3 | The Fornell-Larcker Criterion Test for discriminant validity which requires values in bold to be greater than the remaining values in each column.

	Course design	Educator style	Life skills	Happiness	Life satisfaction
Course design	0.750				
Educator style	0.713	0.776			
Life skills	0.707	0.570	0.771		
Happiness	0.210	0.210	0.231	0.888	
Life satisfaction	0.202	0.240	0.209	0.564	0.836

TABLE 4 | The Heterotrait-Monotrait (HTMT) Criterion Test for discriminant validity which requires values not to exceed 0.85.

	Course content	Educator style	Learning outcomes	Happiness	Life satisfaction
Course design					
Educator style	0.782				
Life skills	0.749	0.606			
Happiness	0.222	0.220	0.237		
Life satisfaction	0.232	0.269	0.230	0.645	–

TABLE 5 | Path coefficients, *t* statistics, and effect size (*f*²).

Hypothesised path	Std β	Std error	<i>t</i> statistic	<i>p</i> value	<i>f</i> ²
Course design → Life skills	0.610	0.065	9.401	0.000	0.372
Educator style → Life skills	0.135	0.068	1.983	0.024	0.018
Life skills → Happiness	0.231	0.056	4.161	0.000	0.057
Life skills → Life satisfaction	0.209	0.052	4.038	0.000	0.046

Course Design

For many interviewees, this was the first time they had engaged in self-reflection to know themselves better and in life skills training to develop personal effectiveness. Moreover, interviewees recognised and appreciated unique opportunities open to them that had not been offered at school nor were available to their friends at other universities.

before. So I think that really helps me reflect how a person I am, in what way."

Students were also able to articulate how they had developed certain personal effectiveness skills, particularly leadership (Learning outcome L2), team-working and communication (L6) and time management (L7).

- Student 1 said: "I kind of enjoy this because. In our public school system in our country we don't really have something like this. So to come to university and to see something like this, it's something that I appreciate, I can appreciate that personally. ... When I first saw that there was a course like this, I was personally quite delighted because this is based off my previous experiences in other universities, they don't really put that much emphasis on programmes like this."
- Student 3 said: "So I told them [my friends]" and their response was "Wow, it's a very good initiative." My friends are from other university, and they mentioned that there is a lack of talks and all those stuff. And they are surprised to see, "Wow, you really have a programme that motivates you to attend those talks."
- Student 5 said: "All these project managing, working as a team, all our leadership skills, other universities do not have that."
- Student 12 said: "Cause before I joined the university [...] I was probably a bit lost in several ways. But during the EmPOWER program, they introduced us to some concepts and some personality traits that I've never really think about before. So I think that really helps me reflect how a person I am, in what way."
- Student 3 said: "I have a good leadership skill. ... I try to improve myself, like joining those clubs and societies, so that when I put it in the LinkedIn post. ... so people will get to notice you, your leadership skills."
- Student 4 said: "I think it's a really good platform for me to start leading people such as. I pick up the responsibility of saying, "Okay, I will do this. I will lead the group. I will delegate a task and then communicate with my members so that my members would say without a leader in the group, I think everything would be chaos. I took the initiative to become a leader. And because of that, I think it helps me to improve my leadership skills. I have definitely improved on my time management skills, because we managed to finish everything on time. We were actually ahead of schedule. And despite certain setbacks from the conflict, we actually still managed to get back ahead of schedule also."
- Student 13 said: "Toward the end, before we conduct our social project, we got into a little bit of a disagreement and then to solve that, we did it as team, so I think that was a good experience 'cause everyone set aside our differences and we got together to solve something."

Students were also able to articulate how they had developed greater emotional awareness (Learning outcome L3), self-regulation (L4), ability to express gratitude (L8) and ability to reframe thoughts more positively (L11).

- Student 4 said: *"I think one of the weakness. My last one was actually something about love. And I do not know how to harness that in my networking among my university friends and lecturers. But maybe I would just say that, through these few months I was in uni, maybe I would say I will just increase. Further strengthen my love toward myself and then my love toward friendships, friendships and professional development."*
- Student 5 said: *"A positive impact it did for me is that, after writing the gratitude letter, it kind a opened my eyes to everything else that I should be grateful for."*
- Student 7 said: *"I'm not usually the type to show love, but because of the gratitude letter, it helps me to show more love to people, so it did help me also."*
- Student 8 said: *"It was really good. I actually did it for my grandfather, who has recently passed away, so. He was happy. He was very happy. I'm really glad I did it."*

The course instilled a genuine sense of excitement to learn.

- Student 5 said: *"I guess, my highlight [...] was the Community Project itself, because it kinda opened my eyes to [...] how work will be, it kinda gives this. forces us to work as a team or work on our own, so we kinda have an idea on how we can apply certain things later on."*
- Student 9 said: *"I remember the attendance was also many because all of us was very excited and interested to hear this workshop."*

Interviewees mentioned on a number of recurring aspects of the course design about how the content had enabled them *'to develop our personal skill which we cannot learn during lectures'* (Student 10). The purpose work (Learning outcome L1) left a lasting impression. Overall, the purpose activities instilled a greater sense of who students are and where they are heading, particularly at a formative time in year 1 when they had further years of study ahead.

- Student 3 said: *"Impact statement, what I find it about is, it's very interesting because even. It knows what you wanna achieve in the future. . . I think that it helps to know what you're capable of in the future, it helps to realise what are your goals, what are your dreams, what you wanna achieve."*
- Student 7 said: *"It helped me a bit because before the impact statement, I don't know what actually am I doing. I'm just doing accounting because I thought this is the easiest for me. But with the impact statement, at least I have a purpose. . ."*
- Student 12 said: *"EmPOWER course is very helpful for especially year one students, who, some of them just entered the proper university environment. And it really helped students to find their purpose and set their goals during the 3–4 years of study in university. And I think it also helps to encourage students to think more about themselves and to really find what kind of person they are."*

Educator Style

The active and positive style of delivery played a contributory role in making the learning activities engaging and informative and encouraging the students to learn new things (Learning outcome L10).

- Student 9 said: *"They reach out to us very often, and remind us on the task, and then it's like more, how to say, closer to us. We can feel that they are very genuinely want to help us, to improve."*
- Student 12 said: *"Let's say the workshop facilitators were very empowering. They were very active. And also, they really help us to follow the instructions and progress through the workshops smoothly."*
- Student 13 said: *"I would say that the instructor was a really helpful person firstly, he gave us certain points, like what to do in this section? What to put in this section? Why it's working and why is not? And all that. I think that was always useful. And he also told us dos and don'ts."*

The hands-on activities encouraged students to apply what they had learned, and this reinforced all life skills that were taught.

- Student 7 said: *"And then we need to do a project like how to clean the river. So actually we did. Like a hands-on activity there because we discuss in the group how to do. How to build something to clean the river. She did the lecture for 1 h and she gave us like 30 min to do the project and another 30 min for presentation. So for me, that is quite interesting because we don't actually sit there for 2 h straight. We're actually like doing something and we're actually using the techniques that she's giving."*

Not all aspects of teaching were well received, and some students commented on the variability in quality of the teaching staff. This may have affected the model findings, particularly the relationship between educator style and life skills.

- For example, student 7 recognised that that not all workshops embedded interactive activities: *"I would prefer it will be like 1 h lecture and another 1 h for us to use that skill, like we try it ourselves."*
- Student 4 said: *"... in all honestly, because they are just reading from the slides, to be honest [...] instead of maybe giving lectures and stuff, like giving talks for 2 h, you should give something more hands-on, more activities, something that requires thinking."*
- Student 1 said: *"Diversifying the power speakers is a good idea [...] But the problem with that is consistency [...] They don't deliver it to that similar standard that we come to have expected from the previous week."*

Life Satisfaction

No themes on happiness were observed, but there was some evidence that the course contributed to life satisfaction as it encouraged students to engage in cognitions and behaviours that promote success. For example, student 13 mentioned: *"I would say that it's worth the shot. I would recommend it. . . it's interesting"*

to know that there are certain things that you think you know yourself, but you don't really know. You don't know the kind of potential that you could unleash. So this programme kinda helped me to find my potential and things that I could do better or improve on." Student 7 also reflected on how the life skills training made them reflect on what was important in their life to achieve personal satisfaction. *"I think more programmes about why. I mean why we need to help more people, something like that, instead of focussing on your work, instead of focussing like to achieve. Instead of focussing to achieve money and stuff, maybe we can say that there's more to life like we need to be like a good people, so everyone can enjoy life together, I think. Yeah, because for me money is everything, but at the same time, life. Like to help others is more important for me."*

DISCUSSION

To our knowledge, this is the first report of a holistic development programme conceived by the senior managerial team, fully integrated into the curriculum and completed by all university students enrolled across an academic year group. Moreover, the study was conducted in an Asian setting. The principal finding from the quantitative data demonstrated improved life skills, mostly achieved through the course design. The qualitative data added value to this finding by illustrating how learning was not limited to knowledge acquisition, but that the students recognised what psychosocial skills, attitudes and values are needed for positive behaviour change to occur (c.f. UNICEF, 2012).

Although there was no detectable change in subjective happiness and life satisfaction scores over time, perceived life skills did exert a small influence on wellbeing at the end of the course. The qualitative data indicated that important seeds had been planted which could promote a mindset for flourishing in adulthood. Longitudinal studies have shown that positive education interventions are at their best when they exert measurable benefits that evolve over time, such as by enhancing academic application through greater engagement in the learning process (e.g., Denovan et al., 2020). In their systematic narrative review, Nasheeda et al. (2019) concluded that many life skills programmes in developing countries (including those across Southeast Asia) fail to achieve sustained benefits because they focus on one-shot or short-term interventions rather than on ongoing activities. The same criticism can also be levied at positive education programmes across Asia (Hall et al., 2021a). The programme under study here is intended to foster lifelong impact as students progress through the levels of the EmPOWER programme throughout their student journey. Although our study design did not include a long-term follow-up assessment, we would recommend testing intervals over the course of the student journey to examine the learning trajectory. This would require institutional resource for systematic planning, implementation, monitoring and evaluation (UNICEF, 2012).

From our cross-sectional modelling data that assessed short-term effects, we conclude that effective course content is more critical than educator style for good life skills acquisition. This

finding is consistent with the published educational literature. For example, course design is known to positively influence students' perceived learning (Eom and Ashill, 2016), with high levels of interaction connecting students with their instructors, the course content, and their fellow classmates (McKone, 1999; Boling et al., 2012). In many ways, course design and educator style should be mutually enhancing, yet in the present study, educator style had a relatively small effect on life skills. We suggest that this may be explained by the somewhat variable student experience that was evident from the interviews, for example in the quotes cited from students 1, 4 and 7. In their exploration of thriving in adolescence, Benson and Scales (2009) reflected on the importance of the interaction between the student and the educational environment. They conceptualised that the process of adolescent thriving is animated by a passion for, and the exercise of action to nurture, a self-identified interest, skill, or capacity. Here, they proposed the role of the educator is to offer, find and create opportunities for the young person to express their spark, as well as to know, affirm, celebrate and guide that spark. Hagenauer and Volet (2014) also argued for a wider research agenda on teacher-student relationships as a relevant construct in higher education. In the context of the present study, with many hundreds of students in the class and only a small number of educators, we suggest that may have hindered the ability to create an optimal positive developmental context, even with the most passionate educators. To foster such development, several students talked about the importance of being taught by academic staff whom have the skills and competencies to facilitate an interactive learning environment that enables them to put into practise those life skills. In this regard, our results address the research gap identified by Nasheeda et al. (2019) by defining what is needed to foster optimal positive change. However, further research is needed to evaluate the impact of this recommended change in practise.

Limitations

A large number of 570 students who were enrolled on this course did not contribute data to the study. Overall, 220 (38.6%) had an overall attendance record below the prerequisite 80%, and so they did not meet the study inclusion criteria. Hypotheses (i) and (ii) were tested using data from only 246 students (43.2% of the total sample) since 104 students attended 80% of the course, but did not complete the life skills survey, SHS and SWLS at T2. Attrition bias can weaken external validity, meaning that the study findings cannot be generalised to other populations, especially if those missing data came from students with a systematic pattern of characteristics (such as being less motivated or engaged in the course). Potentially attrition bias can also weaken internal validity, meaning that the relationships between the latent variables may change. We would argue that this is unlikely because hypotheses (iii) and (iv) were tested using data from 61.4% of the total sample (350 students) collected at T3.

Future Research Directions

From our work, we make two recommendations to advance the field; first to conduct studies that assess long-term impact as well as short-term effects and second to understand how to maximise

benefit for students, including those in Asia. In his recent editorial on the state of Positive Education in Asia, King et al. (2016) identified a lack of attention to the socio-cultural context as another important gap in the research literature. Two-thirds of the world's population lives in Asia, yet most of the academic work is conducted in Western countries. This geographical bias is especially awkward given that culture is well-known to play a key role in wellbeing and learning (e.g., Stankov, 2013; Caleon et al., 2015). Many of the developing countries in Asia have traditionally put primacy on learning as a means for social mobility and economic survival which may explain why an Asian higher education tends to emphasise academic achievement and professional qualifications (Lin et al., 2018). Our qualitative data indicated students' appreciation of the opportunity to address wellbeing and life skills in the context of their own learning. They also appeared to be aware that the course they were studying offered something unique compared to other education providers in the country. Other researchers have shown that youths from Asian societies have a stronger sense of family obligation compared to their Western counterparts (King and Ganotice, 2015). In contrast to the work by Caleon et al. (2015), we found no evidence for a role of external motivators such as needing to show parents that they are achieving. Instead, we found many more examples of an intrinsic motivation to learn. These findings highlight the need for more contemporary research on the Gen Z student population so that researchers and educators can better understand changing perceptions and expectations. For those educators and researchers in Asia, our work also poses the cross-cultural issue about whether families of Gen Z students are becoming more like their Western counterparts in terms of their educational expectations and career aspirations. Based on our experience with the EmPOWER programme, Asian students are calling for a shift toward a more collaborative and empowering approach to teaching (see also Hall et al., 2021b). In this regard, we may move even closer to achieving optimal flourishing by combining positive education with coaching psychology practises (Palmer and Whybrow, 2008; Seligman and Adler, 2018).

Conclusion

The teaching of wellbeing and life skills in higher education provides students with valuable personal resources for growth. Our quantitative and qualitative research findings indicate that such a novel educational approach is well-received by Asian students and may sow the seeds for future benefit by positively impacting on their skills, behaviours, attitudes, and values.

REFERENCES

- Alase, A. (2017). The Interpretative Phenomenological Analysis (IPA): a Guide to a Good Qualitative Research Approach. *Internat. J. Educ. Lit. Stud.* 5:2. doi: 10.1111/sms.13953
- Allen, I. E., and Seaman, C. (2007). Statistics roundtable: likert scales and data analyses. *Q. Prog.* 40, 64–65.
- Bai, X., Wu, C., Zheng, R., and Ren, X. (2011). The psychometric evaluation of the satisfaction with life scale using a nationally representative sample of China. *J. Happ. Stud.* 12, 183–197.

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 Social Sciences Ethics Committee Heriot-Watt University on September 2019 (ref: 2019-120). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TG and ZB designed the study. TG organised the database. TG and DH performed the quantitative analysis. ZB and DH performed the qualitative analysis. TG, ZB, and DH wrote the first draft of the manuscript. JL, PL, and DH revised it critically for important intellectual content. All authors contributed to conception of the study, and read and approved the submitted version.

FUNDING

This work was supported by the EmPOWER research fund, Heriot-Watt University Malaysia.

ACKNOWLEDGMENTS

We thank Albert Lee for assisting with the literature review and creating the perceived learning outcome questionnaire, and Renuka Tharuma Rethinam (course leader) for helping with the survey data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.827517/full#supplementary-material>

- Benson, P. L., and Scales, P. C. (2009). The definition and preliminary measurement of thriving in adolescence. *J. Posit. Psychol.* 4, 85–104. doi: 10.1016/s0140-6736(96)01358-x
- Boling, E. C., Hough, M., Krinsky, H., Saleem, H., and Stevens, M. (2012). Cutting the distance in distance education: Perspectives on what promotes positive, online learning experiences. *Inter. High. Educ.* 15, 118–126. doi: 10.1016/j.iheduc.2011.11.006
- Caleon, I. S., Wui, G., Tan, J. P.-L., Tan, C. S., and King, R. B. (2015). Cross-cultural validation of the Academic Motivation Scale: a Singapore investigation. *Child Indicat. Res.* 8, 925–942. doi: 10.1007/s12187-014-9298-7

- Chin, W. W., Marcolin, B., and Newsted, P. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Informat. Syst. Res.* 14, 189–217.
- Denovan, A., Dagnall, N., Macaskill, A., and Papageorgiou, K. (2020). Future time perspective, positive emotions and student engagement: a longitudinal study. *Stud. High. Educ.* 45, 1533–1546. doi: 10.1016/j.concog.2020.102936
- Diener, E., and Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *J. Personal. Soc. Psychol.* 68, 653–663. doi: 10.1037//0022-3514.68.4.653
- Diener, E., Emmons, R. A., Larsen, R. J., and Griffin, S. (1985). The Satisfaction with Life Scale. *J. Personal. Assess.* 49, 71–75.
- Dondi, M., Klier, J., Panier, F., and Schubert, J. (2021). *Defining the skills citizens will need in the future world of work*. Atlanta, GA: McKinsey and Company.
- Eom, S. B., and Ashill, N. (2016). The determinants of students' perceived learning outcomes and satisfaction in university online education: an update. *Dec. Sci. J. Innov. Educ.* 14, 185–215.
- Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Market. Res.* 18, 39–50. doi: 10.2307/3151312
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., and Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med. Res. Methodol.* 13:117. doi: 10.1186/1471-2288-13-117
- Hagenauer, G., and Volet, S. E. (2014). Teacher–student relationship at university: an important yet under-researched field. *Oxf. Rev. Educ.* 40, 370–388. doi: 10.1080/03054985.2014.921613
- Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*, 2nd Edn. Thousand Oaks, CA: SAGE.
- Hall, D., Juliana, J., Manickam, M., Singh, A. S. T. S., Wei, S. T. S., Vuong, P. A., et al. (2021a). Examining the state of positive education in tertiary education across East, South and Southeast Asia: a scoping review protocol. *SocArXiv* doi: 10.31235/osf.io/wa8y5
- Hall, D. A., Ragnathan, T., Wong, L. W.-L., Low, J., and Choong, S. (2021b). COVID-19 as a catalyst to inspire learning and create fresh, meaningful interpersonal interactions: a case study of positive psychology in higher education. *J. Psikol. Malaysia* 35, 234–246.
- Henseler, J., Ringle, C. M., and Sinkovics, R. (2009). The use of partial least squares path modeling in international marketing. *Adv. Internat. Mark.* 20, 277–319. doi: 10.2196/jmir.3122
- Hood, B., Jelbert, S., and Santos, L. R. (2021). Benefits of a psychoeducational happiness course on university student mental well-being both before and during a COVID-19 lockdown. *Health Psychol. Open* 8, 1–12. doi: 10.1177/2055102921999291
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strat. Manag. J.* 20, 195–204. doi: 10.1002/(sici)1097-0266(199902)20:2<195::aid-smj13>3.0.co;2-7
- Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *J. Cons. Res.* 30, 199–218. doi: 10.1086/376806
- King, R. B., and Ganotice, F. A. (2015). Does family obligation matter for your motivation, engagement, and well-being? It depends on your self-construal. *Personal. Individ. Diff.* 86, 243–248.
- King, R. B., Caleon, I. S., Tan, J. P.-L., and Ye, S. (2016). Positive Education in Asia (editorial). *Asia-Pacific Edu. Res.* 25, 361–365.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- Lin, L. L., Junaidi, M., Amirul, M. R., and Thuraya, N. (2018). *The school-to-work transition of young Malaysians*. Kuala Lumpur, Malaysia: Khazanah Research Institute. Available online at: http://www.krinsitute.org/assets/contentMS/img/template/editor/20181212_SWTS%20Presentation%20DR%20LIM%20.pdf (accessed 13 June 2021).
- Lomas, T., Waters, L., Williams, P., Oades, L. G., and Kern, M. L. (2020). Third wave positive psychology: broadening towards complexity. *J. Posit. Psychol.* 16, 660–674.
- Lyubomirsky, S., and Lepper, H. S. (1999). A measure of subjective happiness: preliminary reliability and construct validation. *Soc. Indic. Res.* 46, 137–155. doi: 10.1017/sjp.2017.8
- McKone, K. E. (1999). Analysis of student feedback improves instructor effectiveness. *J. Manag. Educ.* 23, 396–415. doi: 10.1177/105256299902300406
- Nasheeda, A., Binti Abdullah, H., Krauss, S. E., and Binti Ahmed, N. (2019). A narrative systematic review of life skills education: effectiveness, research gaps and priorities. *Internat. J. Adol. Youth* 24, 362–379.
- Norrih, J. M., Williams, P., O'Connor, M., and Robinson, J. (2013). An applied framework for Positive Education. *Internat. J. Wellb.* 3, 147–161.
- Palmer, S., and Whybrow, A. (eds) (2008). *Handbook of coaching psychology: A guide for practitioners*. Milton Park: Routledge/Taylor & Francis Group.
- Pavot, W., and Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychol. Assess.* 5, 164–172.
- Peterson, C., and Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Washington, DC: American Psychological Association.
- Proctor, C. P., Linley, A., and Maltby, J. (2009). Youth life satisfaction measures: a review. *J. Posit. Psychol.* 4, 128–144.
- Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. North Sydney, NSW: William Heinemann Australia.
- Seligman, M. E. P., and Adler, A. (2018). “Positive education,” in *Global Happiness Policy Report: 2018*, eds J. F. Helliwell, R. Layard, and J. Sachs (Global Happiness Council), 52–73.
- Seligman, M. E. P., Ernst, R. M., Gillham, J., Reivich, K., and Linkins, M. (2009). Positive education: positive psychology and classroom interventions. *Oxford Rev. Educ.* 35, 293–311.
- Simel, S., and Bogner, B. (2017). “Positive orientation to education and critical friendship,” in *Creating an Online Community of Action Researchers*, eds B. Bogner and P. Mompont-Gaillard (Strasbourg: Council of Europe Publishing), 149–170.
- Sprangel, J. R., Stavros, J. M., and Cole, M. (2011). Creating sustainable relationships using the SOAR framework, trust, and environmentalism: a research-based case study. *Internat. J. Train. Dev.* 15, 39–57.
- Stankov, L. (2013). Depression and life satisfaction among European and Confucian adolescents. *Psycholog. Assess.* 25, 1220–1234. doi: 10.1037/a0033794
- Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., and Lauro, C. (2005). PLS path modeling. *Comput. Stat. Data Anal.* 48, 159–205.
- Teo, T. S. H., Srivastava, S. C., and Jiang, L. (2008). Trust and electronic government success: an empirical study. *J. Manag. Inform. Syst.* 25, 99–132.
- UNICEF (2012). *Evaluation Report. Global Evaluation of Life Skills Education Programmes*. New York, NY: United Nations Children's Fund.
- Wong, P. T. P. (2020). The unheard cry of a successful Asian psychologist. *J. Psychol.* 154, 619–631. doi: 10.1080/00223980.2020.1820430
- Wong, P. T. P., and Roy, S. (2017). “Critique of positive psychology and positive interventions,” in *The Routledge International Handbook of Critical Positive Psychology*, eds N. J. L. Brown, T. Lomas, and F. J. Eiroa-Orosa (London: Routledge).
- Young, T., Macinnes, S., Jarden, A., and Colla, R. (2020). The impact of a wellbeing program imbedded in university classes: the importance of valuing happiness, baseline wellbeing and practice frequency. *Stud. High. Educ.* 2020:1793932. doi: 10.1080/03075079.2020.1793932

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 Gan, Beevi, Low, Lee and Hall. 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 L2 Motivational Self System and Grit in EFL Learners' Willingness to Communicate: A Study of Public School vs. Private English Language Institute Learners

Faramarz Ebn-Abbasi, Musa Nushi and Nazila Fattahi*

Department of English Language and Literature, Shahid Beheshti University, Tehran, Iran

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Abdullah Alamer,
Imam Muhammad Ibn Saud Islamic
University, Saudi Arabia
Yongliang Wang,
Henan University, China
Jing Yu,
Tianjin Normal University, China

*Correspondence:

Musa Nushi
m_nushi@sbu.ac.ir

Specialty section:

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

Received: 17 December 2021

Accepted: 09 February 2022

Published: 17 March 2022

Citation:

Ebn-Abbasi F, Nushi M and
Fattahi N (2022) The Role of L2
Motivational Self System and Grit
in EFL Learners' Willingness
to Communicate: A Study of Public
School vs. Private English Language
Institute Learners.
Front. Educ. 7:837714.
doi: 10.3389/feduc.2022.837714

Learning a second language (L2) is analogous to a journey replete with setbacks and discouragements. Given the pivotal role of communication in L2 learning, researchers have set out to identify factors that may influence L2 learners' willingness to communicate (L2 WTC). To this end, the present comparative study attempts to investigate the role of the L2 motivational self-system and grit in learners' L2 WTC in two different educational settings: public schools and private language institutes. L2 motivational self-system includes ideal L2 self (ILS) and ought-to L2self (OLS). Grit also comprises two lower-order constructs: perseverance of effort (POE) and consistency of interest (COI), which are examined separately in this study. The data collected through an online survey from 308 participants were analyzed using the Mann Whitney *u*-test, Spearman's rho, and multiple regression analysis. The results revealed a stronger ILS, POE, and a higher level of WTC among language institute learners. In contrast, public school students showed stronger OLS and lower levels of WTC. Moreover, unlike its counterpart, COI did not display a significant correlation or predictive power with/over L2 WTC in either context. The findings suggest that low levels of ILS and lack of POE can lead to lower levels of L2 WTC.

Keywords: willingness to communicate, L2 motivational self-system, grit, classroom enjoyment, EFL learners

INTRODUCTION

Promoting the target language use is an inseparable part of second language acquisition (SLA) pedagogy and research. For a long time, SLA was dominated by structural approaches, focusing on proficiency in vocabulary and grammar. However, with the blooming of communicative approaches, there has been a significant shift in scholars' perspectives toward SLA. They realized that communication should be considered not just as an essential part but also as the goal of second/foreign language learning (Fallah, 2014; Yashima et al., 2018). However, many scholars have pointed out that people differ in their willingness to communicate (WTC) generally and in the process of L2 acquisition specifically. WTC, in essence, highlights the fact that some individuals

seek chances to speak while others escape them and remain silent (MacIntyre, 2020). During the past decades, WTC has attracted vast research attention from L1 and L2 researchers as a critical factor influencing communication behavior.

Regarding L2 WTC, researchers' main concern was the dysfunctional aspects of non-communication (Richmond and McCroskey, 1989) while aiming to improve communication by removing these hindering constructs that prevent people from initiating or participating in oral communication. Still, with the blooming of positive psychology (PP) and the works of educational psychologists, more attention has been directed to the link between positive variables and L2 WTC (MacIntyre et al., 2019).

Grit is a recently proposed positive internal variable that has not been studied sufficiently in relation to SLA as much as other internal variables such as motivation and language aptitude. Defined as "perseverance and passion for long-term goals" by Duckworth et al. (2007, p. 1087), grit is believed to be a significant predictor of success and performance. Grit can play a crucial role in the L2 acquisition process since the process requires extended effort, perseverance, and patience (Alamer, 2021).

Another variable investigated in relation to WTC is L2 motivation. Early researchers used Gardner's (1985) socio-educational model as the basis of their investigation. Nevertheless, due to the limited applicability of this model in EFL contexts, Dörnyei's (2009) L2 motivational self-system (L2MSS) model garnered more attention recently. This new conceptualization of L2 motivation re-oriented the concept of motivation and dealt with possible selves and future self-guides in learning an L2. According to this model, there exist three primary sources of motivation in learning an L2: ILS, OLS, and L2 learning experience. Drawing upon L2MSS, researchers have examined the link between EFL students' ILS and OLS with L2 WTC (Peng, 2015; Yung, 2019; Lee and Lee, 2020; Lee et al., 2020).

Many language scholars have investigated the connection between L2 WTC and positive variables (e.g., MacIntyre and Mercer, 2014; Dewaele and Dewaele, 2018; MacIntyre et al., 2019). Nonetheless, there is still a paucity of comparative research on L2 WTC, especially between EFL learners studying English at different educational institutions. This paucity is strongly felt in regard to English programs in Iran which have always been subject to criticism in relation to L2 WTC. The significance of communication inside the language classroom is more highlighted in EFL contexts, such as Iran, where EFL learners find few if any, opportunities to communicate with native speakers of English.

REVIEW OF THE LITERATURE

Positive Psychology in L2

In the last two decades, positive psychology (PP) has caused an everlasting revolution in mainstream psychology, deviating from an exclusive focus on problems to more positive dimensions of humans' minds (Dewaele et al., 2019). Before the advent of PP, psychology was regarded as a science to repair the damaged.

It was replete with studies on "psychological disorders and the negative effects of environmental stressors such as parental divorce, death, and physical and sexual abuse" (Gallagher and Lopez, 2009, p. 4). Around the turn of the millennium, psychologists realized that psychology is not just about mental illnesses and how to heal them. With the rise of PP, these researchers started investing in positive aspects of human experience, such as hope, love, joy, optimism, etc. (Dewaele et al., 2019). Seligman and Csikszentmihalyi (2000), as the pioneers of the field, explain that PP aims to "catalyze a change in the focus of psychology from preoccupation only with repairing the worst things in life to also building positive qualities" (p. 5).

Throughout the years, a cognitive perspective dominated the SLA and applied linguistics resulting in sheer ignorance of the human psyche (Dewaele et al., 2019). With the flourishing of humanistic approaches such as the *whole-person approach* in the SLA domain, affective variables were brought into the spotlight. However, researchers believe that even after thriving of scholars' interest in affective variables, positive emotions have been overshadowed by negative emotions (Seligman et al., 2006). Therefore, little attention has been given to the role of positive variables in second and foreign language teaching and learning. Under the umbrella of PP, there has also been a shift from a deep-rooted focus on negative emotions, especially anxiety (Jiang and Dewaele, 2019) and cognitive perspectives (Li, 2020) in the SLA, to a more holistic analysis of emotions. Scholars have started to recognize the positive and negative emotional dimensions of language learning as significant assets. Before that, emotions were "poorly studied, poorly understood, seen as inferior to rational thought" (Swain, 2013, p. 11).

In accordance with positive psychology, researchers have started a *positive renaissance* or, in other words, a *flowering of positive psychology in the SLA*, too (Lee, 2020). Based on this view, MacIntyre and Gregersen (2012), pointed out that L2 teachers can help learners increase the degree of students' positive emotions in the classroom, thereby "broaden a person's perspective [and] opening the individual to absorb the language" (p. 193). PP in the SLA domain has been studied in relation to a variety of perspectives such as academic achievement, positive L2 self (e.g., Kikuchi and Lake, 2021), motivation (e.g., Martin, 2005), emotional intelligence (Li, 2020), WTC (e.g., Mystkowska-Wiertelak, 2021), and classroom enjoyment (Jin and Zhang, 2018). With the advent of PP, another variable gardening attention in the SLA domain is grit, a positive, non-cognitive trait which has been discussed in detail in the next section.

Grit

As the pioneers of the grit construct, Duckworth et al. (2007) defined the notion in a broad sense concerning success in the workplace, military, education, and sports. Grit gains immediate relevance in the SLA field, especially with the blooming of PP in this domain, since successful proficiency of an L2 is highly dependent on learners' sustained effort (Dörnyei and Ushioda, 2013). Researchers have established that grittier EFL students succeed in achieving higher scores in English (Strayhorn, 2014; Khajavy et al., 2021), employ persistent effort during learning

an L2 (Lake, 2015), and experience more positive classroom emotions (Wei et al., 2019).

Grit as a higher-ordered variable comprises two lower-order constructs: *Perseverance of Effort* (POE), which refers to a person's continued investment of energy in long-term pursuits; and *Consistency of Interests and passion for long-term goals* (COI), which refers to a person's durable passion for high-order goals over a long period regardless of failures, disappointments, or challenges (Lee, 2020; Teimouri et al., 2020). Since grit is not a fixed personality trait and can be malleable and teachable, the importance of measuring and estimating grit is gaining unprecedented attention. Duckworth et al.'s (2007) grit is a domain-general variable that could be applied to various fields such as military, workplace, education, and sports. As Teimouri et al. (2020) explained, there are two main reasons behind the incongruent findings of research on personality variables such as grit in the SLA field: "(a) its focus on super-traits instead of lower-traits with more relevancy to L2 learning and (b) its use of general personality measures instead of the situation- or domain-specific measures that assess students' traits in L2 situations" (p. 32). As a result, L2 researchers have recently extended the concept to L2 learning and developed L2 domain-specific grit notions.

Among these scholars, Teimouri et al. (2020) and Alamer (2021) proposed L2 domain-specific grit scales to help researchers in addressing grit in the language studies domain. The former examined the 12-scale validity through principal component analysis (PCA), which resulted in the reduction of the items from 12 items in the original scale to 9 items. They explained that grit as a higher-order factor consisting of two lower-order elements, namely POE and COI, should be examined separately. The latter, however, retained the 12 items and used bifactor-CFA to account for the specific constructs, COI and POE, and the general construct, Grit. The study examined the effect of L2 grit on future attainment of vocabulary. The study illustrated a weak predictive power of initial levels of grit on subsequent L2 learning. However, the author found that only students who sustained their levels of grit over time showed later language achievement. For this study that was conducted in Saudi Arabia, the author developed a 12-item L2 domain-specific grit scale. In the present study, we utilized Teimouri et al.'s (2020) grit scale.

L2 Motivational Self-System

Motivation is of great significance within the field of SLA. To highlight the importance of motivation in language learning, Dörnyei (2005) argues that motivation is the key impetus to start learning an L2 and to carry on through the long and arduous journey. He also points out that "all the other factors involved in SLA presuppose motivation to some extent" (p. 65). While early studies on L2 motivation investigated the issue through Gardner's (1985) socio-educational model, the concept was soon challenged by researchers (e.g., Dörnyei, 2005, 2009) who questioned the model's efficiency for addressing the expanding FL learning environment. As an answer to a call for a reconceptualization of motivation theory and inspired by self-discrepancy theory (Higgins, 1987) and possible selves theory (Markus and Nurius, 1986), Dörnyei (2005, 2009) presented a prominent model of learner motivation that is L2 motivational self-system. L2MSS

has three components; ILS is the L2-specific facet of one's ideal self; OLS refers to the attributes that one believes one ought to possess to avoid possible negative outcomes (Dörnyei, 2009, p. 29). And L2 Learning Experience focuses on the learner's present experience, covering a range of situated, "executive" motives related to the immediate learning environment (Dörnyei and Ryan, 2015, p. 88).

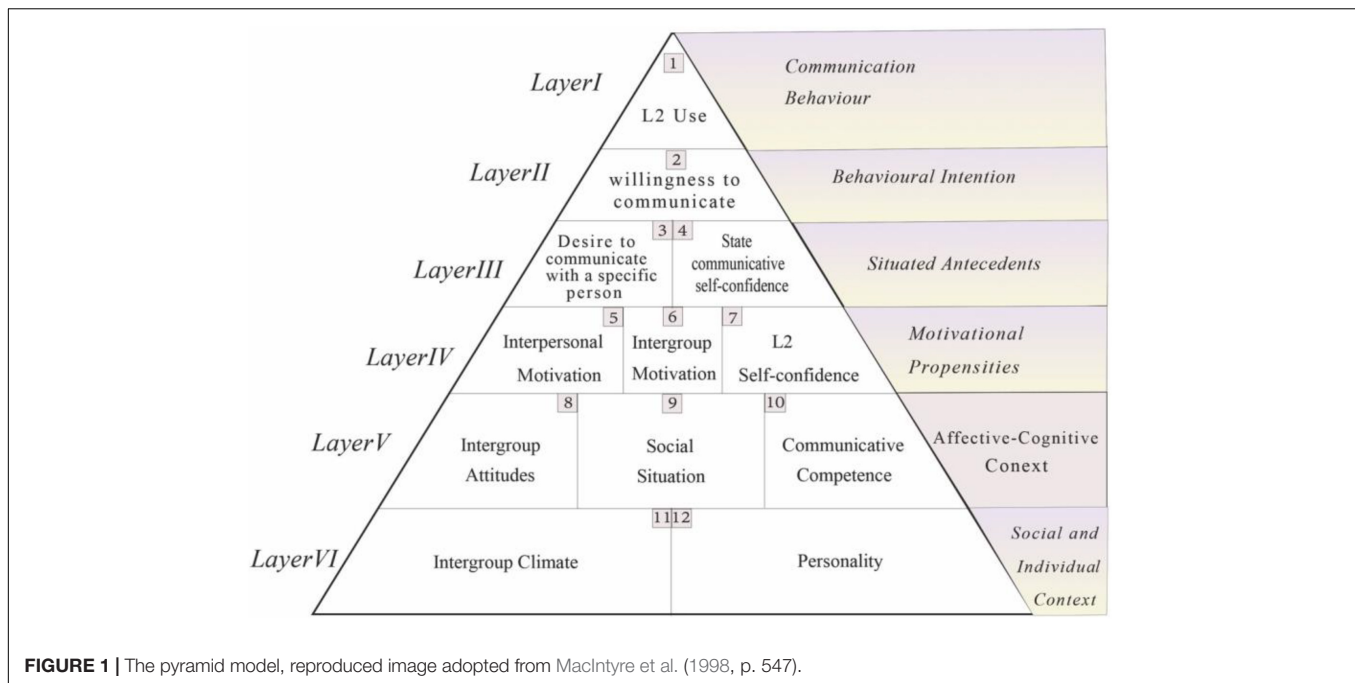
Over the past years, multiple empirical studies have been conducted to test and validate L2MSS in various learning environments (e.g., Csizér and Kormos, 2009; MacIntyre et al., 2009; Ryan, 2009; Taguchi et al., 2009; Kim and Kim, 2014). All the studies in the literature report solid confirmation for the proposed L2MSS. Inspired by Dörnyei's (2005) argument regarding a positive relationship between the ILS and L2 WTC, a new strand of research has focused on the relationship between L2 WTC and L2MSS. Munezane (2013) was the first researcher who tested the path between the ILS and L2 WTC and approved Dörnyei's (2005) hypothesis. Ever since, various studies have lent further support to the relationship between L2MSS components and L2 WTC (see Munezane, 2015, 2016; Darling and Chanyoo, 2018; Zulkepli, 2020).

L2 Learners' Willingness to Communicate

Verbal communication holds a central place among humans. The importance of *talk* in human interpersonal communication is an undeniable fact. As Khajavy et al. (2018) put it, the plethora of investigations on communication is because having communicative competence does not necessarily imply a willingness to use the language for authentic communication by itself. This can explain why some people with an acceptable linguistic competence may avoid communication and stay silent. In contrast, some with partial competence seek every opportunity to speak, which brings us to the issue of WTC.

Throughout the last decades, considerable research has been conducted on the various dimensions involved in WTC as an influential aspect of communication in L2. MacIntyre et al. (2003, p. 590) define WTC as "the probability of initiating communication, specifically talking when the opportunity arises." In its early stages, WTC was considered as a personality variable that is stable across time and in different situational contexts. More recently, however, L2 WTC has been presented not only as a stable tendency toward communication but also as a situational trait. Nowadays, researchers believe that internal variables such as L2 motivation and attitude toward L2 and external variables such as inter-group climate and social support can influence learners' L2 WTC (Zhang et al., 2019).

In their groundbreaking model of L2 WTC known as the Heuristic model of WTC in L2 (Figure 1; MacIntyre et al., 1998) emphasized that L2 WTC is dependent on complex interactions of varying factors such as context, personality traits, cognition, and emotion. This model illustrates the multifaceted effects of different individual and situational variables, including the variables of interest to the present study (i.e., grit and L2MSS), in each layer on L2 communication. This model consists of six layers which include 12 constructs. In this model, the above layers



(I, II, and III) include “more changeable, and context-dependent” (state) variables, and at the bottom layers of the pyramid (IV, V, and VI) more “stable social individual context” (trait) variables are placed (Dewaele and Pavelescu, 2021, p. 3).

Previous Studies

Lee (2020) examined the effects of grit and classroom enjoyment on Korean FL students' WTC. To obtain the data, he utilized Duckworth et al.'s (2007) grit scale, Dewaele and MacIntyre's (2014) FLE scale, and a five-item WTC scale (Lee et al., 2020). Lee realized that POE and classroom enjoyment (CE) correlate positively with L2 WTC among the participants. He explained that students with higher levels of POE who experience a positive environment in the classroom are more likely to show higher levels of WTC. On the other hand, COI was not observed as a significant predictor of L2 WTC. Another important finding of this study was the positive relationship between L2 WTC and the length of time devoted to learning English. In other words, students with more experience learning English tend to communicate more than others; on the contrary, Lee found out that there is a negative correlation between the age of students and their level of L2 WTC.

Cheng (2021) investigated the effect of language-specific grit and future self-guides on Taiwanese college English students' L2 WTC. In his study, L2 grit was measured using the 9-item scale developed by Teimouri et al. (2020), and L2 selves were estimated by Tseng et al.'s (2020) 16-item scale measuring four different L2 self-configuration (i.e., ILS/own, ILS/other, OLS/own, OLS/other). The data regarding the L2 WTC was collected using a questionnaire developed by Lee and Hsieh (2019). In disagreement with Lee (2020), the results of this study indicated that both POE and COI could predict L2 WTC. Yet, Cheng (2021) realized that COI showed a far less predictive power

compared to POE. Besides L2 grit, future L2 self-guides exerted a predictive power on L2 WTC. While both ILS, own and other, played a predictive role, only OLS/own could affect WTC, and OLS/other failed to do so.

Bursali and Öz (2017) focused on one of the components of L2MSS, that is, the ILS, and explored its relationship with L2 WTC of Turkish EFL learners. The data collected through a questionnaire from 56 students showed a positive and significant relationship between learners' ILS and their L2 WTC inside the classroom. The findings indicated that 32.1 percent of the participants had high, 30.4 percent had moderate, and 37.5 percent had low L2 WTC inside the classroom. However, they have noted that due to this study's limited sample, a generalization cannot be made concerning the relationship between L2MSS and L2 WTC at large.

Ghasemi (2018) set out to probe deeper into the WTC of Iranian English language major students and determine the variables contributing to L2 WTC in a single mode. The interrelationships between motivation, perceptual learning styles, cognitive variables, and language proficiency were under focus. Specifically, the role of the ILS, visual learning styles, and L2 confidence in predicting learners' language proficiency and WTC in real life were analyzed. 150 university freshmen students answered a five-section questionnaire. The results revealed that visual learning styles positively impact the ILS and are indirectly connected to L2 proficiency through the mediating function of motivation. These relationships support the fact that learners with more vivid images of the distance between their actual and ILS find language learning a goal-oriented activity and, consequently, are encouraged to attain higher levels of language proficiency. The desire to fill the gap between their current and ideal selves increases learner self-confidence, leading to higher levels of WTC.

To better understand the link between the ILS and L2 WTC, Lan et al. (2021) conducted a study that proposed a moderated mediation model incorporating grit and psychological shyness. To this end, the data were collected through a four-section questionnaire from 842 undergraduate students in three Chinese universities. Their findings revealed both direct and mediated links between the ILS and L2 WTC. ILS was positively associated with L2 WTC, and grit played a mediating role in this relationship. The authors argue that language learners with clear ILS images are grittier in achieving their language learning goals. Furthermore, learners with higher grit levels are more successful in facing learning tasks, such as participating in communicative activities.

In the present study, the aim is investigating the role of L2MSS and grit in EFL learners' L2 WTC in two different educational settings of public schools and private language institutes. To this end, the following research questions have been posited:

1. Is there a significant difference between the level of L2 WTC of Iranian EFL learners in public schools and private institutes?
2. To what extent do L2 grit (POE and COI) and L2MSS correlate with Iranian EFL public school students' L2 WTC?
3. To what extent do L2 grit (POE and COI) and L2MSS correlate with Iranian EFL private institute learners' L2 WTC?

MATERIALS AND METHODS

Participants

A total of 308 Iranian EFL learners from both public schools and private institutes completed a survey. Based on convenience sampling, the participants selected from the public schools were studying in the 11th and 12th grades at the time. As for the institute sample, language learners with elementary (A1) and pre-intermediate (A2) levels of English proficiency were selected so that the two groups of participants were at the same level of language proficiency. These levels were based on the Common European Framework of Reference for Languages (CEFR), which describes language ability on a six-point scale, from A1 for beginners to C2 for advanced students. Gender was not considered a variable in our study; hence, the inequality of gender distribution was not an issue. **Table 1** represents the demographic characteristics of each group that took part in this study.

Instruments

The data for each variable of interest were collected using an online survey consisting of four parts, viz. demographic information, L2 WTC questionnaire (Peng and Woodrow, 2010), grit questionnaire (Teimouri et al., 2020), and L2MSS questionnaire (Taguchi et al., 2009). In the first part of the survey, demographic information, two items in the public schools' questionnaire version were added to detect and remove those public-school students who were simultaneously attending private institutes and to identify students' grades. The

TABLE 1 | The participants' demographic information.

		Public school (N = 132)	Private institute (N = 176)
Gender	Male	63 (48%)	109 (61%)
	Female	69 (52%)	67 (39%)
	Total	132	176
Educational stage		11th G* (45.5%)	A1* (51%)
		12th G (55.5%)	A2* (49%)
Age	Mean	16.5	17.2

*G = Grade, A1 = Elementary, A2 = Pre-intermediate.

demographic section of the questionnaire administered to the learners in the private institutes also contained an additional item regarding their current English proficiency level. Moreover, to prevent age differences from affecting our results, the data collected from those institute learners older than 18 and younger than 15 were excluded from the analysis. The details of the different parts of the questionnaire are detailed below.

The L2 Learners' Willingness to Communicate Questionnaire

A modified version of L2 WTC by Weaver (2005) adopted by Peng and Woodrow (2010) was used in this study. The participants answered the 10 items such as *I am willing to do a role-play standing in front of the class in English* on a five-point Likert scale ranging from 1 (definitely not willing) to 5 (definitely willing).

The Grit Questionnaire

The researchers employed the L2 domain-specific grit scale developed by Teimouri et al. (2020) that consisted of two separate sections, namely POE and COI, in the classroom. Overall, the questionnaire consisted of 9 items including 5 items regarding POE such as *I am a diligent English language learner* and 4 items regarding COI such as *I am not as interested in learning English as I used to be*. The participants could select their responses on a five-point Likert scale ranging from 1 (not like me at all) to 5 (very much like me).

The L2 Motivational Self-System Questionnaire

To investigate L2MSS, 20 questions were adopted from Taguchi et al.'s (2009) study of L2MSS in three contexts of Japan, China, and Iran. To suit the purpose of the present study, among the 76 items of this scale, ten items measuring the ILS including *I can imagine a situation where I am speaking English with foreigners* and ten items such as *study English because close friends of mine think it is important* measuring the OLS were adopted. A five-point Likert scale ranging from strongly agree to strongly disagree was utilized in the present study.

Data Collection Procedure

To collect the required data for the present research, the following procedure was adopted. First, the questionnaires were translated into Persian by two professional translators. Google Forms was chosen to design an online survey. A pilot study was conducted with 50 students (25 public school students and 25 private

TABLE 2 | The reliability statistics for the three questionnaires of the survey.

Survey sections	Cronbach's alpha	N
WTC	0.97	10
POE	0.95	5
COI	0.95	4
ILS	0.93	10
OLS	0.76	10

language institute learners) from the target population. Based on the participants' comments, the translation of some of the items were edited to avoid confusion. After obtaining the schools ($N = 5$) and institutes ($N = 6$) board approval, the survey link was sent to the participants. The participants were assured of anonymity, and their participation was entirely voluntary.

RESULTS

Descriptive Statistics

Cronbach's alpha was used to estimate the reliability of each questionnaire comprising the survey. As seen in **Table 2**, the questionnaires enjoyed high levels of reliability to elicit L2 WTC, Grit, and L2MSS from all the respondents in public schools and private institutes.

For the first research question, the mean and standard deviation for each group were calculated (**Table 3**). The results of the descriptive analysis revealed that the mean of L2 WTC in public schools was 2.56, and the standard deviation was 1.29. **Table 3** displays the descriptive statistics for WTC, L2 Grit (POE and COI), and L2MSS (ILS and OLS) in public schools and private language institutes. The respondents in public schools mean performance regarding these variables was above 2, which indicates a medium performance on L2 Grit (POE and COI), and L2MSS (ILS and OLS).

The mean of L2 WTC in private institutes was 3.82, and the standard deviation was 1.22, revealing that the participants' level of L2 WTC was also medium. The mean for POE, COI, ILS, and OLS, which were 3.90, 2.33, 2.61, and 3.31, respectively, indicated

TABLE 4 | Mann-Whitney U -test results.

Mann-whitney U	Wilcoxon W	Z	Asymptotic. Sig. (2 tailed)
6789.50	15567.500	-6.278	0.000

that the participants' POE, COI, ILS, and OLS were at medium range for private institute learners. The POE has the highest mean (3.90) compared with that obtained for COI, ILS, and OLS.

To investigate how public school students and private institute learners differ in their level of L2 WTC, Kolmogorov-Smirnov and Shapiro-Wilk tests were first conducted to see if the data were normally distributed. The results showed the data did not enjoy normal distribution; therefore, the Mann-Whitney U -test, the non-parametric equivalent of the independent t -test, was used. As shown in **Table 4**, the Mann-Whitney U -test showed a significant difference between the level of L2 WTC in private and public school learners ($U = 6789.50$, $p < 0.001$).

Correlational Analyses

The Spearman correlation was employed to investigate the relationship between L2 WTC, L2 Grit (POE and COI), and L2MSS (ILS and OLS) among the EFL learners in public schools

TABLE 5 | Spearman's rho correlation, grit, L2MSS, and WTC for public school students and private institute learners.

			PS students' WTC	PI learners WTC
Spearman's rho	Grit (POE)	Correlation coefficient	0.588**	0.352**
		Sig. (2-tailed)	0.000*	0.000
	Grit (COI)	Correlation coefficient	0.110	0.135
		Sig. (2-tailed)	0.210	0.073
	ILS	Correlation coefficient	0.340**	0.372**
		Sig. (2-tailed)	0.000*	0.000*
	OLS	Correlation coefficient	0.679**	-0.117
		Sig. (2-tailed)	0.000*	0.123

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

TABLE 3 | Descriptive Statistics for WTC, L2 Grit (POE and COI), and L2MSS (ILS and OLS).

		N	Mean	SD	Skewness	Std. error skewness	Kurtosis	Std. error kurtosis
PS ^a	WTC	132	2.56	1.29	0.76	0.21	-1.29	0.41
	POE	132	3.00	1.40	0.22	0.21	-1.85	0.41
	COI	132	2.86	1.39	0.28	0.21	-1.76	0.41
	ILS	132	3.32	0.76	0.10	0.21	-0.82	0.41
	OLS	132	2.66	1.45	0.46	0.21	-1.63	0.41
PI ^b	WTC	176	3.82	1.21	-1.28	0.18	-0.23	0.36
	POE	176	3.90	1.04	-1.39	0.18	0.24	0.36
	COI	176	2.33	1.13	0.70	0.18	-0.87	0.36
	ILS	176	3.31	1.20	-0.28	0.18	-1.55	0.36
	OLS	176	2.61	0.92	0.04	0.18	-1.34	0.36

^aPublic schools.

^bPrivate institutes.

TABLE 6 | Multiple regression analysis for public school students.

Model summary				
Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.900 ^a	0.809	0.803	0.57427

^aPredictors: (Constant), OLS, ILS, COI, POE.**TABLE 7 |** The ANOVA results for public school students.

Institutes	Model	Sum of squares	Df	Mean square	F	Sig.
Public schools	1 Regression	177.710	4	44.428	134.716	0.000 ^a
	Residual	41.883	127	0.330		
	Total	219.593	131			

^aPredictors: (Constant), OLS, ILS, COI, POE.

and private institutes (questions two and three). The results revealed that, in public schools, there was a significant positive association between POE and WTC, [$r(132) = 0.588, p < 0.001$] (Table 5). In the case of L2MSS, the results indicated a significant positive correlation between OLS and L2 WTC, [$r(132) = 0.679, p < 0.001$] and, between ILS and L2 WTC, [$r(132) = 0.340, p < 0.001$]. Grit (COI), on the other hand, did not show a significant correlation with L2 WTC, [$r(132) = 0.110, p > 0.05$]. Regarding the private institute learners, the results indicated significant positive associations between POE, [$r(176) = 0.352, p < 0.001$], and ILS, [$r(176) = 0.372, p < 0.001$], and L2 WTC in private institutes' context. COI, [$r(176) = 0.135, p > 0.05$], and OLS, [$r(176) = -0.117, p > 0.05$], on the other hand, did not show significant correlations with private institutes' EFL learners' L2 WTC.

Regression Analyses

Two separate regression analyses were also performed to estimate the predictive power of L2 Grit (POE and COI) and L2MSS (ILS and OLS) for L2 WTC. Table 6 shows information about the regression model as a whole. The four variables together positively correlated with the total score at 0.90, which is high. The adjusted R^2 indicated that the model significantly predicted 81 percent of the variance in the population.

TABLE 8 | Coefficients of grit and L2MSS for public school students.

Coefficients ^a						
Model		Unstandardized coefficients		Standardized coefficients		Sig.
		B	Std. error	Beta	T	
Public schools	1 (Constant)	-0.851	0.280		-3.040	0.003
	POE	0.544	0.042	0.589	13.007	0.000
	COI	-0.025	0.037	-0.026	-0.669	0.505
	ILS	0.242	0.069	0.143	3.488	0.001
	OLS	0.392	0.041	0.439	9.600	0.000

^aDependent variable: WTC.**TABLE 9 |** Multiple regression analysis for private institute learners.

Model summary				
Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.508 ^a	0.258	0.241	1.06010

^aPredictors: (Constant), OLS, ILS, COI, POE.

The results of the ANOVA test (Table 7) revealed that this model differed significantly from the hypothetical model ($4,127 = 134.716, p < 0.001$).

Moreover, the coefficients table (Table 8) showed that POE as the Grit's first component ($\text{Beta} = 0.589, p < 0.001$), OLS ($\text{Beta} = 0.439, p < 0.001$), and the ILS ($\text{Beta} = 0.143, p = 0.001$) wielded significant predictive power over public schools' students L2 WTC. On the other hand, COI did not show any significant predictive power on L2 WTC among learners ($\text{Beta} = -0.026, p = 0.05$).

Table 9 demonstrates that the four variables positively correlate with the total score at $R = 0.508$, which is high. R-squared (R^2) shows 26 percent of the variance. The adjusted R^2 in the model reveals that L2 Grit (POE and COI) and L2MSS (ILS and OLS) could predict 24 percent of EFL learners' L2 WTC.

Examining the ANOVA results (Table 10), one can realize that this model differed significantly from the hypothetical model [$F(4, 171) = 14.855, p < 0.001$].

The table of coefficients (Table 11) displays that POE as the Grit's first component ($\text{Beta} = 0.265, p = 0.003$), and the ILS ($\text{Beta} = 0.266, p = 0.007$) exerted a significant predictive power

TABLE 10 | The ANOVA results for private institute learners.

The ANOVA results						
Private institute	Model	Sum of squares	Df	Mean square	F	Sig.
	1 Regression	66.777	4	16.694	14.855	0.000 ^a
	Residual	192.172	171	1.124		
	Total	258.949	175			

Dependent variable: WTC.

^aPredictors: (Constant), OLS, ILS, COI, POE.

TABLE 11 | Coefficients of Grit and L2MSS for private institute learners.

		Coefficients ^a					
Private institutes	Model	Unstandardized coefficients		Standardized coefficients		T	Sig.
		B	Std. Error	Beta			
1	(Constant)	1.335	0.533			2.503	0.013
	POE	0.309	0.101	0.265		3.047	0.003
	COI	0.110	0.085	0.103		1.295	0.197
	ILS	0.268	0.098	0.266		2.726	0.007
	OLS	0.052	0.108	0.040		0.487	0.627

^aDependent variable: WTC.

over learners' L2 WTC in the private institutes setting. COI and OLS, however, did not wield a significant effect on L2 WTC.

DISCUSSION

This study is among the first to investigate and compare the role of a positive variable (grit) and a motivational variable (L2 MSS) in L2 WTC among EFL learners in two educational settings. To some extent, the present study's results replicate but mainly extend previous findings on EFL learners' L2 WTC. To conclude the findings presented in the previous section and observe how the results of this particular study stand in relation to similar studies, a thorough discussion of the results is provided.

Willingness to Communicate in Public Schools and Private English Language Institutes

The descriptive data and the Mann-Whitney *U*-test demonstrated higher levels of L2 WTC among private institutes learners than public school students. From a socio-educational perspective, several reasons can be highlighted concerning this phenomenon.

Teaching Methodology

One way to interpret the significant difference between level L2 WTC is by highlighting the differences in the dominant methodology in these two settings. For decades, Iranian public schools' primary pedagogy of English learning suffered a sheer focus on "reading, translation, memorization, and grammar" (Hosseini Goodrich, 2020, p. 9). In 2013, the Ministry of Education succumbed to the long-called-for radical paradigm change from grammar-translation to a modified communicative language teaching (CLT). Nonetheless, criticism was still directed at the quality and effectiveness of the new English teaching method and its textbooks (Hosseini Goodrich, 2020). In essence, the revision in the English curriculum and textbooks aimed mainly at inserting the "Islamic-Iranian identity, national culture and local beliefs" (Mirhosseini and Khodakarami, 2015, p. 25), and in practice, public schools' English courses still focused on reading comprehension, grammar, and vocabulary development (Sadeghi and Richards, 2016). Writing practice is limited to composing decontextualized sentences, listening cannot be seen

anywhere in the syllabus, and speaking does not move beyond a few drills and short dialogues (Rahimi, 2009; Sadeghi and Richards, 2016; Khoshsima and Hashemi Toroujeni, 2017).

In contrast, by having the advantage of a decentralized system, private institutes utilize various international coursebooks and different approaches to teaching English. The majority of the private English institutes design their courses based on the principles of the CLT approach (Zhang and Rahimi, 2014) to prepare learners to communicate in both spoken and written modalities (Rahimi and Zhang, 2015; Moradkhani and Shirazizadeh, 2017). As a result, compared to private institute EFL learners, the lack of willingness to communicate among public school students is not surprising.

The Classroom Atmosphere

The majority of English classrooms in public schools suffer a strict teacher-dominated atmosphere which can discourage students' communicative participation. Compared to public schools, due to the less age difference between the EFL learners and their teachers, more appropriate teaching methodology, and the employment of novel activities, private institute classrooms are less strict and teacher-centered. This can be further confirmed by Zarei et al.'s (2019) findings suggesting that the teacher-centered style is among the debilitating factors regarding L2 WTC.

Teachers' Role and Test-Oriented Classrooms

Reflection upon teachers' role is crucial due to their effects on the educational process which ensures the quality of human and social resources of society (Zlatkovic and Petrovic, 2011). The EFL teachers play an influential role in students' WTC in many ways, including topic selection (Zarrinabadi, 2014), teaching style (Zarei et al., 2019), and appropriate activities selection (MacIntyre et al., 1998). Public school EFL teachers are forced/accustomed to avoiding employing L2 tasks, which can encourage and facilitate communication in the target language (Khoshsima and Hashemi Toroujeni, 2017; Hosseini Goodrich, 2020). This can be due to the centralized curriculum, the limited classroom time, excessive workload, inappropriate coursebooks, and crowded classes. This gains relevance in the Iranian public schools' context. English study for most public-school students in Iran is undertaken only to meet the need for passing examinations where performance-based tasks are disregarded in

the final exam criteria (Farhady and Hedayati, 2009). A similar observation has been reported for English language learning in the Chinese context where Peng and Woodrow (2010) found that EFL teachers and learners prioritize test-related skills such as vocabulary, reading, and writing over speaking in an exam-oriented context.

On the other hand, thanks to the decentralized curriculum, longer classroom hours, and more appropriate coursebooks, Iranian private institutes' teachers experience more freedom in choosing activities. Therefore, they are more likely to employ L2 communicative tasks such as information gap. Furthermore, these performance-based tasks are included in their examination criteria. Yashima et al. (2017) and Lee and Lee (2020) also further confirmed that inserting performance-based activities in the examination criteria can enhance students' WTC in Japanese and Korean contexts.

Grit and L2 Learners' Willingness to Communicate

In line with Teimouri et al. (2020), Cheng (2021), and Khajavy et al. (2021), who confirmed the superiority of a two-factor model over a single-factor model of language-specific grit, in the present study, the researchers examined grits' components (POE and COI) separately. This is in disagreement with Duckworth et al. (2007), who saw grit as a single variable. The results of Spearman's rho and multiple regression analyses revealed that grit (POE) has a positive correlation and significant predictive power with/over L2 WTC in both contexts, whereas grit (COI) showed no significant correlation or predictive power with/over EFL learners L2 WTC in either context. In a similar vein, Credé et al. (2017) proposed that "perseverance is a much better predictor of performance than either consistency or overall grit." (p. 502).

The positive relationship between grit (POE) and WTC can be reviewed in light of Fredrickson's (2003) "broaden and build" theory: The "broaden" side of this theory explains in what way positive emotions play a motivating role in encouraging the learners to explore new experiences and seek more learning opportunities. This theory expounds why learners who experience positive emotions in the classroom are more willing to participate in communicative activities. Therefore, we can infer from the analysis that grittier students who show determination and hard work are more likely to engage in communication despite setbacks. This is in agreement with Dörnyei and Ushioda's (2013) claim that success in an L2 is at the mercy of learners' sustained effort. These findings align well with those of the prior studies that confirmed a positive relationship between positive emotions and L2 WTC (Peng, 2015; Khajavy et al., 2018, 2021; Jiang and Dewaele, 2019).

Regarding grit (COI), during the lengthy and tedious language learning process, learners' interest in language learning may suffer some ups and downs due to the numerous setbacks and failures. Nonetheless, learners' hard work and sustained effort are not reliant on unwavering interest in the process. Similar findings in other contexts such as Korea (Lee, 2020), Switzerland (Karlen

et al., 2019), and China (Feng and Papi, 2020) further confirms the results of this article.

L2 Motivational Self-System and L2 Learners' Willingness to Communicate

According to the Spearman correlation, both ILS and OLS showed a significant relationship with L2 WTC in public school setting. However, in the case of the learners in the private institutes, only ILS correlated with L2 WTC. Additionally, the multiple regression revealed that OLS was the second predictor of WTC in public school, while it failed to predict WTC in private language institutes. On the other hand, the ILS was the second predictor of L2 WTC in private language institutes and the third in public schools.

As noted, between the two components of L2MSS, OLS plays a more vital role on L2 WTC of public school students, which suggests that these students tend to become motivated to learn English to meet the expectations of their significant others, such as parents and teachers. Although the OLS usually appears as the weakest component of L2MSS (Csizér and Kormos, 2009; Aubrey, 2014), research in collectivist cultures has proved the opposite. It has been claimed that learners with collectivist cultures, such as Thai and Chinese EFL learners, tend to have stronger OLS (Patterson and Smith, 2003; Taguchi et al., 2009; Lanvers, 2016; Darling and Chanyoo, 2018). In concordance with these findings, the results of the present study invite the possibility that Iranian collectivist culture has led to a significant relationship between OLS and school students' L2 WTC.

The strong impact of the ILS on the L2 WTC of private language institute learners denotes that in this group of learners, an ideal future image rather than external expectations and obligations aid language learners to put aside any fears they might have of communication and engage in the communicative act. The general tenor of previous studies on L2MSS has introduced the ILS as the strongest dimension of the model (Csizér and Kormos, 2009; Dörnyei and Ushioda, 2009; Papi, 2010; Kim, 2012). Thus, the present study's findings tie in well with the literature. Furthermore, the results of the current research are in line with studies that have reported a positive relationship between ILS and L2 WTC among Turkish (Bursali and Öz, 2017, 2018; Sak, 2020), Thai (Darling and Chanyoo, 2018), Japanese (Munezane, 2013, 2015, 2016), Chinese (Peng, 2015; Lan et al., 2021; Li and Liu, 2021), and Iranian (Ghasemi, 2018) EFL learners.

CONCLUSION

The present study set out to investigate if/how L2MSS and grit correlate/predict L2 WTC in two educational settings of public schools and private language institutes. The findings pointed out that public school students with stronger ILS and POE enjoyed higher levels of WTC; meanwhile, private institute learners with a stronger OLS were less willing to communicate compared to their counterparts. Moreover, while POE was the strongest predictor of WTC in both settings, COI did not display a

significant correlation or predictive power with/over L2 WTC in either context.

The implications of this study mainly call for changes in the overall Iranian system of education to improve WTC among EFL learners. The present research highlights the role of grit (POE) and L2MSS (ILS) in this process. The fact that learners with higher levels of ILS (private institute learners) were more willing to communicate than those learners with strong OLS (public school students) underlines Dörnyei's (2009) call for the promotion of motivation in academic settings, which can be achieved by enhancing what he calls learner's vision. As the centerpiece of classroom instruction, teachers play an essential role in helping students hold and build more vivid ILS images. Creating vivid ILS images may be achieved by visualization and goal-setting activities such as writing future-oriented autobiographies, creating L2 goals for the semester and for the next 10 years (Al-Murtadha, 2019; Safdari, 2019); establishing short-term goals that are tangible and achievable through more active participation in communicative classroom activities (Lan et al., 2021). Meanwhile, by having high levels of POE, EFL learners would be more likely to exert themselves to achieve these goals (communication in the target language in this case) despite setbacks and failures.

Grit has a malleable nature meaning that it can be increased through practice and instruction. EFL teachers should be aware of the importance of grit and how it can help learners in the possible tedious and challenging (e.g., participating in communicative activities) language learning process (Wang et al., 2021). They can ensure that their learners are familiar with the positive role of POE by giving lectures, introducing successful people, and encouraging them to be persistent to achieve their objectives. Keegan (2017) explains that being "gifted" is not the only predictor of success in language learning. Therefore, EFL students should be informed of the malleable nature of intelligence which explains that besides talent, other appreciated factors such as assiduousness and perseverance can aid them to achieve their goals in relation to L2 WTC. This gains relevance in the EFL context since many teachers, regardless

of where they are teaching, struggle to engage their students in communication.

It needs to be acknowledged that generalizations based on the findings of this study must be made with caution. Since the researchers intended to compare the results obtained from public schools with those of private institutes, students with the same level of proficiency in these contexts had to be selected. Due to the limited curriculum in Iranian public schools, English proficiency among students remains low, almost equal to the lowest level of proficiency in private institutes. Moreover, more research is called for to compare the WTC of learners in different educational settings and studies that cover the effect of gender on EFL learners' grit and L2 MSS in relation to L2 WTC.

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' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

NF and FE-A conceived the research idea and wrote up the article which was read and commented on multiple times by MN. NF and FE-A revised the article which was given final approval by MN. MN submitted the manuscript and did all the corresponding with the journal. All authors contributed to the article and approved the submitted version.

REFERENCES

- Alamer, A. (2021). Grit and language learning: construct validation of L2-Grit scale and its relation to later vocabulary knowledge. *Educ. Psychol.* 41, 1–18. doi: 10.1080/01443410.2020.1867076
- Al-Murtadha, M. (2019). Enhancing EFL learners' willingness to communicate with visualization and goal-setting activities. *TESOL Q.* 53, 133–157. doi: 10.1002/tesq.474
- Aubrey, S. (2014). Development of the L2 motivational self system: english at a university in Japan. *JALT J.* 36, 153–174. doi: 10.37546/JALTJJ.36.2-1
- Bursali, N., and Öz, H. (2017). The relationship between ideal L2 self and willingness to communicate inside the classroom. *Int. J. Higher Educ.* 6, 229–239. doi: 10.5430/ijhe.v6n4p229
- Bursali, N., and Öz, H. (2018). The relationship between L2 motivational self system and willingness to communicate in learning English as a foreign language. *J. Lang. Linguistic Stud.* 14, 1–11.
- Cheng, H. F. (2021). The effect of language-specific grit and future self-guides on willingness to communicate in the second language. *English Teach. Learn.* 45, 283–301. doi: 10.1007/s42321-021-00081-7
- Credé, M., Tynan, M. C., and Harms, P. D. (2017). Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Person. Soc. Psychol.* 113, 492–511. doi: 10.1037/pspp0000102
- Csizér, K., and Kormos, J. (2009). "Learning experiences, selves and motivated learning behavior: a comparative analysis of structural models for Hungarian secondary and university learners of English," in *Motivation, Language Identity and the L2 Self*, eds Z. Dörnyei and E. Ushioda (Bristol: Multilingual Matters), 98–209.
- Darling, W. E., and Chanyoo, N. (2018). Relationships of L2 motivational self-system components and willingness to communicate in English among Thai undergraduate students. *NIDA J. Lang. Commun.* 23, 1–22.
- Dewaele, J. M., 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. M., and Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the FL classroom. *J. Eur. Second Lang. Assoc.* 2, 24–37. doi: 10.22599/jesla.37
- 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/ssl.2014.4.2.5

- Dewaele, J. M., and Pavlescu, L. M. (2021). The relationship between incommensurable emotions and willingness to communicate in English as a foreign language: a multiple case study. *Innovat. Lang. Learn. Teach.* 15, 66–80. doi: 10.1080/17501229.2019.1675667
- Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2009). "The L2 motivational self system," in *Motivation, Language Identity and the L2 Self*, eds Z. Dörnyei and E. Ushioda (Bristol: Multilingual Matters), 9–42.
- Dörnyei, Z., and Ryan, S. (2015). *The Psychology of the Language Learner Revisited*. New York, NY: Routledge.
- Dörnyei, Z., and Ushioda, E. (eds) (2009). *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters.
- Dörnyei, Z., and Ushioda, E. (2013). *Teaching and Researching: Motivation*. Milton Park: Routledge.
- Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *J. Person. Soc. Psychol.* 92, 1087–1101. doi: 10.1037/0022-3514.92.6.1087
- Fallah, N. (2014). Willingness to communicate in English, communication self-confidence, motivation, shyness and teacher immediacy among Iranian English-major undergraduates: a structural equation modeling approach. *Learn. Individ. Differ.* 30, 140–147. doi: 10.1016/j.lindif.2013.12.006
- Farhady, H., and Hedayati, H. (2009). Language assessment policy in Iran. *Annu. Rev. Appl. Linguistics* 29, 132–141. doi: 10.1017/S0267190509090114
- Feng, L., and Papi, M. (2020). Persistence in language learning: the role of grit and future self guides. *Learn. Individ. Differ.* 81, 1–10. doi: 10.1016/j.lindif.2020.101904
- Fredrickson, B. L. (2003). The value of positive emotions: the emerging science of positive psychology is coming to understand why it's good to feel good. *Am. Sci.* 91, 330–335.
- Gallagher, M. W., and Lopez, S. J. (2009). Positive expectancies and mental health: identifying the unique contributions of hope and optimism. *J. Posit. Psychol.* 4, 548–556. doi: 10.1080/17439760903157166
- Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. London: Edward Arnold.
- Ghasemi, A. A. (2018). Ideal L2 self, visual learning styles, and L2 self confidence in predicting language proficiency and L2 WTC: a structural equation modeling. *English Teach. Learn.* 42, 185–205. doi: 10.1007/s42321-018-0010-8
- Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychol. Rev.* 94:319. doi: 10.1037/0033-295X.94.3.319
- Hosseini Goodrich, N. (2020). English in Iran: a sociolinguistic profile. *World Englishes* 39, 482–499. doi: 10.1111/weng.12491
- 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. *J. Multilingual Multicultural Dev.* 41, 246–263. doi: 10.1080/13670050.2018.1526253
- Karlen, Y., Suter, F., Hirt, C., and Merki, K. M. (2019). The role of implicit theories in students' grit, achievement goals, intrinsic and extrinsic motivation, and achievement in the context of a long-term challenging task. *Learn. Individ. Differ.* 74:101757. doi: 10.1016/j.lindif.2019.101757
- Keegan, K. (2017). Identifying and building grit in language learners. *Engl. Teach. Forum* 55, 2–9.
- 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
- Khajavy, G. H., MacIntyre, P. D., and Hariri, J. (2021). A closer look at grit and language mindset as predictors of foreign language achievement. *Stud. Second Lang. Acquisition* 43, 379–402. doi: 10.1017/S0272263120000480
- Khoshshima, H., and Hashemi Toroujeni, S. M. (2017). Context in English language teaching program: differentiating between Iranian public schools and private institutes. *J. English Lang. Teach. Linguistics* 2, 13–36. doi: 10.21462/jeltl.v2i1.35
- Kikuchi, K., and Lake, J. (2021). "Positive psychology, positive L2 self, and L2 motivation: a longitudinal investigation," in *Positive Psychology in Second and Foreign Language Education*, eds K. Budzińska and O. Majchrzak (Cham: Springer), 79–94.
- Kim, T. Y. (2012). The L2 motivational self system of Korean EFL students: cross-grade survey analysis. *English Teach.* 67, 29–56. doi: 10.15858/engtea.67.1.201203.29
- Kim, T.-Y., and Kim, Y.-K. (2014). A structural model for perceptual learning styles, the ideal L2 self, motivated behavior and English proficiency. *System* 46, 14–27. doi: 10.1016/j.system.2014.07.007
- Lake, J. (2015). *Positive Psychology and Second Language Motivation: Empirically Validating a Model of Positive L2 Self*. Unpublished Doctoral dissertation. Philadelphia, PA: Temple University.
- Lan, G., Nikitina, L., and Woo, W. S. (2021). Ideal L2 self and willingness to communicate: a moderated mediation model of shyness and grit. *System* 99:102503. doi: 10.1016/j.system.2021.102503
- Lanvers, U. (2016). Lots of selves, some rebellious: developing the self discrepancy model for language learners. *System* 60, 79–92. doi: 10.1016/j.system.2016.05.012
- Lee, J. S. (2020). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *J. Multilingual Multicultural Dev.* 1–17. doi: 10.1080/01434632.2020.1746319
- Lee, J. S., and Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System* 82, 63–73. doi: 10.1016/j.system.2019.03.002
- Lee, J. S., and Lee, K. (2020). Role of L2 motivational self system on willingness to communicate of Korean EFL university and secondary students. *J. Psycholinguist. Res.* 49, 147–161.
- Lee, J. S., Sylvén, L. K., and Lee, K. (2020). Cross-cultural insights into Korean and Swedish secondary school students' willingness to communicate in a second language. *J. Multilingual Multicultural Dev.* 52, 1–15. doi: 10.1007/s10936-019-09675-6
- Li, C. (2020). A positive psychology perspective on Chinese EFL students' trait emotional intelligence, foreign language enjoyment and EFL learning achievement. *J. Multilingual Multicultural Dev.* 41, 246–263. doi: 10.1080/01434632.2019.1614187
- Li, T., and Liu, Z. (2021). Exploring effects of the second language motivational self system on Chinese EFL learners' willingness to communicate in English and implications for L2 education. *J. Higher Educ. Res.* 2:169.
- MacIntyre, P. D. (2020). Expanding the theoretical base for the dynamics of willingness to communicate. *Stud. Second Lang. Learn. Teach.* 10, 111–131. doi: 10.14746/ssl.2020.10.1.6
- MacIntyre, P. D., Baker, S., Clément, R., and Donovan, L. A. (2003). Talking in order to learn: willingness to communicate and intensive language programs. *Canadian Modern Lang. Rev.* 59, 589–608. doi: 10.3138/cmlr.59.4.589
- MacIntyre, P. D., Clément, R., Dörnyei, Z., and Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. *Modern Lang. J.* 82, 545–562. doi: 10.1111/j.1540-4781.1998.tb05543.x
- MacIntyre, P. D., 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., Mackinnon, S. P., and Clément, R. (2009). "Toward the development of a scale to assess possible selves as a source of language learning motivation," in *Motivation, Language Identity and the L2 Self*, eds Z. Dörnyei and E. Ushioda (Bristol: Multilingual Matters), 193–214. doi: 10.21832/9781847691293-011
- MacIntyre, P. D., and Mercer, S. (2014). Introducing positive psychology to SLA. *Stud. Second Lang. Learn. Teach.* 4, 153–172. doi: 10.3389/fpsyg.2020.02142
- MacIntyre, P. D., Ross, J., and Sparling, H. (2019). Flow experiences and willingness to communicate: connecting Scottish Gaelic language and traditional music. *J. Language Soc. Psychol.* 38, 536–545. doi: 10.1177/0261927x19867364
- Markus, H., and Nurius, P. (1986). Possible selves. *Am. Psychol.* 41, 954–969. doi: 10.1037/0003-066X.41.9.954
- Martin, A. J. (2005). The role of positive psychology in enhancing satisfaction, motivation, and productivity in the workplace. *J. Organ. Behav. Manag.* 24, 113–133. doi: 10.1300/J075v24n01_07
- Mirhosseini, S. A., and Khodakarami, S. (2015). "A glimpse of contrasting de jure–de facto ELT policies in Iran," in *English Language Teaching in the Islamic*

- Republic of Iran: *Innovations, Trends and Challenges*, ed. C. Kennedy (London: British Council), 23–32.
- Moradkhani, S., and Shirazizadeh, M. (2017). Context-based variations in EFL teachers' reflection: the case of public schools versus private institutes in Iran. *Reflective Pract.* 18, 206–218. doi: 10.1080/14623943.2016.1267002
- Munezane, Y. (2013). Attitudes, affect and ideal L2 self as predictors of willingness to communicate. *Eurosla Yearbook* 13, 176–198. doi: 10.1075/eurosla.13.09mun
- Munezane, Y. (2015). Enhancing willingness to communicate: relative effects of visualization and goal setting. *Modern Lang. J.* 99, 175–191. doi: 10.1111/modl.12193
- Munezane, Y. (2016). Motivation, ideal self and willingness to communicate as the predictors of observed L2 use in the classroom. *Eurosla Yearbook* 16, 85–115. doi: 10.1075/eurosla.16.04mun
- Mystkowska-Wiertelak, A. (2021). "The link between different facets of willingness to communicate, engagement and communicative behavior in task performance," in *Positive Psychology in Second and Foreign Language Education*, eds K. Budzińska and O. Majchrzak (Cham: Springer), 95–113. doi: 10.1007/978-3-030-64444-4_6
- Papi, M. (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: a structural equation modelling approach. *System* 38, 467–479. doi: 10.1016/j.system.2010.06.011
- Patterson, P. G., and Smith, S. (2003). A cross-cultural study of switching barriers and propensity to stay with service providers. *J. Retail.* 79, 107–120. doi: 10.1016/S0022-4359(03)00009-5
- Peng, J., and Woodrow, L. (2010). Willingness to communicate in English: a model in the Chinese EFL classroom context. *Lang. Learn.* 60, 834–876. doi: 10.1111/j.1467-9922.2010.00576.x
- Peng, J. E. (2015). L2 motivational self system, attitudes, and affect as predictors of L2 WTC: an imagined community perspective. *Asia-Pacific Educ. Researcher* 24, 433–443. doi: 10.1007/s40299-014-0195-0
- Rahimi, M. (2009). The role of teacher's corrective feedback in improving Iranian EFL learners' writing accuracy over time: is learner's mother tongue relevant? *Read. Writ.* 22, 219–243. doi: 10.1007/s11145-008-9139-5
- Rahimi, M., and Zhang, L. J. (2015). Exploring non-native English-speaking teachers' cognitions about corrective feedback in teaching English oral communication. *System* 55, 111–122. doi: 10.1016/j.system.2015.09.006
- Richmond, V. P., and McCroskey, J. C. (1989). "Willingness to communicate and dysfunctional communication processes," in *Interpersonal Communication Processes*, eds C. V. Roberts and K. W. Watson (New Orleans, LA: SPECTRA Inc.), 292–318.
- Ryan, S. (2009). "Self and identity in L2 motivation in Japan: the ideal L2 self and Japanese learners of English," in *Motivation, Language Identity and the L2 Self*, eds Z. Dörnyei and E. Ushioda (Bristol: Multilingual Matters), 43–120.
- Sadeghi, K., and Richards, J. C. (2016). The idea of English in Iran: an example from Urmia. *J. Multilingual Multicultural Dev.* 37, 419–434. doi: 10.1080/01434632.2015.1080714
- Safdari, S. (2019). Operationalizing L2 motivational self system: improving EFL learners' motivation through a vision enhancement program. *Lang. Teach. Res.* 25, 282–305. doi: 10.1177/1362168819846597
- Sak, M. (2020). The role of ideal L2 self in predicting L2 willingness to communicate inside and outside the classroom. *Eur. J. Appl. Linguistics* 6, 189–203. doi: 10.32601/ejal.775798
- Seligman, M. E. P., and Csikszentmihalyi, M. (2000). Positive psychology [Special issue]. *Am. Psychol.* 55, 5–14. doi: 10.1037/0003-066X.55.1.5
- Seligman, M. E., Rashid, T., and Parks, A. C. (2006). Positive psychotherapy. *Am. Psychol.* 61, 774–788. doi: 10.1037/0003-066X.61.8.774
- Strayhorn, T. L. (2014). What role does grit play in the academic success of black male collegians at predominantly white institutions? *J. Afr. Am. Stud.* 18, 1–10. doi: 10.1007/s12111-012-9243-0
- Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Lang. Teach.* 46, 195–207. doi: 10.1017/S0261444811000486
- Taguchi, T., Magid, M., and Papi, M. (2009). "The L2 motivational self system among Japanese, Chinese and Iranian learners of English: a comparative study," in *Motivation, Language Identity and the L2 Self*, eds Z. Dörnyei and E. Ushioda (Bristol: Multilingual Matters), 66–97. doi: 10.21832/9781847691293-005
- Teimouri, Y., Plonsky, L., and Tabandeh, F. (2020). L2 grit: passion and perseverance for second-language learning. *Lang. Teach. Res.* doi: 10.1177/1362168820921895
- Tseng, W. T., Cheng, H. F., and Gao, X. (2020). Validating a motivational self-guide scale for language learners. *Sustainability* 12:6468. doi: 10.3390/su12166468
- Wang, Y., Derakhshan, A., and Zhang, L. J. (2021). Researching and practicing positive psychology in second/foreign language learning and teaching: the past, current status and future directions. *Front. Psychol.* 12:731721. doi: 10.3389/fpsyg.2021.731721
- Weaver, C. (2005). Using the Rasch model to develop a measure of second language learners' willingness to communicate within a language classroom. *J. Appl. Meas.* 6, 396–415.
- Wei, H., Gao, K., and Wang, W. (2019). Understanding the relationship between grit and foreign language performance among middle school students: the roles of foreign language enjoyment and classroom environment. *Front. Psychol.* 10:1508. doi: 10.3389/fpsyg.2019.01508
- Yashima, T., MacIntyre, P. D., and Ikeda, M. (2018). Situated willingness to communicate in an L2: interplay of individual characteristics and context. *Lang. Teach. Res.* 22, 115–137. doi: 10.1177/1362168816657851
- Yashima, T., Nishida, R., and Mizumoto, A. (2017). Influence of learner beliefs and gender on the motivating power of L2 selves. *Modern Lang. J.* 101, 691–711. doi: 10.1111/modl.12430
- Yung, K. W. H. (2019). Exploring the L2 selves of senior secondary students in English private tutoring in Hong Kong. *System* 80, 120–133. doi: 10.1016/j.system.2018.11.003
- Zarei, N., Saeidi, M., and Ahangari, S. (2019). Exploring EFL teachers' socio-affective and pedagogic strategies and students' willingness to communicate with a focus on Iranian culture. *Educ. Res. Int.* 2019:3464163. doi: 10.1155/2019/3464163
- Zarrinabadi, N. (2014). Communicating in a second language: investigating the effect of teacher on learners' willingness to communicate. *System* 42, 288–295. doi: 10.1016/j.system.2013.12.014
- Zhang, L. J., and Rahimi, M. (2014). EFL learners' anxiety level and their beliefs about corrective feedback in oral communication classes. *System* 42, 429–439. doi: 10.1016/j.system.2014.01.012
- Zhang, J., Beckmann, N., and Beckmann, J. F. (2019). One situation doesn't fit all: variability and stability of state willingness to communicate in a Chinese College English classroom. *Lang. Teach. Res.* doi: 10.1177/1362168819891330
- Zlatkovic, B., and Petrovic, D. (2011). Pre-service teacher training in Serbia: analysis of the compatibility of teachers faculties curricula. *J. Educ.* 60, 651–663.
- Zulkepli, N. (2020). L2 motivational self-system as predictors of out-of-class willingness to communicate. *J. Lang. Linguistic Stud.* 16, 1898–1908. doi: 10.17263/jlls.851017

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 Ebn-Abbasi, Nushi and Fattahi. 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.



Identifying Determinants of Dyslexia: An Ultimate Attempt Using Machine Learning

Sietske Walda*, Fred Hasselman and Anna Bosman

Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

OPEN ACCESS

Edited by:

Alessandro Antonietti,
Catholic University of the Sacred
Heart, Italy

Reviewed by:

Laurice Joseph,
The Ohio State University,
United States
Iza Sazanita Isa,
Universiti Teknologi MARA, Cawangan
Pulau Pinang, Malaysia

*Correspondence:

Sietske Walda
Sietske.walda@ru.nl

Specialty section:

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

Received: 04 February 2022

Accepted: 09 March 2022

Published: 07 April 2022

Citation:

Walda S, Hasselman F and
Bosman A (2022) Identifying
Determinants of Dyslexia: An Ultimate
Attempt Using Machine Learning.
Front. Psychol. 13:869352.
doi: 10.3389/fpsyg.2022.869352

Research based on traditional linear techniques has yet not been able to clearly identify the role of cognitive skills in reading problems, presumably because the process of reading and the factors that are associated with reading reside within a system of multiple interacting and moderating factors that cannot be captured within traditional statistical models. If cognitive skills are indeed indicative of reading problems, the relatively new nonlinear techniques of machine learning should make better predictions. The aim of the present study was to investigate whether cognitive factors play any role in reading skill, questioning (1) the extent to what cognitive skills are indicative of present reading level, and (2) the extent to what cognitive skills are indicative of future reading progress. In three studies with varying groups of participants (average school-aged and poor readers), the results of four supervised machine learning techniques were compared to the traditional General Linear Models technique. Results of all models appeared to be comparable, producing poor to acceptable results, which are however inadequate for making a thorough prediction of reading development. Assumably, cognitive skills are not predictive of reading problems, although they do correlate with one another. This insight has consequences for scientific theories of reading development, as well as for the prevention and remediation of reading difficulties.

Keywords: dyslexia, cognitive skills, reading development, machine learning, word decoding and reading outcomes

INTRODUCTION

Dyslexia is characterized by severe problems with learning basic subskills of reading and spelling, often resulting in delays in broader literacy skills and unequal opportunities in education and society. In the Netherlands, approximately 10–15% of children in primary and secondary education are diagnosed with dyslexia (Inspectie van het Onderwijs, 2019). Research indicates that dyslexia is sometimes associated with impairments in various cognitive skills, such as phonological awareness, orthographical awareness, syntactic awareness, working memory, as well as general cognitive deficits such as impairments of attention, rapid naming, and self-control (Vellutino et al., 2004; Catts, 2017).

The role of cognitive skills in dyslexia has been subject to varying interpretations. Some authors advocate the existence of different types of reading problems, explaining them by a dysfunction of different cognitive functions. Others interpret these cognitive dysfunctions as a consequence of a more general deficit that causes both cognitive dysfunctions and reading problems

(Parrila et al., 2019). Most theories, however, suggest that dyslexia is caused by some sort of brain dysfunction, resulting in an “all or nothing” diagnosis of dyslexia: Children with the brain dysfunction do and children without the brain dysfunction do not have dyslexia. Still other theories suggest that symptoms of dyslexia result from a multi-factorial interplay of (genes), cognitive skills and environment, resulting in a continuum of more to less severe reading and spelling problems (Vellutino et al., 2004; Catts, 2017).

The relevance of the role of cognitive skills in reading problems differs between varying perspectives. From a theoretical perspective, knowledge about cognitive skills involved in learning to read can lead to insights in the nature of reading and processes that accompany the development of automated reading skill (e.g., see Hammill, 2004). From a prevention perspective, knowledge about cognitive skills predictive of reading difficulties can help identify children at risk for developing reading difficulties. If children at risk are identified early, instruction and exercises can be adjusted to lower the chance of developing reading problems (e.g., see Hammill, 2004; Al Otaiba and Fuchs, 2006; Stuebing et al., 2015). From a remediation perspective, it is argued that knowledge about cognitive skills affecting the development of automated reading skill will provide more effective interventions for children lagging behind. When strengths and weaknesses on these particular cognitive skills are known, they could be part of the remediation process. This knowledge is of great importance, because a substantial part of children with dyslexia receiving early literacy intervention respond below expectations and are labeled treatment resisters (e.g., see Al Otaiba and Fuchs, 2006; Stuebing et al., 2015).

The aim of the present study was to find cognitive factors of varying origin that play a role in the development of reading skill in children with dyslexia and children with typical reading development. In this introduction, we will first discuss results of previous research (see **Table 1**). Previous research demonstrates two concerns: (1) empirical identification of causal factors, and (2) the application of research models adequate for identifying factors involved in reading *level* and in reading *progress*. Next, ideas on potential models of reading development are considered, resulting in a discussion of potential utility of machine learning techniques in an ultimate attempt to identify cognitive skills that are involved in learning to read. Finally, we will present the aims and research questions of the present study.

The relationship between reading level and various cognitive skills has been well established by previous research. The statistical technique most frequently used for identifying factors that are important in the process of learning to read is correlation analysis. Even though correlations are bivariate and cannot determine any directional or causal influences, researchers assume that a correlational analysis provides information concerning factors that might be involved in learning to read. In search for causal factors, significant correlations are insufficient to establish causal relationships with learning to read. However, researchers assume that in the search for causal factors, significant correlations should at least be present (corresponding to the fifth axiom of Spinoza, 1678/1928, p. 15). Therefore, factors that prove to correlate with reading level are of interest for researchers who

seek to find causal relationships between cognitive factors and reading skill (Hammill, 2004; Stuebing et al., 2015).

In the first three rows of **Table 1** meta-analyses on cognitive skills and reading level are presented. Several cognitive skills tend to correlate significantly (not necessarily substantially) with reading level. Age or reading skill often have moderating effects on these correlations (see Hammill and McNutt, 1981; Swanson et al., 2003; Melby-Lervåg, 2012). However, there are three issues that impede the interpretation of correlation between cognitive skills and reading level: (1) For some variables, the amount of overlap with reading skill is unclear. For example, variables such as academic achievement, spelling, pseudoword reading and letter identification could be considered reading skills or at least literacy skills. These skills rather mirror subskills of reading and literacy than function as independent potential cognitive determinants of reading skill. (2) Correlation analyses have not yet led to desirable correlations between cognitive determinants and reading level in such a way that they imply an exhaustive explanatory model of reading skill. Skills close to the reading process (e.g., spelling, pseudoword reading, letter identification) correlate more strongly with reading level than with other cognitive skills (e.g., IQ, memory, attention). Still, correlation coefficients between cognitive skills and reading level rarely exceed $r = 0.50$ (see **Table 1**), which corresponds to a proportion of explained variance of 0.25; thus 75% of variance remains unexplained. (3) Indices of heterogeneity between studies, when specified and not biased by small samples of studies, indicate a medium to high degree of heterogeneity between studies included in these meta-analyses (see Swanson et al., 2003; Melby-Lervåg, 2012; based on Huedo-Medina et al., 2006). High heterogeneity among the selected studies indicates that fixed effects models, applied when interpreting unique effect sizes, are unsuitable for a comparison of results, because effect sizes tend to differ between studies and as such, results of the studies included in these meta-analyses are not unequivocal.

Several authors studied discrepancies in cognitive skills of children with dyslexia and typically developing children. The meta-analyses displayed in rows 5–7 of **Table 1** indicate that children with dyslexia have an increased chance of lower abilities on several cognitive skills compared to typically developing children. Effect sizes (converted to correlation measures in **Table 1**) suggest that cognitive skills of children with dyslexia tend to differ from cognitive skills in typically developing children, although indices of heterogeneity (when specified and not substantially biased) again indicate that results were not unequivocal (Araújo and Faisca, 2019; Lonergan et al., 2019; Parrila et al., 2019). Meta-analyses that distinguished age-matched controls from reading-matched controls, however, reveal that cognitive skills of children with dyslexia do not significantly differ from those of reading-matched controls (e.g., Araújo and Faisca, 2019; Parrila et al., 2019). This indicates that at least some of these lower abilities might have resulted from lower reading level by origin, that is, implying circular causality. On that note, it is suggested that children with dyslexia are identified by decoding skills (i.e., translating printed words in speech) in itself rather than by cognitive skills.

TABLE 1 | Results of meta-analyses on the relation between reading level and cognitive skills, and reading progress and cognitive skills.

Article	Measure	Study aim	Sample characteristics ^a	Cognitive skill ^b	k	Effect Size ^{c,d}	ρ^2 e,f
Hammill and McNutt, 1981	Reading level and progress	Concurrent and longitudinal (after > 1 year) correlation with word recognition, composite reading or reading comprehension ^g	Predominantly mixed samples; few learning disabled/disabled reader; aged Kindergarten 2, Grade 3-12, and other. Language: n.s.; Orthography: n.s.	General academic achievement	14 / 14	0.85 / 0.74	n.s.
				Phonics knowledge	4 / n.s.	0.71 / n.s.	n.s.
				Intelligence	100 / 25	0.61 / 0.52	n.s.
				Readiness	20 / 19	0.56 / 0.50	n.s.
				Spoken Language	8 / 3	0.51 / 0.48	n.s.
				Perceptive abilities	14 / 4	n.s. / n.s.	n.s.
				Motor generalizations	19 / 5	n.s. / n.s.	n.s.
				Affect (emotional functioning)	30 / 8	n.s. / .25	n.s.
Swanson et al., 2003	Reading level	Concurrent correlation with word reading (real and pseudowords)	Average and poor readers; Language: English, Dutch Orthography: n.s.	Spelling	6	0.78	0.00
				Pseudoword reading	24	0.69	79.02
				Phonological awareness	194	0.43	74.03
				Rapid naming	107	0.42	68.46
				Vocabulary	37	0.38	58.16
				Orthography	61	0.41	72.23
				IQ	35	0.42	55.99
				Memory span	46	0.37	25.26
Melby-Lervåg, 2012	Reading level	Concurrent correlation with text decoding	Average school-aged; Language: n.s.; Orthography: n.s.	Phoneme awareness	7	0.56	67.64
				Rhyme awareness	7	0.41	67.01
				Verbal short-term memory	7	0.28	51.21
Parrila et al., 2019	Reading level	Comparison of children with and without dyslexia (reading level-matched [RL] and chronological age-matched [CA]) ^{h,i}	Children with dyslexia and reading level controls (< 13 years) Language: Finnish, Greek, Spanish, Hungarian, Icelandic, Italian, Swedish, Norwegian, German; Orthography: highly consistent European alphabetic	Phonological awareness	19 / 19	-0.21 / -0.49	87.31 / 83.52
				Nonword reading	12 / 10	-0.05 / -0.76	89.71 / 91.94
				Rapid naming	14 / 14	-0.01 / -0.57	74.61 / 80.43
				Verbal short term memory	12 / 12	0.09 / -0.58	85.72 / 86.69
				Auditory temporal processing	5 / 5	-0.24 / -0.21	85.56 / 76.88
Parrila et al., 2019	Reading level	Comparison of children with and without dyslexia ^j	Dyslexic readers; age-matched controls; (age 7-37) Language: English, French, Italian, Finnish, Spanish, Dutch, German, European-Portuguese, Chinese Orthography: "opaque," "transparent," "medium"	Rapid automatized naming:			
				- accuracy	21	0.23	9.91
				- fluency	216	0.50	78.53
				- letters	67	0.50	86.81
				- numbers	127	0.54	83.24
				- objects	93	0.54	75.12
				- colors	56	0.32	64.39
Araújo and Faisca, 2019	Reading level	Comparison of children with and without dyslexia ^j	Children with and without dyslexia; (age 5-18) Language: Dutch, English, German, Norwegian, Italian, Greek, French, Brazilian, Portuguese, Polish, Algerian, Chinese Orthography: n.s.	Inhibition			
				- reaction time	10	0.29	0.00
				- error rate	5	0.24	0.00
				- accuracy	2	0.48	0.00
				Switching attention	3	0.33	0.00
				- reaction time	6	0.41	0.00
				- error rate	2	0.72	0.00
				- accuracy	14	0.48	0.00
				Auditory working memory			
				- accuracy			

(Continued)

TABLE 1 | (Continued)

Article	Measure	Study aim	Sample characteristics ^a	Cognitive skill ^b	<i>k</i>	Effect Size ^{c,d}	<i>f</i> ^{2 e,f}
Scarborough, 1998	Reading progress	Correlation with future (after 1, 2, or 3 years of instruction) word reading, composite reading score or rarely reading comprehension	Unselected samples, few high risk samples; Language: n.s.; Orthography: n.s.	Concepts of print	7	0.46	0.00
				Letter-sound and reading skills	22	0.56	0.00
				Letter identification	24	0.52	0.00
				Phonological awareness	27	0.46	0.00
				Speech discrimination	11	0.22	0.00
				Speech production	4	n.s.	0.00
				IQ full scale	11	0.41	0.00
				Verbal IQ	12	0.37	0.00
				Performance IQ	8	0.26	0.00
				Receptive vocabulary	20	0.33	0.00
				Expressive vocabulary	5	0.45	0.00
				Rapid naming:	9	0.37	0.00
				- Colors, objects	8	0.41	n.s.
				- Digits, letters	9	≤0.37	0.00
				Receptive language skills	11	0.24	0.00
				- Syntax/morphology	11	0.32	0.00
				- Semantic/unspecified	18	0.33	0.00
				Expressive language skills	11	0.45	0.00
				Verbal memory	5	0.22	0.00
				- Words, digits	6	0.16	0.00
				- Story, sentences	8	0.31	0.00
				Visual perception	5	0.25	
				Visual-motor integration			
				Visual memory			
				(Motor skills)			

(Continued)

TABLE 1 | (Continued)

Nelson et al., 2003	Reading progress	Treatment effectiveness of early literacy interventions	Students at risk for reading disabilities; Preschool-3rd Grade; Language: German and n.s. Orthography: n.s.	Rapid naming	7	0.47	0.00
				(Problem behavior)	6	0.43	0.00
				Phonological awareness	17	0.40	0.00
				- phonemic	13	0.35	0.00
				- rhyming	4	0.49	0.00
				Alphabetic principle	18	0.34	0.00
				Memory	11	0.30	0.00
				- short term	8	0.29	0.00
				- long term	3	0.32	0.00
				IQ	8	0.25	0.00
				(Demographic)	5	0.07	0.00
				- (disability / retention)	3	0.10	0.00
				- (ethnicity)	1	0.10	0.00
				- (grade)	1	-0.24	0.00
Tran et al., 2011	Reading progress	Comparing reading level of responders and low-responders during interventions in reading: pretest and posttest ^k	Children at risk for reading disabilities; Language/orthography: n.s. Orthography: n.s.	General IQ	2/9	0.36 / 0.11	0.00 / 91.56
				Verbal IQ	1/3	0.34 / 1.07	n.s. / 79.25
				Real-word identification	21/11	1.06 / 1.53	77.99 / 79.40
				Rapid naming speed	13/15	1.31 / 0.74	82.02 / 0.00
				Phonological awareness	28/9	1.15 / 0.82	85.25 / 69.21
				Pseudo word reading	19/11	1.10 / 1.28	89.24 / 88.59
				Vocabulary	4/8	0.71 / 1.19	0.00 / 93.83
				Reading comprehension	8/18	0.45 / 1.43	70.56 / 72.56
				(Spelling)	1/1	1.85 / 0.79	n.s. / n.s.
				Phonological memory	2/8	0.41 / 0.92	0.00 / 64.77
				(Behavior)	6/6	0.15 / -0.51	81.82 / 72.97
				(Reading fluency)	2/2	0.70 / 0.66	95.03 / 94.99
				(General reading achievement)	1/2	-0.32 / -0.50	100 / 0.00

n.s. = not specified.

^a Information about age/Grades, dyslexia diagnosis, and language or orthography (when specified).

^b Skills that were reported by the articles, but cannot be considered cognitive skills are placed between brackets.

^c Effect sizes in the meta-analyses were reported using d , g , and r . In order to ease comparison of effect sizes between meta-analyses we inferred r from the information provided by the authors. When d measures were reported as effect size metric, r was inferred using the formula: $r = \frac{d}{\sqrt{d^2 + a}}$; when g measures were reported as effect size metric, r was inferred using the formula: $r = \frac{g}{\sqrt{g^2 + a}}$, where a is a correction factor that depends on the ratio of the sample sizes. Tran et al. (2011) provided only d as effect size and provided insufficient information to infer r . Therefore, d measures are displayed in the effect size column for the meta-analysis of Tran et al. (2011).

^d All meta-analyses provided mean weighted effect sizes, except Hammill and McNutt (1981), who reported median correlation coefficients, only when correlations were significant.

^e When Q measures were reported as homogeneity metric, I^2 was inferred using the formula: $I^2 = \frac{Q - (k-1)}{Q} \times 100\%$ for $Q > (k-1)$, $I^2 = 0$ for $Q \leq (k-1)$, where k is the number of studies. I^2 is used to quantify heterogeneity among the studies included in a meta-analysis, and is defined as "a percentage of heterogeneity, that is, the part of total variation that is due to between-studies variance τ^2 ." (Huedo-Medina et al., 2006, p. 197). Homogeneity statistics could not be derived for results of Hammill and McNutt (1981), because they reported median correlation coefficients as effect size metric.

^f Please note that I^2 is biased when the number of studies is small. Interpretation of I^2 is problematic when $k < 20$ see Huedo-Medina et al. (2006).

^g results on concurrent concurrent reading / results on longitudinal reading.

^h for comparison with RL group / for comparison with CA group.

ⁱ positive correlations indicate highest group means for the dyslexia group.

^j positive correlations indicate highest group means for the (non-dyslexic) control group.

^k for pretest reading level / for posttest reading level.

Thus, several cognitive skills tend to correlate with reading level and children with dyslexia tend to differ from age-matched controls in their cognitive skills. Considering the relationship between cognitive skills and reading level, it can be concluded that the relationships are evident, although the magnitudes of these relationships are limited.

Investigating the relationship between reading *progress* and cognitive skills is less straightforward than the relation with reading level. As discussed in the previous paragraph, correlations between reading level and cognitive skills indicate possible determinants or causal factors of reading progress: In order to be predictive of future reading progress, a factor should at least correlate to some extent with present reading level (e.g., Nelson et al., 2003; Tran et al., 2011). However, even when factors correlate with reading level, one cannot exclude the possibility of the effects of a third variable, influencing both cognitive skills and reading level, and thereby generating a correlation between them. Also, correlations never express directions of relationships between variables: The relationships could be opposite to what was expected (instead of cognitive skills causing variations in reading level, reading level might be causing variations in cognitive skills), or being bidirectional (cognitive skills and reading level might mutually influence each other over time). Bidirectional relationships between variables form a considerable possibility in research on reading skill, as is evidenced by the relationship between word reading and phonological awareness (see Castles and Coltheart, 2004). According to Catts (2017), such bidirectional relationships could inflate the correlations that were found over time. Catts probably meant to point out that a correlation measure is not suitable to indicate the strength of a causal relationship: One obvious reason is that factors with reciprocal relationships will produce strong correlations over time when they keep mutually influencing each other all the time. Thus, correlations between reading level and cognitive skills do not necessarily indicate that (lacking) cognitive skills cause lower reading level. To predict reading development, a factor should at least be associated with gains in reading skill over time. In other words, in order to prove that a factor affects reading development, its causal role should be observed in an experiment. In the field of reading development, possibilities are limited, because participants with dyslexia and typical reading development cannot be randomly assigned to groups. Therefore, only quasi-experimental designs can meet this demand. Several authors discuss possibilities for quasi-experimental designs, resulting in roughly three suggestions, listed by ascending validity in proving causal relationships:

- (1) Models that explain variance in growth in reading/spelling over time by cognitive skills at baseline (unconditional models, see Stuebing et al., 2015);
- (2) Models that explain outcome of reading/spelling skills by baseline reading/spelling skills as well as baseline cognitive skills (conditional models, see Stuebing et al., 2015);
- (3) Models that explain progress in reading/spelling skills from progress in cognitive skills (see Vellutino et al., 2004).

Unconditional models fit theoretical research questions about correlations between baseline cognitive skills and reading/spelling development, but do not aim at explaining any causal relationships. If these correlations prove to be compelling, one could ask whether baseline cognitive skills could add to baseline reading/spelling skills during an intervention, which would fit the conditional model. Vellutino et al. (2004), however, argue that possible causal relationships can only be inferred from a model that explains progress in reading skill from progress in cognitive skills. This would be the only opportunity to approximate the demand of observing the causal trajectory in an (quasi-)experiment.

Research results on the relationship between reading progress and cognitive skills are less common and less clear than research results on the relationship with reading level. Overall, studies using unconditional models tend to identify more baseline cognitive skills related to outcome reading level measures and stronger relationships than studies using conditional models (e.g., Stuebing et al., 2015). Although some studies indicate that a number of baseline cognitive skills are related to outcome reading-level measures, it is questionable to what extent these cognitive skills are truly separable from reading skill itself. Meta-analyses on the relation between baseline cognitive skills and progress in reading are presented in the first row and the last three rows of **Table 1**. The results of the meta-analyses presented in **Table 1** originated from the results on unconditional models, albeit Tran et al. (2011) also presented results on conditional models. These meta-analyses reveal little evidence for a relationship between baseline cognitive skills and progress in reading skill. Factors that were most strongly related to results of early literacy training overlapped with reading skill (e.g., general achievement, word reading, pseudoword reading, and reading comprehension), whereas other cognitive skills (e.g., IQ, memory, rapid naming speed, and phonological awareness) produced less strong results (see Hammill and McNutt, 1981; Scarborough, 1998; Nelson et al., 2003; Tran et al., 2011).

Again, when specified, indices of heterogeneity between studies included in the meta-analyses indicated that results were not unequivocal. Findings from empirical research provide suggestions for causes for the varying results on the relationship between reading (level and progress) and cognitive skills. Studies tend to differ in sample characteristics, measurement instruments, methodological approach, and study design (e.g., see Vellutino et al., 2004; Stuebing et al., 2015). As such, comparing studies on the role of cognitive skills in reading development seems like comparing apples to oranges. Thus, research on the relationship between reading progress and cognitive skills is limited in quantity and methodological strength. Moreover, little evidence is found for a relationship between cognitive skills and progress in reading skill and results tend to vary between studies.

Analyses based on traditional linear techniques such as correlations and variances, as presented in the previous paragraphs, may not be applicable to a multifactorial and multidirectional system such as reading development. As Parrila et al. (2019) state, no single factor alone can be accountable for development of a skill as complex as reading, especially

not regarding development over time. Even traditional multi-factorial approaches do not have the capacity to encompass all possible factors and relationships (e.g., Connor and Morrison, 2017). Catts (2017) proposed that dyslexia could be the outcome of multiple interacting factors, moderated by several positive and negative influences. Dyslexia, then, is understood as a condition that follows the risk-resilience framework, in which specific factors serve as moderators of risk, determining different outcomes in individuals with similar precursors (Catts, 2017).

Research based on traditional linear techniques has not been able to clearly identify the role of cognitive skills in reading problems, presumably because the process of reading and the factors that are associated with reading reside within a system of multiple interacting and moderating factors that cannot be captured within traditional statistical models. In contrast, the view of a broad set of interacting variables fits the model of complex adaptive systems: “(...) systems that have a large number of components, (...) that interact and adapt or learn” (Holland, 2006, p. 1). The most important characteristics of complex adaptive systems are:

- (1) Components act simultaneously;
- (2) Components only act upon rules (specific conditions e.g., actions of other components, environmental circumstances);
- (3) Within a component, several rules can combine into specific sequences of rules to, for example deal with novel situations;
- (4) Components can adapt over time, that is, they change their actions in order to abide to the rules and the sequences of rules. Usually these changes are not random, but designed to improve the outcome based on prior experiences (Holland, 2006).

If lags in cognitive skills play a critical role in the emergence and persistence of reading problems, analyses based on complex adaptive systems should be able to at least identify which cognitive skills are involved, and possibly also how and to what extent they are involved. A relative novel way of investigating complex systems is the use of machine learning, in which a set of data together with a set of algorithms seek to find the best solution given the data. Machine learning roughly falls apart in two types of learning: unsupervised and supervised. Unsupervised learning is used to find patterns in the input data without using any output data, particularly to find clusters or dimensions. In supervised learning, the model is confronted with input and output data in order to find the best function between them. Supervised learning is mostly used for predicting future events (Russell and Norvig, 2010).

Research applying the technique of machine learning to the field of reading skill is scarce. To our knowledge only two studies have been published, both applying the unsupervised learning technique of Self Organizing Maps (SOM; Loizou and Laouris, 2010; Astle et al., 2018). Loizou and Laouris (2010) used the SOM technique to make different clusters of participants based on measures of cognitive skills and word reading skill. Subsequently, they identified which (sub)tests made the strongest contribution

to the assignment of participants to clusters. Their results showed that information of only four tests (auditory memory, navigation, word identification and word attack, and rapid naming of pictures) were sufficient to classify 94.64% of the participants in the identified clusters. Also, the fifth strongest factor contributing to the classification was age. Astle et al. (2018) adopted a slightly different approach. They used the SOM technique to distinguish between different clusters of participants based only on measures of cognitive skills (“cognitive profiles,” based on seven measures: nonverbal reasoning, vocabulary, phonological processing, and four measures of short-term memory) and afterwards compared these clusters to initial referral routes and diagnoses. Their results showed that participants allocated to the cluster “lag in a broad spectrum of cognitive skills,” showed the most severe problems in reading, spelling, and math skills. More importantly, learning-problems characteristics and how participants were diagnosed were not related to the cognitive profiles determined by the SOM technique. Thus, some efforts were made studying the role of cognitive skills in reading skill, revealing some preliminary results: Only a small number of cognitive and reading skills were needed to make up some clusters, albeit when clusters were based on cognitive skills only, they were not related to learning skills.

Although both studies on the role of cognitive skill in reading development did consider the multifactorial nature of reading skill, neither was capable of identifying cognitive factors predictive of reading skill. These studies used the unsupervised learning SOM technique, which aims at classification, and both input (cognitive) variables and output (reading level) variables are from the same moment of measurement. As such, participants are labeled according to their reading level, and not according to their reading progress. To come back to the relevance of the role of cognitive skills in reading problems, this mainly serves the theoretical perspective about the nature of reading and processes that accompany the development of automated reading skill. Questions such as which cognitive skills are predictive of reading difficulties (prevention perspective) and which cognitive factors can affect the development of automated reading skill (remediation perspective) are left unanswered.

To answer these questions, supervised machine learning techniques should be applied, because these techniques are capable of making predictions about future events (Russell and Norvig, 2010; Lantz, 2019). More specifically, machine learning based on neural networks seems the most likely technique to identify cognitive factors predictive of reading skill. According to Lantz (2019), neural networks belong to so-called black-box methods and these are, more than any other technique, capable of modeling complex patterns. Moreover, neural networks pose few assumptions on the input data.

The aim of the present study was to find cognitive factors of varying origin that play a role in the development of reading skill in children with dyslexia and children with typical reading development. The present study will address the following research questions:

1. To what extent are cognitive skills indicative of present decoding level?

2. To what extent are cognitive skills predictive of future decoding progress?

The present study is unique in applying supervised machine learning techniques to the field of reading development. Moreover, the present study includes analyses on a relatively homogeneous group of reading disabled children as well as a heterogeneous group representing a sample of the population of school age children to correct for the effects of restriction of range.

GENERAL METHODS

Overview

Three datasets will be subjected to machine-learning modeling and reported on in three studies. Study 1 was performed on the dataset of Verhoeven and Keuning (2018) on cognitive precursors of reading and reading level of 2007 Dutch primary-school children. For Study 2, data of a previous study (Walda et al., 2014) were used and supplemented with new cases (i.e., more participants with dyslexia) and new cognitive skills variables, that is, in the Walda et al. study a number of specific executive functions were investigated, whereas the current study 2 considers cognitive skills in a more general way (noem hier wat voorbeelden van specifiek en algemeen). The dataset of the present study contained data on cognitive skills, reading and spelling in 383 Dutch children with dyslexia. The data consist of results on cognitive precursors of reading and reading level prior to reading and spelling remediation as well as reading level after three months of reading and spelling remediation in a Dutch clinic for the assessment and remediation of learning disorders. Study 3 was conducted on the dataset of Braams and Bosman (2000) pertaining to kindergarten predictors of reading and spelling level of two cohorts (117 and 82, respectively) Dutch primary-school children in Grade 1. Study 1 was aimed at answering research question 1. Studies 2 and 3 also aimed at answering research question 1 (Study 2a and Study 3a) as well as research question 2 (Study 2b and Study 3b). For a schematic overview of the particular characteristics of the studies, see **Figure 1**.

Although the original datasets contained more cases, only results on complete cases are presented in the present study: Cases with a missing record on one or more variables (word decoding or one of the cognitive skills variables described in **Supplementary Material B**) were excluded from the analyses because they do not contribute to the building of a model. Concerning research question 1 (word decoding level) in Studies 1, 2a, and 3a, input data of cognitive skills on T1 were modeled on output data of word decoding skill at T1. Concerning research question 2 (word decoding progress) in Studies 2b and 3b, input data of cognitive skills on T1 were modeled on output data of word decoding level at T2, corresponding to what was considered an unconditional model in the "Introduction" section.

Data Analysis

The models were trained in R (CRAN version). The data of Studies 1 and 2 were normalized to a zero to one range. Those of Study 3 had to be normalized using Z-scores to be able to compare between participants. To maximize the opportunity of selecting a model with a good fit, each dataset was subjected to four different techniques of supervised machine learning. Two techniques building single predictive models were used, namely, neural networks and k-nearest neighbors and two ensemble techniques, namely, random forests and extreme gradient boosting (Xg-boost). In addition, a model using General Linear Model (GLM) technique was used to compare the results of all four nonlinear techniques to a linear one. For a description of the Xg-boost technique, see Lesmeister (2019); for a description of the other four techniques see Lantz (2019).

Model Building

Each model was built to result in a topology with several input factors measuring cognitive skills, which predicted one outcome measure, that is, word-decoding level. To compare the fit of most models to the data, outcome measures were treated binomially, distinguishing between participants with the 20% lowest word-decoding raw scores and participants with the 80% best word-decoding raw scores in Studies 1 and 3. In Study 2, standardized scores (i.e., c-scores) were used for the outcome measure, resulting in binomial scores distinguishing between c-scores 0 and 1 and c-scores 2–9, which corresponds to approximately 10% lowest word-decoding efficiency and 90% highest scoring in the regular population.

Parameters of the four models were tuned until they approximated the model that best fit the data in the dataset (for more information about building these models see Lantz, 2019; Lesmeister, 2019). Which parameters can be tuned depends on the specific machine learning technique, and is specified in **Supplementary Material A**. In Studies 1 and 2, 90% of the datasets were used for tuning and training and 10% for testing the model. Because of the relatively small sample of data in Study 3, 75% of the dataset was used for tuning and training and 25% for testing the model. To overcome the problem of differential findings as a result of the seed set, a number of models were built for each technique and results were run on 100 different seeds. Results are reported with a 95% confidence interval. For more details about building, tuning, training, and testing of the models see **Supplementary Material A**.

Model Evaluation

Subsequently, model performance was evaluated computing summary statistics for the predictive ability of the model and visualizing performance tradeoff for all five (one linear and four nonlinear, see **Figure 1**) models. We used R's `set.seed` function (R Core Team, 2018) to generate random initializations of R's internal Random Number Generator, which was set to the "Mersenne Twister" algorithm (cf. Matsumoto and Nishimura, 1998). To make the analyses based on random sampling exactly reproducible, we stored the random seeds as variables (for details, see the analysis scripts in the **Supplementary Material**).

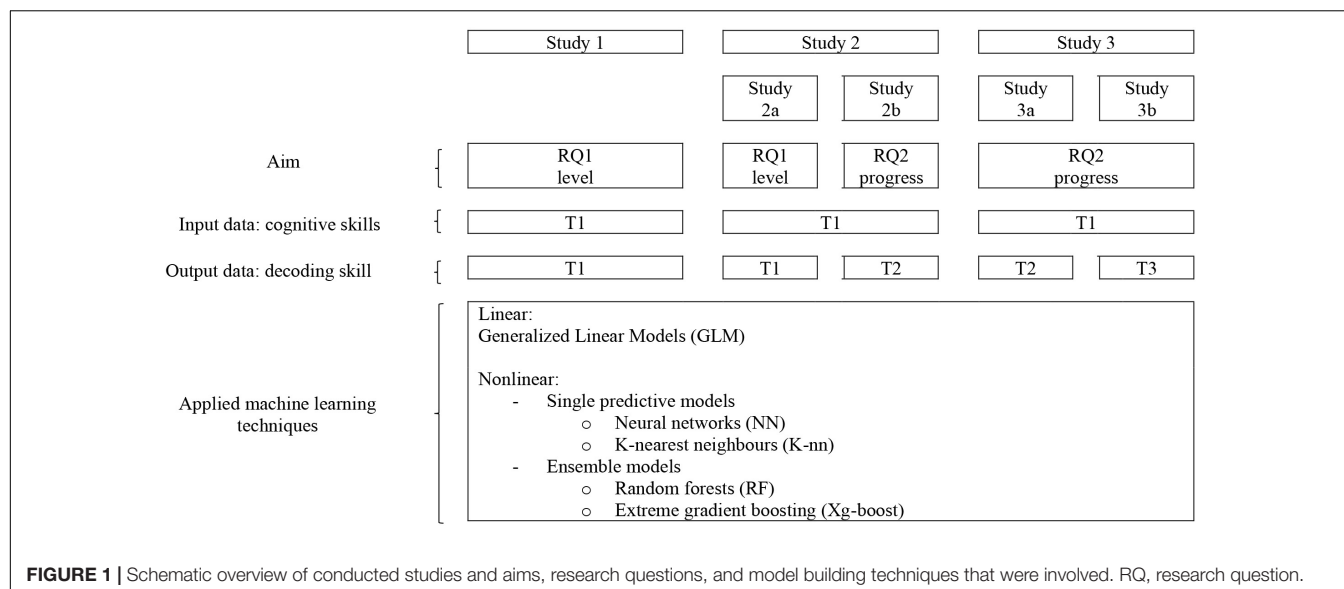


TABLE 2 | Confusion matrix for positive class and negative class allocation by the models.

Group membership according to model	Group membership according to word decoding test		Performance statistic (usefulness)
	Lowest decoding	Not lowest decoding	
Lowest decoding	TP	FP	PP
Not lowest decoding	FN	TN	NP
Performance statistic (identification)	SE	SP	Accuracy

TP, True positives; FP, False positives; FN, False negatives; TN, True negatives; PP, Positive predictive value; NP, Negative predictive value; SE, Sensitivity; SP, Specificity.

TABLE 3 | Descriptive statistics for input variables and output variable of the models ($n = 2009$).

	Range	M	SD
Input variables			
Nonword repetition	11–40	33.72	4.77
Naming speed – digits	43–166	96.25	17.09
Naming speed – letters	1–157	91.79	19.13
Naming speed – pictures	5–129	61.82	11.93
Phoneme segmentation	0–20	18.55	3.07
Phoneme manipulation	0–20	18.55	2.52
Output variable			
Word decoding efficiency	0.5–120.75	57.47	20.88

The predictions vectors were set to contain probabilities. Summary metrics were computed based on average performance using 10-fold cross validation.

Performance tradeoff was evaluated using metrics based on the confusion matrix of model predictions and actual class membership according to the decoding test. **Table 2** presents the confusion matrix and metrics that were considered. Both metrics indicating usefulness of the model for assessment of low decoding skill (i.e., positive predictive value [PP] and negative predictive value [NP]), and accurate identification of children with low decoding skill by the model (i.e., accuracy, sensitivity

[SE], specificity [SP]) were considered. Usefulness of the model for assessment of low decoding skill concerns the likelihood that the constructed model can successfully identify whether children indeed have the lowest decoding skill or not, as illustrated by the last column of **Table 2**. Accurate identification of children with low decoding skill by the model concerns the similarity of the allocation of children to the lowest decoding group by the model to actual lowest decoding skill according to the word decoding test, as illustrated by the last row of **Table 2** (Trevethan, 2017).

Positive Predictive Value

The proportion of positive cases that were accurately allocated to the target category (lowest decoding level group) by the model, that is, when the model allocates members to the low decoding level group, how many belong to this group according to the word decoding test? $PP = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$

Negative Predictive Value

The proportion of negative cases that were accurately allocated to the non-target category (not lowest decoding level group), that is, when the model allocates members to the not-lowest decoding level group, how many belong to this group according to the word decoding test? $NP = \frac{\text{true negatives}}{\text{true negatives} + \text{false negatives}}$

Sensitivity

The proportion of cases that were accurately allocated to the target category (lowest decoding level group) by the model, that

is, of the children who belong to the lowest decoding level group according to the test, how many were allocated to the lowest decoding group by the model? $SE = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$

Specificity

The proportion of cases that were accurately allocated to the non-target category (not lowest decoding level group), that is, of the children who do not belong to the lowest decoding level group according to the test, how many were not allocated to the lowest decoding group by the model? $SP = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$

Accuracy

The proportion of cases that were accurately allocated to the non-target and to the target group.

$$\text{Accuracy} = \frac{\text{true positives} + \text{true negatives}}{\text{true positives} + \text{true negatives} + \text{false positives} + \text{false negatives}}$$

Cohen's Kappa (κ)

Statistic with values between 0 and 1, indicating the agreement between predictions of the model (decoding level within lowest decoding category or not) and the true membership of decoding level category (according to the word decoding test). The results of Cohen's Kappa were evaluated to account for the possibility of accurate prediction by chance alone, which is an evident risk because of class imbalance in the present study. In the present study, Cohen's κ will be interpreted as suggested by Lantz (2019, p. 324):

- Poor agreement = less than 0.20
- Fair agreement = 0.20 to 0.40
- Moderate agreement = 0.40 to 0.60
- Good agreement = 0.60 to 0.80
- Very good agreement = 0.80 to 1.00

Receiver Operating Characteristic Curve

Visualization of the tradeoff between the sensitivity and the proportion of cases that were falsely allocated to the target category by the model.

Area Under the Curve

Statistics with values varying between 0.5 and 1, with higher values indicating better predictive models. AUC is based on the tradeoff between sensitivity and the proportion of cases that were falsely allocated to the target category by the model. Although some guidelines for classifying AUC are available, AUC is best evaluated in a comparative way. In the present study, AUC will be classified using the convention suggested by Lantz (2019, p. 333):

- No Discrimination = 0.5 to 0.6
- Poor = 0.6 to 0.7
- Acceptable / Fair = 0.7 to 0.8
- Excellent / Good = 0.8 to 0.9
- Outstanding = 0.9 to 1.0

Due to the fact that the binary classes were not evenly distributed (80% / 20%), some distortion of the metrics were expected. Specifically, the uneven distribution in favor of the non-target category leads to a higher chance of correct allocation to the non-target category. Therefore, results were interpreted primarily using Cohen's Kappa, ROC curve, and AUC.

STUDY 1

In this study, we built a model from an existing dataset on phonological abilities and word-decoding accuracy in Dutch children. The present database originates from research conducted by Verhoeven (Expertisecentrum Nederlands) and Keuning (Cito) on precursors of dyslexia in Dutch children (see Verhoeven and Keuning, 2018). Study 1 aimed at answering the question: To what extent are cognitive skills indicative of present decoding level? Baseline cognitive skills in Study 1 consisted of phonological awareness skills, rapid naming skills, and working memory skills (for a detailed description see **Supplementary Material B**).

Materials and Methods

Participants

Participants were Dutch children attending 68 elementary schools: 782 in Grade 3, 707 in Grade 4, 263 in Grade 5, and 255 in Grade 6 after deletion of cases with missing variables.

Materials

Measures of one output variable and six input variables were collected using assessments. The output variable consisted of a test score for word-decoding efficiency and the input variables of test scores for nonword repetition, naming speed, phoneme segmentation, and phoneme deletion. A detailed description of the tests that were used can be found in the **Supplementary Material B**.

Procedure

Assessment of input and output variables took place in an individual setting by trained graduate students. The sequence of the tests within a session was randomized. All students were assessed halfway the school year. For more details about the methods, see Verhoeven and Keuning (2018).

Results

Means and standard deviations for all input variables and the output variable of the models are presented in **Table 3**. The output variable (word-decoding efficiency) was transformed into binary classes, with 438 cases in the 20% lowest decoding level class and 1571 cases in the alternative class.

Evaluation results of the models built with five machine learning techniques are presented in **Table 4**. Concerning identification of low decoding skill by the models, Positive Predictive (PP) and Negative Predictive (NP) values were evaluated. The results on the 95% confidence-intervals of PP appear between 0.25 and 0.37. Thus, between 25 and 37% of children allocated to the low decoding-category by the model, truly performed within the low decoding category when assessed with a decoding test; 63 to 75% of children allocated to the low decoding-category did not. The results on the 95% confidence-intervals of NP revealed that 92 to 96% of children not allocated to the low decoding-category by the model, truly did not perform within the low decoding category when assessed with a decoding test; 4 to 8% of children not allocated to the low decoding-category did.

TABLE 4 | Confidence intervals of summary statistics for the predictive ability of the models built with five machine learning techniques.

Technique	PP 95% CI	NP 95% CI	SE 95% CI	SP 95% CI	Acc 95% CI	κ 95% CI	AUC 95% CI
Neural network	[0.32, 0.34]	[0.94, 0.95]	[0.61, 0.65]	[0.83, 0.84]	[0.81, 0.82]	[0.29, 0.34]	[0.77, 0.79]
K-nn	[0.27, 0.30]	[0.94, 0.95]	[0.58, 0.62]	[0.83, 0.83]	[0.80, 0.81]	[0.27, 0.30]	[0.73, 0.75]
Random Forests	[0.32, 0.35]	[0.93, 0.94]	[0.57, 0.61]	[0.83, 0.84]	[0.80, 0.81]	[0.30, 0.33]	[0.78, 0.79]
Xg-boost	[0.34, 0.37]	[0.92, 0.93]	[0.55, 0.59]	[0.84, 0.84]	[0.80, 0.81]	[0.31, 0.34]	[0.77, 0.79]
GLM	[0.25, 0.28]	[0.96, 0.96]	[0.64, 0.68]	[0.82, 0.83]	[0.81, 0.81]	[0.27, 0.30]	[0.77, 0.79]

PP, positive predictive value; NP, negative predictive value; SE, sensitivity; SP, specificity; Acc, accuracy; κ , Kappa; AUC, area under the ROC; CI, confidence interval.

Concerning the usefulness of the models of detecting low decoding skill, accuracy, sensitivity, and specificity were evaluated. The results of the 95% confidence-intervals of accuracy revealed that between 80 and 82% of the children were correctly allocated to the right decoding skill category by the models. The results on the 95% confidence-intervals of sensitivity showed that between 55 and 68% of children who truly performed within the low decoding category when assessed with a decoding test, were indeed allocated to the low decoding group by the model; 45% to 32% of children with actual low decoding skills were not. The results on the 95% confidence-intervals of specificity revealed that between 82 and 84% of children who truly performed not within the low decoding-category when assessed with a decoding test, were indeed allocated to the not low decoding group by the model; 16 to 18% of children without actual low decoding skill were falsely allocated to the low decoding group by the model. The results of Cohen’s κ were evaluated to account for the possibility of accurate prediction by chance alone. The results on the 95% confidence-intervals of the Cohen’s κ appear between 0.27 and 0.34, and indicate fair agreement between the models’ predictions and the true values. See “Materials and Methods” section for suggested interpretation of Cohen’s κ .

The charts in the first column of **Figure 2** visualize results of identification of first-quintile word-decoding efficiency for all models. The curves indicate positive predictive abilities of all models. The results on the 95% confidence-intervals of the AUC-statistics appear between 0.73 and 0.79, and indicate acceptable identification of first quintile word decoding for all models. See “Materials and Methods” section for suggested interpretation of AUC statistics. Visual inspection of the curves confirms the similarity between the five models pertaining to results of the AUC statistic.

Conclusion

Study 1 aimed at building a model of five cognitive skills input variables to predict which participants achieved within the 20th percentile of decoding skill of children attending regular Dutch schools. The results indicated that model building by nonlinear machine learning techniques yields results comparable to model building by more traditional linear techniques. The predictive ability of both the linear model and the four nonlinear models appeared to be acceptable, as indicated by the ROC curves and AUC statistics. Balancing the tradeoff between sensitivity and specificity reveals best results in terms of specificity, at the expense of sensitivity,

that is, the models tend to identify about 80% of children without decoding problems correctly (specificity), but only about 60% of children who indeed have low decoding skill (sensitivity), which results in accuracy of about 80%. The results of Cohen’s κ , however, indicate only fair agreement between the expected decoding category and the true decoding category. Recall that accuracy can be overestimated due to class imbalance, which is the case in the present study (see “General Methods” section). The weak agreement according to Cohen’s κ demonstrates that accuracy scores were affected by the imbalanced categories.

STUDY 2

In this study, we built a model from a dataset on performance on cognitive tests and word-decoding accuracy in Dutch children with dyslexia. Two types of model design were evaluated: Study 2a models fourteen cognitive predictors and uses decoding accuracy as an outcome variable at the same moment of measurement, pertaining to the question: To what extent are cognitive skills indicative of present decoding level? Study 2b models fifteen cognitive predictors at T1 and uses decoding accuracy after three months of reading and spelling remediation (T2) as an outcome variable, pertaining to the question: To what extent are cognitive skills predictive of future decoding progress? Baseline cognitive skills in Study 2 consisted of phonological awareness skills, rapid naming skills, working memory skills, nonverbal reasoning skills, and vocabulary skills (for a detailed description, see **Supplementary Material B**).

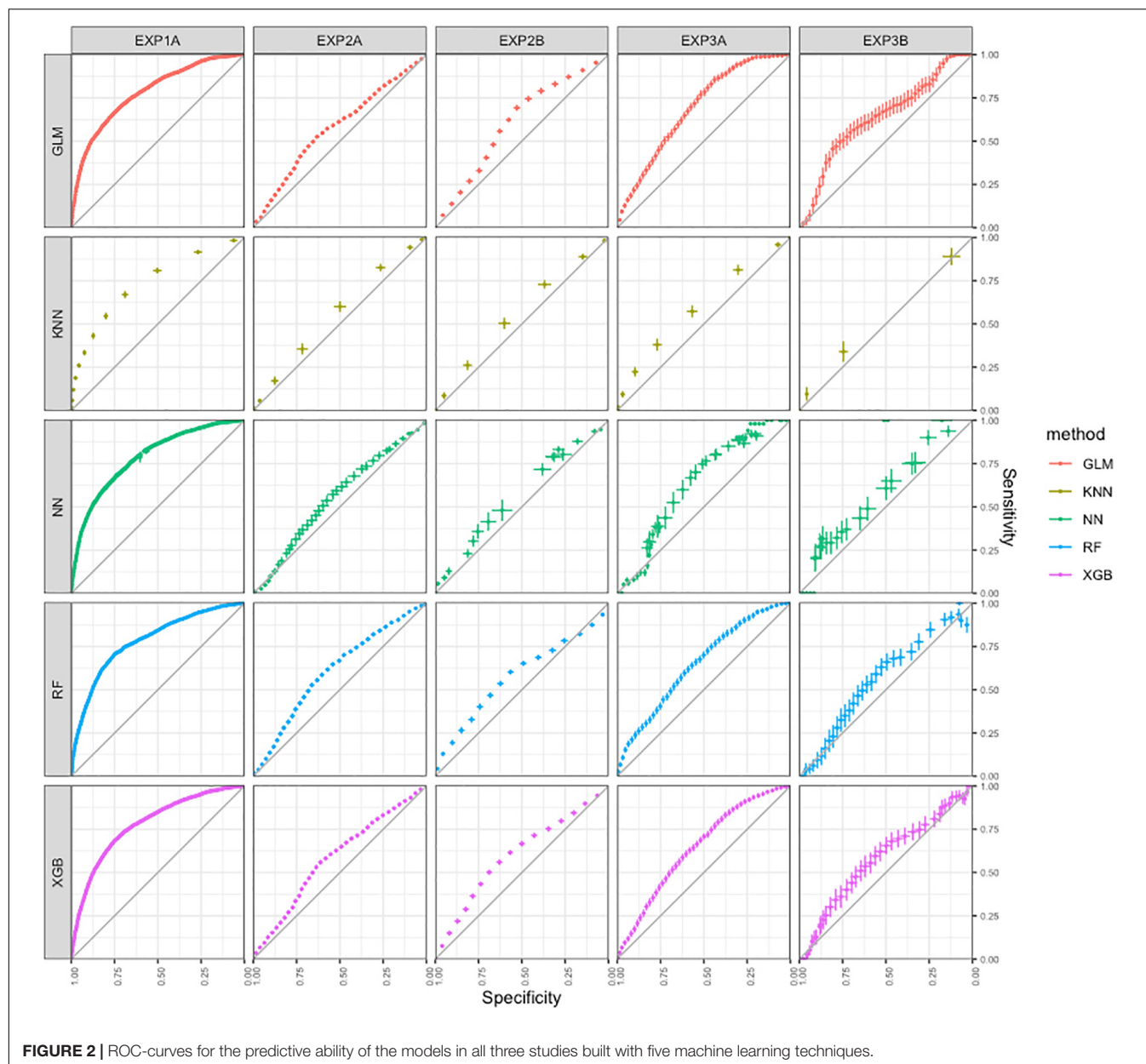
Materials and Methods

Participants

After excluding all cases with missing variables, the data of 383 children attending Braams & Partners (a Dutch clinic for the assessment and remediation of learning disorders) could be used in this study. All participants were attending Dutch primary schools Grade 2–6 and were diagnosed with severe dyslexia according to the criteria of Blomert (2006) and the Stichting Dyslexie Nederland (Dutch Dyslexia Foundation).

Materials

Measures of one output variable and fourteen input variables were collected using the following assessments. The output variable consisted of a test score for word decoding. The input variables consisted of test scores for grapheme-phoneme



identification, grapheme-phoneme discrimination, naming speed, vocabulary, nonverbal reasoning, digit recall, block recall, and word recall. A detailed description of the tests that were used can be found in the **Supplementary Material B**.

Procedure

The tests administered at the clinic assessed whether criteria were met for severe dyslexia and thus for reading and spelling remediation at the clinic. Only children who met the criteria for severe dyslexia were included in the study. The Protocol Dyslexia Diagnosis and Treatment (Blomert, 2006) required the following criteria for a diagnosis of severe singular dyslexia:

- (1) Persistence: The main criterion for referral to a clinic was that students appear not to profit from at least eight weeks of extra reading and spelling remediation in school, which is roughly operationalized in persisting scores below the 10th percentile on reading tests.
- (2) Severity: Upon referral, severity of the reading and spelling impairment was assessed by means of several standardized reading and spelling tests.
- (3) Cognitive profile: Other cognitive skills associated with dyslexia (e.g., phonological processing, rapid naming, verbal working memory) were assessed as well.
- (4) Differential diagnostics: To exclude students whose reading and spelling impairment stem from alternative

TABLE 5 | Descriptive statistics for input variables and output variable of the models.

Variable	Study 2a (n = 384)			Study 2b (n = 183)		
	Range	M	SD	Range	M	SD
Input variables						
Grapheme-phoneme identification						
Speed	1.28–5.60	2.54	0.58	1.28–4.99	2.58	0.57
Accuracy	35.56–100	87.26	8.66	53.33–100.00	86.61	8.19
Grapheme-phoneme discrimination						
Speed	0.83–19.44	1.85	1.04	0.00–3.39	1.80	0.52
Accuracy	17.78–96.67	81.39	10.02	0.98–95.56	80.40	11.74
Naming speed						
Digits	3.54–42.58	10.39	3.03	6.51–42.58	10.57	3.45
Letters	3.78–20.63	11.25	2.53	7.14–19.28	11.32	2.37
Pictures	3.71–31.22	15.66	3.54	9.72–31.22	15.60	3.41
Vocabulary	57–144	110.11	9.96	85–144	110.84	9.26
Nonverbal reasoning	5–119	19.90	6.78	5–119	20.46	9.22
Digit recall						
Forward	6–36	24.15	3.83	11–32	24.25	3.92
Backward	3–39	11.05	3.74	5–30	11.17	3.65
Block recall	2–37	25.58	4.45	2–37	25.76	4.41
Word recall						
Reproduction	16–61	39.86	8.47	16–61	40.21	8.26
Recall	1–15	8.51	2.71	1–15	8.47	2.82
Output variable						
Word decoding efficiency	0–6	1.45	1.29	0–5	1.50	1.49

TABLE 6 | Confidence intervals of summary statistics for the predictive ability of the models built with five machine learning techniques in study 2a.

Technique	PP 95% CI	NP 95% CI	SE 95% CI	SP 95% CI	Acc 95% CI	κ 95% CI	AUC 95% CI
Neural network	[0.33, 0.45]	[0.59, 0.71]	[0.52, 0.58]	[0.51, 0.54]	[0.51, 0.53]	[0.02, 0.06]	[0.55, 0.57]
K-nn	[0.50, 0.54]	[0.53, 0.57]	[0.52, 0.55]	[0.52, 0.55]	[0.52, 0.55]	[0.04, 0.09]	[0.56, 0.58]
Random forests	[0.33, 0.45]	[0.59, 0.71]	[0.52, 0.58]	[0.51, 0.54]	[0.51, 0.53]	[0.02, 0.06]	[0.60, 0.63]
Xg-boost	[0.52, 0.56]	[0.57, 0.61]	[0.55, 0.59]	[0.55, 0.57]	[0.55, 0.58]	[0.10, 0.15]	[0.57, 0.60]
GLM	[0.48, 0.53]	[0.58, 0.62]	[0.54, 0.58]	[0.53, 0.56]	[0.54, 0.57]	[0.07, 0.13]	[0.56, 0.59]

PP, positive predictive value; NP, negative predictive value; SE, sensitivity; SP, specificity; Acc, accuracy; κ , Kappa; AUC, area under the ROC; CI, confidence interval.

TABLE 7 | Confidence intervals of summary statistics for the predictive ability of the models built with five machine learning techniques in study 2b.

Technique	PP 95% CI	NP 95% CI	SE 95% CI	SP 95% CI	Acc 95% CI	κ 95% CI	AUC 95% CI
Neural network	[0.70, 0.79]	[0.20, 0.27]	[0.51, 0.53]	[0.42, 0.53]	[0.49, 0.53]	[−0.05, 0.02]	[0.54, 0.58]
K-nn	[0.59, 0.66]	[0.34, 0.40]	[0.51, 0.54]	[0.44, 0.50]	[0.48, 0.53]	[−0.05, 0.04]	[0.56, 0.60]
Random forests	[0.53, 0.60]	[0.37, 0.45]	[0.50, 0.54]	[0.42, 0.48]	[0.47, 0.51]	[−0.07, 0.02]	[0.56, 0.60]
Xg-boost	[0.48, 0.55]	[0.38, 0.44]	[0.47, 0.52]	[0.40, 0.46]	[0.44, 0.49]	[−0.12, −0.03]	[0.58, 0.62]
GLM	[0.58, 0.64]	[0.34, 0.41]	[0.50, 0.54]	[0.42, 0.49]	[0.47, 0.52]	[−0.07, 0.03]	[0.59, 0.63]

PP, positive predictive value; NP, negative predictive value; SE, sensitivity; SP, specificity; Acc, accuracy; κ , Kappa; AUC, area under the ROC; CI, confidence interval.

“causes,” several additional cognitive skills were assessed (e.g., IQ, nonverbal memory). Children with other diagnoses that can account for language problems, such as SLI and sensory problems, were also excluded.

Once assessment results met all criteria, the student was diagnosed with severe singular dyslexia and was eligible for a subsidized, specialized reading and spelling remediation program

at the clinic. Students who did not satisfy these criteria were provided with specific recommendations for remediation in school or referred to another specialist, depending on assessment results.

Results

Means and standard deviations for all input variables and the output variable of the models are presented in **Table 5**. The output

variable (word decoding efficiency) was transformed into binary classes, with 190 cases in the lowest decoding level class and 194 cases in the alternative class in Study 2a and 99 cases in the lowest decoding level class and 84 cases in the alternative class in Study 2b.

Evaluation results of the models built with five machine learning techniques are presented in **Table 6** for Study 2a and in **Table 7** for Study 2b. Concerning identification of lowest decoding skill by the models, PP and NP were evaluated. The results on the 95% confidence-intervals of PP reveal that between 33 and 56% of children allocated to the lowest decoding category by the model, truly performed within the lowest decoding category when assessed with a decoding test; 54 to 67% of children allocated to the low decoding-category did not. The results on the 95% confidence-intervals of NP reveal that between 53 and 71% of children not allocated to the lowest decoding category by the model, truly did not perform within the lowest decoding category when assessed with a decoding test; 29 to 47% of children not allocated to the lowest decoding category did.

Concerning the usefulness of the models to detect lowest decoding skill, accuracy, sensitivity, and specificity are evaluated. The results of the 95% confidence-intervals of accuracy reveal that between 51 and 58 % of the children were correctly allocated to the right decoding skill category by the models. The results on the 95% confidence-intervals of sensitivity reveal that between 52 and 59% of children who truly performed within the lowest decoding category when assessed with a decoding test, were indeed allocated to the lowest decoding group by the model; 41 to 48% of children with actual lowest decoding skills were not. The results on the 95% confidence-intervals of specificity reveal that between 51 and 56% of children who truly performed not within the lowest decoding-category when assessed with a decoding test, were indeed allocated to the not lowest decoding group by the model; 44 to 49% of children without actual lowest decoding skill were falsely allocated to the lowest decoding group by the model. Cohen's κ -were evaluated to account for the possibility of accurate prediction by chance alone. The results on the 95% confidence-intervals of the κ statistics appear between 0.02 and 0.15, and indicate poor agreement between the models' predictions and the true values. See "Materials and Methods" section for suggested interpretation of Cohen's κ .

The charts in the second column of **Figure 2** visualize results on identification of the lowest decoding level class for all models in Study 2a, and the charts in the third column for Study 2b. For both studies, the curves indicate minimal yet predominantly positive predictive abilities of all models, with little to no differences between models based on different techniques. The results on the 95% confidence-intervals of the AUC-statistics appear between 0.55 and 0.63, and indicate no to poor identification of lowest word decoding for all models. See "Materials and Methods" Section for suggested interpretation of AUC statistics. Visual inspection of the curves confirms the similarity between the five models on results on the AUC statistic.

Conclusion

Study 2 was aimed at building a model of fourteen cognitive skills input variables to predict which participants achieved within the

TABLE 8 | Descriptive statistics for raw scores on input variables and output variable of the model.

Variable	Study 3a (n = 203)			Study 3b (n = 199)		
	range	M	SD	range	M	SD
Input variables						
Rhyme	0–28	17.67	5.11	0–28	17.75	4.96
Rhyme prime	2–26	12.54	5.05	2–26	12.54	5.04
Auditory synthesis	0–24	13.42	4.98	0–24	13.43	4.98
Phoneme deletion	0–29	12.47	7.98	0–29	12.41	7.98
Letter naming	0–20	7.71	6.52	0–20	7.73	6.49
Output variable						
Word decoding efficiency	4–68	16.16	10.39	4–74	27.32	14.16

lowest category of decoding skill within children with severe dyslexia. The results of this study again indicated that results on model building by nonlinear machine learning techniques are comparable to results on model building by more traditional linear techniques. The predictive ability of both the linear model and the four nonlinear models appeared to be poor, as indicated by the ROC curves and AUC statistic. Balancing the tradeoff between sensitivity and specificity indicates that the models tend to identify only about 50% of children without lowest decoding skill correctly (specificity), and only about 50% of children who have indeed lowest decoding skill (sensitivity), which results in a disappointing accuracy of about 50%. Correcting for the imbalanced categories, the results of Cohen's κ leave only poor agreement between the expected decoding category and the true decoding category. Results for prediction of participants achieving within the lowest category of decoding skill at T1 (Study 2a, aimed at word decoding level) were comparable to those at T2 (Study 2b, aimed at word decoding progress).

STUDY 3

In this study, the predictive validity of five phonological-awareness tests on initial reading was investigated in Dutch children attending primary school. Phonological awareness was assessed once: In the last year of kindergarten (T1). Word decoding was assessed twice: Halfway Grade 1 (T2, providing results for Study 3a) and at the end of Grade1 (T3, providing results for Study 3b), both answering to the question: To what extent are cognitive skills predictive of future decoding progress? Baseline cognitive skills in Study 3 consisted of phonological awareness skills (for a detailed description see **Supplementary Material B**).

Materials and Methods

Participants

Participants were Dutch children attending three primary schools. The group that was assessed on phonological awareness halfway the last year of kindergarten (cohort 2) consisted of 121 participants, and the group that was assessed on phonological awareness at the end of the last year of Kindergarten (cohort 1) consisted of 82 participants, after deletion of cases with missing

TABLE 9 | Confidence intervals of summary statistics for the predictive ability of the models built with five machine learning techniques in study 3a.

Technique	PP 95% CI	NP 95% CI	SE 95% CI	SP 95% CI	Acc 95% CI	κ 95% CI	AUC 95% CI
Neural network	[0.10, 0.15]	[0.88, 0.93]	[0.20, 0.30]	[0.80, 0.80]	[0.79, 0.80]	[0.00, 0.01]	[0.64, 0.66]
K-nn	[0.13, 0.17]	[0.91, 0.96]	[0.41, 0.53]	[0.81, 0.82]	[0.78, 0.80]	[0.10, 0.15]	[0.60, 0.63]
Random Forests	[0.20, 0.25]	[0.88, 0.93]	[0.39, 0.46]	[0.82, 0.83]	[0.77, 0.79]	[0.14, 0.20]	[0.63, 0.67]
Xg-boost	[0.17, 0.22]	[0.84, 0.90]	[0.28, 0.35]	[0.81, 0.82]	[0.74, 0.76]	[0.07, 0.13]	[0.63, 0.66]
GLM	[0.08, 0.11]	[0.94, 0.98]	[0.42, 0.57]	[0.81, 0.81]	[0.79, 0.80]	[0.05, 0.10]	[0.67, 0.71]

PP, positive predictive value; NP, negative predictive value; SE, sensitivity; SP, specificity; Acc, accuracy; κ , Kappa; AUC, area under the ROC; CI, confidence interval.

TABLE 10 | Confidence intervals of summary statistics for the predictive ability of the models built with five machine learning techniques in study 3b.

Technique	PP 95% CI	NP 95% CI	SE 95% CI	SP 95% CI	Acc 95% CI	κ 95% CI	AUC 95% CI
Neural network	[0.36, 0.43]	[0.72, 0.80]	[0.40, 0.49]	[0.75, 0.76]	[0.75, 0.76]	[0.00, 0.01]	[0.65, 0.67]
K-nn	[0.20, 0.24]	[0.91, 0.93]	[0.48, 0.58]	[0.78, 0.79]	[0.74, 0.76]	[0.14, 0.19]	[0.63, 0.66]
Random Forests	[0.19, 0.23]	[0.91, 0.93]	[0.46, 0.56]	[0.78, 0.79]	[0.74, 0.76]	[0.14, 0.19]	[0.61, 0.64]
Xg-boost	[0.23, 0.28]	[0.84, 0.87]	[0.35, 0.41]	[0.78, 0.79]	[0.70, 0.72]	[0.10, 0.15]	[0.61, 0.64]
GLM	[0.07, 0.10]	[0.93, 0.95]	[0.28, 0.40]	[0.76, 0.76]	[0.73, 0.74]	[0.01, 0.05]	[0.67, 0.70]

PP, positive predictive value; NP, negative predictive value; SE, sensitivity; SP, specificity; Acc, accuracy; κ , Kappa; AUC, area under the ROC; CI, confidence interval.

variables in Study 3a. In Study 3b, cohort 2 consisted of 82 participants, and cohort 1 consisted of 117 participants, after deletion of cases with missing variables.

Materials

Measures of one output variable and five input variables were collected using assessments. The measures of the input variables were designed by Braams (see Braams and Bosman, 2000). The output variable consisted of a test score for word decoding. The input variables consisted of test scores for rhyme, rhyme prime, auditory synthesis, phoneme deletion, and letter naming. A detailed description of the tests that were used can be found in **Supplementary Material B**.

Procedure

Participants were split in two groups: Cohort 1 and Cohort 2. Participants in both cohorts were assessed on word decoding twice in Grade 1, halfway and at the end of the schoolyear. Participants in Cohort 2 were assessed on the five phonological awareness measures halfway the last year in kindergarten, and participants in Cohort 1 at the end of the last year in Kindergarten. For more details about the methods, see Braams and Bosman (2000).

Results

Means and standard deviations for all input variables and the output variable of the models are presented in **Table 8**. The output variable (word decoding efficiency) was transformed into binary classes, with 42 cases in 20% lowest decoding level class and 161 cases in the alternative class in Study 3a, and 48 cases in 20% lowest decoding level class and 151 cases in the alternative class in Study 3b. Scores of input variables were transformed into Z-scores, in order to even out effects of differing assessment moments between cohort 1 and 2.

Evaluation results of the models built with five machine learning techniques are presented in **Tables 9, 10**. Concerning identification of low decoding skill by the models, PP and NP

were evaluated. The results on the 95% confidence-intervals of PP reveal that between 8% and 25% of children allocated to the low decoding-category by the model, truly performed within the low decoding category when assessed with a decoding test; 75 to 92% of children allocated to the low decoding-category did not. The results on the 95% confidence-intervals of NP reveal that between 88 and 98% of children not allocated to the low decoding-category by the model, truly did not perform within the low decoding category when assessed with a decoding test; 2 to 12% of children not allocated to the low decoding-category did.

Concerning the usefulness of the models to detect low decoding skill, accuracy, sensitivity, and specificity were evaluated. The results of the 95% confidence-intervals of accuracy reveal that between 74 and 80% of the children were correctly allocated to the right decoding skill category by the models. The results on the 95% confidence-intervals of sensitivity reveal that between 20 and 57% of children who truly performed within the low decoding category when assessed with a decoding test, were indeed allocated to the low decoding group by the model; 43 to 80% of children with actual low decoding skills were not. The results on the 95% confidence-intervals of specificity reveal that between 80 and 83% of children who truly performed not within the low decoding-category when assessed with a decoding test, were indeed allocated to the not low decoding group by the model; 17 to 20% of children without actual low decoding skill were falsely allocated to the low decoding group by the model. The results of Cohen's κ -were evaluated to account for the possibility of accurate prediction by chance alone, which is an evident risk because of class imbalance in the present study. The results on the 95% confidence-intervals of Cohen's κ appear between 0.00 and 0.15, and indicate poor agreement between the models' predictions and the true values. See "Materials and Methods" Section for suggested interpretation of Cohen's κ .

The charts in the fourth column of **Figure 2** visualize results on identification of the 20% lowest decoding level class for all models in Study 3a, and the charts in the fifth column for Study 3b. For both studies, the curves indicate minimal yet predominantly

positive predictive abilities of all models, with little to no differences between models based on different techniques. The results on the 95% confidence-intervals of the AUC-statistics appear between 0.60 and 0.67, and indicate poor identification of first quintile word decoding for all models. See “Materials and Methods” Section for suggested interpretation of AUC statistics. Visual inspection of the curves confirms the similarity between the five models on results on the AUC statistic.

Conclusion

Study 3 was aimed at building a model of five cognitive skills input variables to predict which participants achieved within the lowest 20th percentile of decoding skills among children attending Grade 1 of regular Dutch education. The results again indicated that results on model building by nonlinear machine learning techniques are comparable to results on model building by more traditional linear techniques. The predictive ability of both the linear model and the four nonlinear models appeared to be poor, as indicated by the ROC curves and AUC statistics. Balancing the tradeoff between sensitivity and specificity reveals best results in terms of specificity, at the expense of sensitivity, that is, the models tend to identify about 80% of children without decoding problems correctly (specificity), but only about 20–50% of children who have indeed low decoding skill (sensitivity), which results in accuracy of about 75%. The results of Cohen’s κ , however, indicate poor agreement between the expected decoding category and the true decoding category, demonstrating that accuracy scores were affected by the imbalanced categories. Results for prediction of participants achieving within the lowest 20th percentile of decoding skills at T2 (Study 3a) were comparable to those at T3 (Study 3b).

GENERAL DISCUSSION

The main goal of the present study was to find whether cognitive factors of varying origin play any role in the development of reading skill in children with dyslexia and children with typical reading development. The present study focused on the possibilities of nonlinear machine learning techniques compared to traditional, linear statistical techniques.

The results of the present study pertain to two research questions: (1) To what extent are cognitive skills indicative of present decoding difficulties (Studies 1 and 2a), and (2) To what extent are cognitive skills predictive of future decoding difficulties yielded similar findings (Studies 2b, 3a, and 3b). Results of models built with nonlinear machine learning techniques were comparable to results of models built by the more traditional linear (GLM) technique and results on data of average school-aged children were comparable to those on data of children with severe dyslexia.

Cognitive Skills and Decoding Level

Irrespective of the building technique, models built to determine *present* decoding skill level based on an input set of *present* cognitive skills seem to produce the best results in the present study. However, these results indicate an inadequate base for making a potential thorough prediction of reading development.

The best performing models in the present study only produced an AUC 95%-confidence interval of 0.78–0.79, corresponding to a proportion explained variance of 0.23–0.25 (see Ruscio, 2008). Leaving 75% of variance unexplained, these results are comparable to the results on correlations between cognitive skills and reading skill discussed in the “Introduction” section (Hammill and McNutt, 1981; Scarborough, 1998; Swanson et al., 2003; Melby-Lervåg, 2012).

These results are in line with those of Verhoeven and Keuning (2018), who studied the predictive ability of several models on the dataset used for Study 1 of the present study. The present study differed from the study of Verhoeven and Keuning (2018) in two ways: (1) Verhoeven and Keuning (2018) used the logistic regression analyses (based on linear techniques, and using traditional statistical analysis for data modeling) to build the models, whereas the present study used several techniques (linear and nonlinear) of machine learning, and (2) Verhoeven and Keuning (2018) aimed to predict a dyslexia diagnosis (according to formal criteria of the Dutch Dyslexia Foundation) as the outcome variable, whereas the present study aimed to predict low decoding skill. Verhoeven and Keuning (2018) found an AUC of 0.84 when modeling cognitive skills variables on dyslexia diagnosis, which roughly agrees with the results of the AUC 95%-confidence intervals of best performing models in the present study (AUC 0.78–0.79, see **Table 4** in “Results” section). A question arises when we look at the outcome variables, which differed between these particular studies, as noted earlier. In our study, the variable of the dyslexia diagnosis was only to a limited extent related with low decoding skill (only 38% of dyslexic children belonged to the target category “decoding level within lowest 20%”) in the dataset of Verhoeven and Keuning (2018). This suggests that the children in their sample with a dyslexia diagnosis are not necessarily performing worst on decoding skill. Thus, irrespective of using traditional statistical analyses versus machine learning techniques for data modeling, and irrespective of the outcome being low word decoding or a dyslexia diagnosis, these studies on the role of cognitive skills in the field of reading ability suggest that relationships between cognitive skills and reading ability do exist. The relative weakness of this relationship does not justify the conclusion that these skills play a substantial role in reading performance. Therefore, cognitive skills and decoding skill are *moderately related*, leaving room for considering (1) other potential variables that could be related to reading skill, and (2) more complicated relationships than only the unidirectional explanation of cognitive skills being predictive for decoding skill.

Cognitive Skills and Decoding Progress

As Vellutino et al. (2004) pointed out, in order to predict future reading development, a model needs to include a variable that indicates progress in reading skill, which requires at least two moments of measurement. Studies 2b and 3b partly fulfilled this criterion, by modeling input cognitive variables on future decoding skill, corresponding to the unconditional models discussed in the “Introduction” section.

Models built to indicate future low decoding skill based on an input set of *present* cognitive skills seem to produce both weak and highly varying results in Studies 2b and 3b, which

poses doubts on the usefulness of these models. A small caveat has to be made because the relatively low sample size may have troubled the prospects of building adequate models. Results of the present study have yet to be confirmed by future research to strengthen conclusions.

The results on the role of cognitive skills in reading progress of the present study are in line with research based on linear techniques, and using traditional statistical analysis for data modeling. The present study confirms results of correlational meta-analyses discussed in the “Introduction” section, that revealed no evident role of cognitive skills in reading development (Nelson et al., 2003; Tran et al., 2011). Furthermore, the present study confirms correlational studies on decoding skill and cognitive skills in children with dyslexia. Walda et al. (2014) found many significant (80%) correlations between present cognitive skills (e.g., executive functions) and present decoding level, whereas significant correlations between present cognitive skills and progress in decoding skill were nearly absent (8.2%). Walda et al. (2022) studied the role of attentional skill in children with dyslexia and found that the working speed component of attentional skill was related to decoding level, whereas both the working speed component and the distraction component of attentional skill did not affect progress in decoding skill during remediation.

Thus, present and past research suggest that it is unlikely that cognitive skills play a determinant role in future decoding progress. Previous research based on traditional statistical methods discussed in the “Introduction” section revealed unconvincing and unequivocal results. Results of the present study, based on machine learning techniques, although preliminary, confirms results of previous studies. As machine learning techniques enable the model to take into account multidirectional, reciprocal, and concurrent relationships between variables, results of the present study add to previous studies the cautious suggestion that a determinant role for cognitive skills in decoding progress is not evident.

Implications for Future Research and Educational Practice

Apart from low sample size, some other factors may have flawed the potential of the used algorithms for model building. In Study 2, the sample included only children with dyslexia might have caused restriction of range in both input and output variables, as it is known that children with dyslexia tend to have lower abilities in both cognitive skills and decoding skill compared to average school aged children. Therefore, results of Study 2 should not be interpreted without considering results of Studies 1 and 3. Also, the output variable (decoding skill) was recoded in a binary variable with uneven classes (i.e., approximately 20–80% in study 1 and 3, and approximately 50%–50% in Study 2), which might have produced a slight class-imbalance problem, including only a small number of cases in the target category of the test set, especially when small samples are used.

In the present study, we have used several machine learning techniques to compare the results. Although different techniques differ in the extent to which they are sensitive to class imbalance,

this problem seems limited because various techniques agree on the results. This was the case for Studies 1 and 2. In Study 3, however, the results did differ between the techniques that were used. We have tried to limit this problem in Study 3 by expanding the test set to 25% of the sample (instead of 10% in Studies 1 and 2) and by training and testing all models on 100 different seeds of the sample.

Our findings should be considered preliminary due to the fact that previous research using nonlinear techniques and adequate modeling in this field is scarce. We, therefore, suggest future research on the role of cognitive factors in reading development applying nonlinear techniques for model building and to include designs and variables that expand insights in development of future reading skill, as was suggested by Vellutino et al. (2004) and Stuebing et al. (2015). Still, adequate model building is complicated because it requires sufficient sample sizes and longitudinal data gathering on large sets of variables, producing high costs, dedicated and often long-lasting participation of many stakeholders, and time-consuming strategies. Such research on the specific group of children with learning disabilities is even more complicated: Because of limited prevalence, sampling a large number of children with dyslexia is problematic. Pending more substantial results in future research, we would like to propose some thoughts on practical implications of present results.

From a theoretical perspective, the present results reveal that cognitive skills do correlate with reading skill, suggesting a relationship. The nature of this relationship is, however, unclear and it is as yet unlikely that cognitive skills will make good predictions about reading skill.

From a prevention perspective, it seems unlikely that children at risk for reading difficulties could be identified by assessment of cognitive skills, even if these cognitive skills consist of phonological awareness (specifically the case in Studies 1 and 3). In this line, it seems not sensible to use results of individual cognitive skills assessment in identifying children eligible for a dyslexia diagnosis. Instead, individual reading development, operationalized by repeated assessment of (word and text) decoding skills, should make better predictions of longitudinal reading development. Fortunately, in the Netherlands the Protocol Dyslexie Diagnose en Behandeling version 3.0 [Protocol Dyslexia Diagnosis and Treatment] (Tijms et al., 2021) has dropped cognitive skills criteria as requirements to diagnose dyslexia and specifies solely criteria on the reading and writing skill level. Still, the Protocol Dyslexie Diagnose en Behandeling version 3.0 considers cognitive skills as protective and/or risk factors for developing reading disorder. However, pending more effective strategies to identify these children, it seems even more tenable to monitor initial literacy development in an intensive and professional way, and intervene immediately when potential struggles appear in any child (independent of a dyslexia diagnosis). It is noteworthy that the nature of machine learning techniques does not allow to identify which cognitive variables could be relevant from the prevention perspective, and we certainly do not want to fall into the trap of detailed interpretations about individual variables involved in a black box model (e.g., see Rudin, 2019). A slight drawback of machine

learning techniques is that the mechanisms of the results that were produced are difficult to interpret (Lantz, 2019). Note, however, if there is any predictive value of cognitive factors regarding reading development, machine learning techniques should be able to at least identify them. The fact that none of the models in the present study had sufficient predictive value for decoding skill, seems to suggest that these cognitive skills have no valuable role in reading development from the prevention perspective.

In a similar way, from a remediation perspective, it seems unlikely that training cognitive skills could prevent children from developing reading difficulties or that training cognitive skills could stimulate reading development. In this line, we agree with Hammill (2004), that:

(1) professionals interested in improving literacy skills should focus on teaching written language abilities such as print awareness and book handling, letters, phoneme-letter correspondences, word recognition, alphabet knowledge, and comprehension and (2) the current interest in the role of nonprint abilities in reading such as phonological awareness, rapid naming, intelligence, and memory might be overemphasized (p. 453).

Moreover, awaiting results from future research suggested earlier, we propose to reconsider the role of cognitive skills in criteria for dyslexia or severe reading problems. If a lack of cognitive skills does prove to have accidental side effects in some but not all children with low reading skill, assessing cognitive skills does not have any use in diagnosis and selection, let alone remediation, of children who need extra attention. This is in accordance with a proposal of McEaneaney et al. (2006) to move away from defining disabilities relying on factors within individuals and to center instructional needs and future steps in the process of remediation, and with Vellutino et al. (2004), who argue that inadequate instruction and other experiential factors play a crucial role for many children developing reading difficulties, and that cognitive factors such as IQ should have less emphasis in diagnosing reading difficulties. Thus, both assessment and instruction of children who are learning to read should be focused on reading skill itself.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: study 2: <https://doi.org/10.17026/dans-2cq-96v9>; study 3: <https://doi.org/10.17026/dans-xun-4v56>.

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' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AB and SW contributed to conception and design of the study. SW organized the database and wrote the first draft of the manuscript. SW and FH performed the statistical analyses. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This publication is part of the project "Influence of limited cognitive skills on remediation success of children with severe reading and spelling difficulties" (project number 023.010.019), which was (partly) funded by the NWO, the Dutch Research Council.



ACKNOWLEDGMENTS

The present data base of Study 1 originates from research conducted by Ludo Verhoeven (Expertisecentrum Nederlands) and Jos Keuning (Cito) on precursors of dyslexia in Dutch children (see Verhoeven and Keuning, 2018). The nature of dyslexia in a transparent orthography. *Scientific Studies of Reading*, 22(1), 7–23. The research of Study 2 was supported by Braams & Partners, Clinic for Learning Disorders, in Deventer, the Netherlands. Special appreciation is expressed to the children who participated, their parents, teachers, and the reading specialists of Braams & Partners. The dataset analyzed for this study can be found in DANS-EASY (<https://doi.org/10.17026/dans-2cq-96v9>). The present data base of Study 3 originates from research conducted by Tom Braams (Braams & Partners) and Anna Bosman (Radboud University) on phonological awareness, literacy and instruction on reading and spelling (see Braams and Bosman, 2000). (Fonologische vaardigheden, geletterdheid en lees- en spellinginstructie. *Tijdschrift voor Orthopedagogiek*, 39, 199–211). The dataset analyzed for this study can be found in DANS-EASY (<https://doi.org/10.17026/dans-xun-4v56>). Special appreciation is expressed to professor Tom Heskes (Radboud University) for sharing his knowledge about data science and machine learning, and advising our research group how to apply machine learning techniques to the field of learning and learning problems.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.869352/full#supplementary-material>

REFERENCES

- Al Otaiba, S., and Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? *J. Learn. Disabil.* 39, 414–431. doi: 10.1177/00222194060390050401
- Araújo, S., and Faisca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. *Sci. Stud. Read.* 23, 349–368. doi: 10.1080/10888438.2019.1572758
- Astle, D. E., Bathelt, J., The Calm Team, and Holmes, J. (2018). Remapping the cognitive and neural profiles of children who struggle at school. *Dev. Sci.* 22:e12747. doi: 10.1111/desc.12747
- Blomert, L. (2006). *Protocol Dyslexie Diagnostiek en Behandeling [Protocol Dyslexia Diagnostics and Remediation]*. Available online at: <https://docplayer.nl/19724769-Protocol-dyslexie-diagnostiek-en-behandeling-1.html> (accessed March, 21, 2022).
- Braams, T., and Bosman, A. M. T. (2000). Fonologische vaardigheden, geletterdheid en lees- en spellinginstructie [Phonological skills, literacy, and reading and spelling instruction]. *Tijdschr. Orthopedagogiek* 39, 199–211. doi: 10.47635/kjld.2019.16.1.199
- Castles, A., and Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition* 91, 77–111. doi: 10.1016/S0010-0277(03)00164-1
- Catts, H. (2017). “Early identification of reading disabilities,” in *Theories of Reading Development*, eds K. Cain, D. Compton, and R. Parrila (Amsterdam: Benjamins), 311–331. doi: 10.1075/swll.15.18cat
- Connor, C. M., and Morrison, F. J. (2017). “Child characteristics by instruction interactions, literacy, and implications for theory and practice,” in *Theories of Reading Development*, eds K. Cain, D. Compton, and R. Parrila (Amsterdam: Benjamins), 507–524. doi: 10.1075/swll.15.27con
- Hammill, D. D. (2004). What we know about correlates of reading. *Except. Children* 70, 453–469. doi: 10.1177/001440290407000405
- Hammill, D. D., and McNutt, G. (1981). *The Correlates of Reading: The Consensus of Thirty Years of Correlational Research*. Austin, TX: PRO-ED.
- Holland, J. H. (2006). Studying complex adaptive systems. *J. Syst. Sci. Complex.* 19, 1–8. doi: 10.1007/s11424-006-0001-z
- Huedo-Medina, T. B., Sánchez-Meca, J., and Marín-Martínez, F. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychol. Methods* 11, 193–206. doi: 10.1037/1082-989X.11.2.193
- Inspectie van het Onderwijs (2019). *Dyslexieverklaringen. Verschillen Tussen Scholen Nader Bekeken [Dyslexia Diagnoses. A Closer Look at Differences Between Schools]*. Utrecht: Ministerie van OCW.
- Lantz, B. (2019). *Machine Learning with R. Expert Techniques for Predictive Modeling*. Birmingham: Packt.
- Lesmeister, C. (2019). *Mastering Machine Learning with R*, 3rd Edn. Birmingham: Packt.
- Loizou, A., and Laouris, Y. (2010). Developing prognosis tools to identify learning difficulties in children using machine learning technologies. *Cogn. Comput.* 3, 490–500. doi: 10.1007/s12559-010-9052-5
- Loneragan, A., Doyle, C., Cassidy, C., Mahon, S. M., Roche, R. A. P., Boran, L., et al. (2019). A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD. *J. Cogn. Psychol.* 31, 725–749. doi: 10.1080/20445911.2019.1669609
- Matsumoto, M., and Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.* 8, 3–30. doi: 10.1145/272991.272995
- McEaneaney, J. E., Lose, M. K., and Schwartz, R. M. (2006). A transactional perspective on reading difficulties and response to intervention. *Read. Res. Q.* 41, 117–128. doi: 10.1598/RRQ.41.1.7
- Melby-Lervåg, M. (2012). The relative predictive contribution and causal role of phoneme awareness, rhyme awareness and verbal short-term memory in reading skills: a review. *Scand. J. Educ. Res.* 56, 363–380. doi: 10.1080/00313831.2011.594611
- Nelson, J. R., Benner, G. J., and Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: a meta-analytic review. *Learn. Disabil. Res. Pract.* 18, 255–267. doi: 10.1111/1540-5826.00080
- Parrila, R., Dudley, D., Song, S., and Georgiou, G. K. (2019). A meta-analysis of reading-level match dyslexia studies in consistent alphabetic orthographies. *Ann. Dyslexia* 70, 1–26. doi: 10.1007/s11881-019-00187-5
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* 1, 206–215. doi: 10.1038/s42256-019-0048-x
- Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychol. Methods* 13, 19–30. doi: 10.1037/1082-989X.13.1.19
- Russell, S. J., and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Hoboken, NJ: Prentice Hall.
- Scarborough, H. S. (1998). “Early identification of children at risk for reading disabilities: phonological awareness and some other promising predictors,” in *Specific Reading Disability: A View of the Spectrum*, eds P. Accardo, A. Capute, and B. Shapiro (Timonium, MD: York Press), 75–119.
- Spinoza, B. (1678/1928). *Ethica in Meetkundigen Trant Uiteengezet*. (N. van Suchtelen, Trans.). Amsterdam: Maatschappij voor goede en goedkope lectuurbibliotheek.
- Stuebing, K. K., Barth, A. E., Trahan, L. H., Miciak, J., and Fletcher, J. M. (2015). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis. *Rev. Educ. Res.* 85, 395–429. doi: 10.3102/0034654314555996
- Swanson, H. L., Trainin, G., Nevochea, D. M., and Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: a meta-analysis of the correlation evidence. *Rev. Educ. Res.* 73, 407–440. doi: 10.3102/00346543073004407
- Tijms, J., De Bree, E. H., Bonte, M., De Jong, P. F., Loykens, E., and Reij, R. (2021). *Protocol Dyslexie Diagnostiek en Behandeling – Versie 3.0 [Protocol Dyslexia Diagnostics and Treatment – Version 3.0]*. Nederlands Kwaliteitsinstituut Dyslexie. Available online at: <https://www.nkd.nl/app/uploads/2021/09/Protocol-Dyslexie-Diagnostiek-en-Behandeling-3.0-versie-0.99.pdf> (accessed March, 21, 2022).
- Tran, L., Sanchez, T., Arellano, B., and Swanson, H. L. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *J. Learn. Disabil.* 44, 283–295. doi: 10.1177/0022219410378447
- Trevethan, R. (2017). Sensitivity, specificity, and predictive values: foundations, liabilities, and pitfalls in research and practice. *Front. Public Health* 5:307. doi: 10.3389/fpubh.2017.00307
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., and Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *J. Child Psychol. Psychiatry* 45, 2–40. doi: 10.1046/j.0021-9630.2003.00305.x
- Verhoeven, L., and Keuning, J. (2018). The nature of dyslexia in a transparent orthography. *Sci. Stud. Read.* 22, 7–23. doi: 10.1080/10888438.2017.1317780
- Walda, S. A. E., van Weerdenburg, M., Van der Ven, A. H. G. S., and Bosman, A. M. T. (2022). Literacy progress in children with dyslexia and the role of attention. *Read. Writ. Q.* 38, 19–32. doi: 10.1080/10573569.2021.1897910
- Walda, S. A. E., van Weerdenburg, M., Wijnants, M. L., and Bosman, A. M. T. (2014). Progress in reading and spelling of dyslexic children is not affected by executive functioning. *Res. Dev. Disabil.* 35, 3431–3454. doi: 10.1016/j.ridd.2014.08.013

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 Walda, Hasselman and Bosman. 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.



Determinants of Major Choice and Academic Expectations: Testing a Prediction Model Across Gender

Sonia Alfonso¹, António M. Diniz², Angeles Conde¹ and Mar García-Señorán^{1*}

¹ Department of Evolutionary Psychology, University of Vigo, Ourense, Spain, ² Department of Psychology, Research Center in Education and Psychology, School of Social Sciences, University of Évora, Évora, Portugal

With this study, we aim to test the predictive relationships between determinants of major choice (DMC) and academic expectations (AEs) and to analyze gender differences, using six items of the Determinants of Major Choice Scale and the Academic Perceptions Questionnaire to assess AEs. A convenience sample of Portuguese ($n = 839$) and Spanish ($n = 1,001$) first-year students (age-range = 17–23 years), mostly composed of women (56.9%, $n = 1,047$), was selected from two public universities. The invariance of the multivariate regression model with latent variables of the effect of DMC on AEs, with determinants linked to Personal Characteristics (PCs; e.g., capacities) and Mediating Agents (MAs; e.g., parents) as AE predictors, was tested across gender with LISREL. The invariance test of the multivariate regression model across gender fit the data well and revealed an equivalence of slopes between women and men, which allows a unique interpretation of the model's predictive relationships for both genders. We also found statistically significant predictive relationships of PCs for six AE factors and MAs for five AE factors. The results showed theoretical relationships with the self-determination theory. At a practical level, they indicated the importance of PCs and MAs to design AE intervention programs in Higher Education (HE) institutions.

Keywords: major choice, academic expectations, first-year students, gender, structural equation modeling

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

Reviewed by:

Maria Do Céu Taveira,
University of Minho, Portugal
Teresa Pozo-Rico,
University of Alicante, Spain

*Correspondence:

Mar García-Señorán
mseñoran@uvigo.es

Specialty section:

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

Received: 03 January 2022

Accepted: 28 February 2022

Published: 07 April 2022

Citation:

Alfonso S, Diniz AM, Conde A and
García-Señorán M (2022)
Determinants of Major Choice
and Academic Expectations: Testing
a Prediction Model Across Gender.
Front. Psychol. 13:847843.
doi: 10.3389/fpsyg.2022.847843

INTRODUCTION

Educational expectations in Higher Education (HE) are students' aspirations and desires about what they hope to achieve during their academic life. Educational expectations are cognitive and motivational, so they represent a key element for adaptation to the university context and the decision to prolong studies (Cabrera et al., 1993; Tinto, 1993; Braxton et al., 2014). They constitute the mental foundations of everyday engagement, which guides student behavior inside and outside educational institutions toward higher-order educational goals such as high school or college graduation (Lorenz et al., 2020). Educational expectations represent the lower or realistic limit of a continuum of educational alternatives ordered from a lower to a higher level of difficulty where aspirations are at the upper or idealistic limit. Along this continuum, students make decisions about educational alternatives based on their potential for fulfillment and their personal value for the students (Zimmermann, 2020). Students adapt and revise their expectations in a continuous Bayesian updating process (Morgan, 2005) in response to new information received or educational experiences. Thus, educational expectations provide early insight into students' subjective perceptions of opportunities and constraints to further study (Anders, 2017).

The development of the continuum, represented by educational expectations and aspirations, requires time and mental maturity. Children broaden their knowledge of the social structure and

their position in it through the family, the peer group, and the school, and later on, as youths, they become aware of their potential, and of the classification of occupations according to their prestige, and they recognize the interaction between income, occupation, and educational level (Gottfredson, 2002). Although the students must evolve to discover their desired place in society and to be able to distinguish between ideal or real aspirations (expectations), youth is where a potentially realizable educational choice is made throughout their educational trajectory (Zimmermann, 2020). These predictions or expectations, which may be more or less idealized or realistic, influence behavior: if it occurs, the planned behavior will be performed in line with the interpretation of the situation. This influence occurs by translating experiences, knowledge, attitudes, motivations, and beliefs into academic actions (Kuh et al., 2005). Understanding one's subsequent behavior grants an adaptive role to expectations that may or may not match the attainment of students' desires and hopes.

In the context of HE, a student's career choice is a crucial moment that will have a high impact on their professional life and future achievements (Ahmed et al., 2017). An appropriate career choice will make the student feel satisfied and motivated, whereas an inappropriate career choice can lead to the abandonment of their studies (Rodríguez-Muñoz et al., 2019). As for the change of major subject, one of the aspects that differentiate students who change from those who do not is that the former attribute less primacy to their interests and skills when choosing a course (Diniz and Almeida, 2007). On another hand, students' motivation and positive expectations toward studies (educational expectations) are variables that promote academic permanence and success (Castillo-Sánchez et al., 2020).

There are various ways of classifying the determinants of career choice that condition success or failure: individual and contextual (Zacher et al., 2019); internal (personal characteristics, interests: Päßler and Hell, 2012; Zafar, 2013) and external (socioeconomic background: Ma, 2009; Parker et al., 2012). In addition, a third type, interpersonal determinants, is related to the influence of mediating agents (parental support, peer influence, and the interaction with teachers and other educators: Whiston and Keller, 2004; Fouad et al., 2010; Lerkkanen et al., 2012). These classifications are an anchor in the Social Cognitive Career Theory (SCCT; Lent et al., 1994). Among the determinants of career choice, the internal determinants seem to be the most important for the choice of major subjects in both genders (Päßler and Hell, 2012; Zafar, 2013).

From the interaction of the set of internal, external, and contextual factors, students obtain various types of information that consolidate into a series of perceptions about the curriculum of the chosen degree program. These perceptions or "a priori expectations" will be modified over time, based on the actual knowledge of the studies acquired while pursuing the degree (Navarro and Soler, 2014), and will determine whether or not the career choice was appropriate (Dias, 2011). Thus, a close relationship between determinants of career choice and educational expectations can be assumed.

Current studies link internal determinants of career choice with aspirations and expectations of job satisfaction,

development of autonomy and self-efficacy, and enhancement of learning experiences, in addition to deepening of knowledge in the area of interest (Eren, 2017; McLean et al., 2019). External determinants of career choice relate to aspirations and expectations associated with pursuing a particular profession that ensures appropriate financial remuneration, job security, and promotion opportunities (Akosah-Twumasi et al., 2018). Parents, teachers, and friends also play a relevant role in consolidating college students' expectations, especially those linked to opportunities for socialization in that context (Akosah-Twumasi et al., 2018; Lorenz et al., 2020) and those related to satisfying the desires of others (Guan et al., 2015).

PURPOSE OF THE PRESENT STUDY

The main objective of this study is to test a predictive model between determinants of major choice (DMC) and academic expectations (AEs) of first-year university students, analyzing their gender invariance.

From a multifactorial conception of AEs, they were grouped into seven factors related to future employability, personal and social development, student mobility, political/citizen engagement, social pressure, quality of education, and social interaction. These factors were analyzed considering country (Deaño et al., 2015) and country and gender (Diniz et al., 2018), resulting in a cross-cultural validation of a questionnaire to assess them (Almeida et al., 2018), which was used in this study. The DMC considered in this study involves the influence of personal characteristics (PCs) and mediator agents (MAs). We used six items retrieved from the Determinants of Major Choice Scale (Diniz, 2008).

Taking into account the target variables and the instruments used for their operationalization, as a first hypothesis, we expected that their psychometric properties would allow testing a prediction model between the DMC and the factors of expectations. Furthermore, considering the variability and number of factors that affect the DMC and AEs, country and gender were introduced as variables to be considered in the study.

As a second hypothesis, we conjectured that the DMC factors would have a significant impact on the factors of expectations. To our knowledge, no research has considered both types of constructs conjointly through a prediction model while considering gender, a key variable of major selection (Parker et al., 2012). The exploration of the DMC's impact on AEs through a multivariate regression model with latent variables across gender can shed light on these relationships.

MATERIALS AND METHODS

Sample

A convenience sample of 1,840 Portuguese ($n = 839$) and Spanish ($n = 1,001$) first-year university students (age-range = 17–23 years; $Mdn = 18$) was selected from two public universities. More of the voluntary participants in this study were enrolled in the humanities and social studies area (60.9%, $n = 1,121$) than

TABLE 1 | 9-Factor oblique model: factorial invariance across countries and gender.

Model	$SB\chi^2_{(df)}$	RMSEA _[90% CI]	SRMR		CFI
Countries			Spain (<i>n</i> = 1,001)	Portugal (<i>n</i> = 839)	
M1	7314.14 _(2,088)	0.052 _[0.051–0.054]	0.075	0.077	0.975
M2	9846.39 _(2,268)	0.060 _[0.059–0.062]	0.101	0.094	0.963
M3	9284.72 _(2,267)	0.058 _[0.057–0.060]	0.101	0.095	0.966
M1-M3					ΔCFI = −0.009
Gender			Men (<i>n</i> = 793)	Women (<i>n</i> = 1,047)	
M1	7007.23 _(2,088)	0.051 _[0.049–0.052]	0.070	0.076	0.974
M2	7974.66 _(2,268)	0.052 _[0.051–0.054]	0.099	0.079	0.970
M1-M2					ΔCFI = −0.006

M1, form invariance; M2, M1 fully invariant; M3, M2 with the intercept of Item 37 of the Academic Perceptions Questionnaire freely estimated across countries (Spain 5.61, Portugal, 3.53). RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; CFI, comparative fit index; Δ = difference between tested model and baseline model.

in the scientific and technological area. Most participants were women (56.9%, *n* = 1,047).

Materials

We used six items retrieved from two factors of a scale for the assessment of HE students' valuation of the DMC (Diniz, 2008): determinants linked to PCs (e.g., capacities) and MAs (e.g., parents). Items are rated on a 5-point Likert-type scale, ranging from 1 (*not at all important*) to 5 (*decisive or extremely important*).

To assess AEs, we used the Academic Perceptions Questionnaire (APQ; Almeida et al., 2018) which measures seven factors of expectations, with six items each: (1) Training for Employment (TE), (2) Personal and Social Development (PSD), (3) Student Mobility (SM), (4) Political Engagement and Citizenship (PEC), (5) Social Pressure (PS), (6) Quality of Education (QE), and (7) Social Interaction (SI). Items are rated on a 6-point Likert-type scale ranging from 1 (*totally disagree*) to 6 (*totally agree*).

Data Collection and Analysis

We collected the data before the SARS-CoV-2 pandemic, at the beginning of the first semester (mid-October and the beginning of November), after obtaining students' informed consent. We used the IBM SPSS Statistics for Windows (version 24.0) for descriptive data analysis and to deal with missing values (substituted by the respective distributional median).

The invariance of the multivariate regression model with latent variables, or factors, of the effect of DMC on AEs, with both PCs and MAs as AE predictors, was tested across gender with LISREL 8.80 (Jöreskog and Sörbom, 2006).

We performed the multivariate regression model invariance test across gender only after the invariance testing, through confirmatory factor analysis, of its analogous 9-factor oblique model across countries and gender, and the subsequent inspection of its psychometric properties to complete the model's structural validity study (Anderson and Gerbing, 1988; Jöreskog and Sörbom, 1993). We followed Fornell and Larcker (1981) criteria of factors' convergent and discriminant validity (CV and

DV), as well as their composite reliability (CR). CV is based on the items' average variance extracted (AVE), which should be at least 0.50, and DV is based on the comparisons of any two factors' shared variance (φ^2 ; squared disattenuated correlation) and the AVE of each factor, which should be higher than φ^2 . Factors' reliability should be at least 0.70, and 0.80 is desirable for group comparisons (Nunnally and Bernstein, 1994).

Because the observed variables were ordinal, model estimation and testing were performed with the underlying bivariate normal approach (Jöreskog, 2005), using the robust Satorra–Bentler (SB) scaled correction for maximum likelihood (Satorra and Bentler, 1994). This approach involves the estimation in PRELIS 2 (Jöreskog and Sörbom, 1996) of the means and the asymptotic covariance matrix of the polychoric covariances of each group's latent normal counterparts of the observed variables, under thresholds fixed to the pooled thresholds estimated in the combined group. The result of this multi-group analysis was used as input for model estimation and testing with the SIMPLIS command language (Jöreskog and Sörbom, 1993) under the independence of the items' error measurement, or residuals (uniqueness and random error), and the factor's identification was ensured by setting to one (1.00) the path to one of its items.

The analysis was conducted by comparing the 9-factor oblique model's form of invariance (all parameters freely estimated across groups) with a more restrictive model (i.e., with more degrees of freedom), the fully invariant model, which is invariant across groups at measurement (factor scores, intercepts, and residuals) and factor levels (variances and covariances).

The multivariate regression model was specified by freely estimating the error covariances between the criteria (AE factors), assuming that the predictors (PCs and MAs) do not capture the totality of their correlations. The model was tested across gender with different slopes and then, with equal slopes (Jöreskog and Sörbom, 1993) to assess the invariance of the model with equal slopes.

The assessment of model invariance was based on the variation (Δ) of the comparative fit index (CFI) and, in addition, on the following goodness-of-fit indices and recommended benchmarks to indicate a good fit (Hu and Bentler, 1998): a

TABLE 2 | Fully invariant 9-factor oblique model by country and gender: common metric's robust maximum likelihood estimates completely standardized, convergent validity, and composite reliability ($N = 1,840$).

Item (Factor)	Countries		Gender	
	β	R^2	β	R^2
Personal aptitudes and capacities (PCs)	0.71	0.50	0.70	0.49
Interest in the professional area	0.67	0.45	0.68	0.46
Way of being and personal characteristics	0.74	0.54	0.73	0.54
AVE/CR	0.50/0.75		0.50/0.75	
Parents, siblings, or other relatives (MAs)	0.71	0.50	0.70	0.50
Friends, colleagues, or girlfriend/boyfriend	0.70	0.50	0.71	0.50
Teachers	0.53	0.28	0.54	0.29
AVE/CR	0.43/0.69		0.43/0.69	
1. Achieve a prestigious profession (TE)	0.59	0.35	0.57	0.32
8. Have better career opportunities in the job market	0.79	0.62	0.77	0.59
15. Obtain training to achieve a good job	0.83	0.69	0.83	0.69
22. Qualify to achieve professional success	0.84	0.71	0.84	0.71
29. Ensure a successful professional career	0.82	0.67	0.83	0.69
36. Achieve in-service training to facilitate access to work	0.72	0.52	0.72	0.52
AVE/CR	0.59/0.90		0.58/0.89	
2. Improve my identity, autonomy, and self-confidence (PSD)	0.69	0.48	0.68	0.46
9. Develop my personality traits	0.72	0.52	0.72	0.52
16. Gain self-confidence in my potential	0.76	0.58	0.76	0.58
23. Have goals in life	0.74	0.55	0.73	0.53
30. Deal autonomously with life's difficulties	0.74	0.55	0.73	0.53
37. Acquire skills to be a responsible adult	0.77	0.59	0.73	0.53
AVE/CR	0.54/0.88		0.52/0.87	
3. Participate in student exchange programs (SM)	0.80	0.64	0.80	0.64
10. Accomplish a stay in another country	0.84	0.71	0.84	0.71
17. Obtain training that allows me to achieve international employment	0.77	0.59	0.76	0.58
24. Obtain international quality training	0.69	0.48	0.73	0.53
31. Spend some of my study time in another country	0.87	0.76	0.73	0.53
38. Achieve an international title	0.86	0.74	0.77	0.59
AVE/CR	0.65/0.92		0.60/0.90	
4. Contribute to improving the world and society (PEC)	0.77	0.59	0.77	0.59
11. Solve problems that disadvantaged people face	0.77	0.59	0.77	0.59
18. Develop a critical view of the world	0.71	0.50	0.71	0.50
25. Participate in volunteer activities	0.68	0.46	0.68	0.46
32. Be an educated citizen committed to society	0.75	0.56	0.75	0.56
39. Contribute to the improvement of the human condition	0.83	0.69	0.82	0.67
AVE/CR	0.57/0.89		0.56/0.89	
5. Meet my family's expectations (SP)	0.77	0.59	0.77	0.59
12. Not obtain worse grades than other classmates	0.55	0.30	0.55	0.30
19. Not disappoint my family or friends because of my grades	0.83	0.69	0.83	0.69
26. Seize the educational opportunity provided by my family	0.54	0.29	0.54	0.29
33. Fulfill the desire of people close to me who encourage my higher education	0.75	0.56	0.75	0.56
40. Achieve a close to or higher level of education than that obtained by my parents (or older siblings)	0.57	0.32	0.57	0.32
AVE/CR	0.46/0.83		0.46/0.83	
6. Participate in debates or scientific conferences (QE)	0.52	0.27	0.53	0.28
13. Deepen my knowledge of specific subjects	0.61	0.37	0.62	0.38
20. Participate in research projects	0.53	0.28	0.53	0.28
27. Correspond to society's investment in higher education	0.65	0.42	0.64	0.41
34. To get a satisfactory academic performance to conform a good curriculum	0.71	0.50	0.72	0.52
41. To have teachers with recognized capacity in the area of training they teach	0.60	0.36	0.59	0.35
AVE/CR	0.37/0.78		0.37/0.78	

(Continued)

TABLE 2 | (Continued)

Item (Factor)	Countries		Gender	
	β	R^2	β	R^2
7. Enjoy living with others and having fun (SI)	0.75	0.56	0.74	0.55
14. Engage in extracurricular activities	0.50	0.25	0.51	0.26
21. Establish a weekly schedule that allows for other activities	0.63	0.40	0.62	0.38
28. Attend university student parties	0.65	0.42	0.65	0.42
35. Have a group of friends with whom I can relax and socialize outside of class	0.85	0.72	0.85	0.72
42. Socializing/connecting with a new group of friends	0.77	0.59	0.77	0.59
AVE/CR	0.49/0.85		0.49/0.85	

PCs, Personal Characteristics; MAs, Mediating Agents; TE, Training for Employment; PSD, Personal and Social Development; SM, Student Mobility; PEC, Political Engagement and Citizenship; SP, Social Pressure; QE, Quality of Education; SI, Social Interaction. β , standardized factor loading; R^2 (communality), $1 - \epsilon$ (standardized residual); AVE, average variance extracted; CR, composite reliability.

TABLE 3 | Fully invariant 9-factor oblique model by country and gender: Disattenuated correlations from the common metric completely standardized solution ($N = 1,840$).

	PCs	MAs	TE	PSD	SM	PEC	SP	QE	SI
Countries									
PCs	1.00								
MAs	0.06	1.00							
TE	0.31	0.06	1.00						
PSD	0.38	0.07	0.78	1.00					
SM	0.16	0.06	0.41	0.48	1.00				
PEC	0.38	0.07	0.51	0.74	0.54	1.00			
SP	0.05	0.41	0.48	0.46	0.29	0.36	1.00		
QE	0.35	0.14	0.80	0.82	0.52	0.78	0.66	1.00	
SI	0.16	0.10	0.57	0.65	0.45	0.48	0.42	0.58	1.00
Gender									
PCs	1.00								
MAs	0.06	1.00							
TE	0.33	0.06	1.00						
PSD	0.40	0.07	0.79	1.00					
SM	0.17	0.07	0.41	0.47	1.00				
PEC	0.38	0.07	0.53	0.74	0.53	1.00			
SP	0.00	0.41	0.49	0.46	0.29	0.36	1.00		
QE	0.35	0.15	0.82	0.82	0.52	0.78	0.66	1.00	
SI	0.16	0.12	0.56	0.64	0.45	0.47	0.43	0.58	1.00

Refer to **Table 2** for abbreviations.

comparative fit index (CFI) close to or above 0.95, a root mean square error of approximation (RMSEA) close to or below 0.06, and a standardized root mean square residual (SRMR) close to or below 0.08. Values of δ CFI between a restricted model and a baseline model of less than -0.01 indicate non-invariance of the restricted model (Cheung and Rensvold, 2002).

RESULTS

Oblique Model Invariance Across Countries and Gender

The 9-factor oblique model was fully invariant across countries and gender, as shown in **Table 1**, but with the nuance of a differential item functioning between countries presented by

Item 37 of the APQ. This difference occurred in its intercepts (Spain = 5.61; Portugal = 3.53), not indicating different levels of item ambiguity but merely differences in the item's attractiveness to the samples (Ferrando, 1996).

In **Tables 2, 3**, we present the psychometric data for the fully invariant 9-factor oblique model by country and gender.

In **Table 2**, it can be seen that all items represented their respective factors well ($\beta > 0.50$), and also that the factors' CV (AVE) and CR ranged between acceptable to good, except for MAs and, mainly, for QE. The QE factor showed similar problems with other Portuguese and Spanish samples (Diniz et al., 2018). Without Items 6 and 20, the factor CV improves and, in the opposite, its reliability deteriorates in both samples ($VME = 0.41$; $FC = 0.73$). Considering previous psychometric results (Diniz et al., 2018), and that the items showed standardized factor

TABLE 4 | Comparisons of the means of the model's predictor factors by gender.

Predictor factors	Men (<i>n</i> = 739)	Women (<i>n</i> = 1,047)	SE	<i>T</i>
	<i>M</i>	<i>M</i>		
Personal characteristics	0.00	0.41	0.07	5.86***
Mediating agents	0.00	−0.05	0.05	−1.02

****p* < 0.001.

loadings higher than 0.50 (all $R^2 > 0.26$, high effect size; Cohen, 1988), they were retained in the model for further analysis. Furthermore, the SM factor presented a better CV in countries than in gender.

Regarding factors' DV, taking **Tables 2, 3** conjointly, undesirable correlations (φ ; **Table 3**) and, consequently, shared variances (φ^2), were found between TE–QE (countries, $\varphi^2 = 0.64$; gender, $\varphi^2 = 0.67$), PSD–QE (countries and gender, $\varphi^2 = 0.67$), and PEC–QE (countries and gender, $\varphi^2 = 0.61$), jeopardizing their DV (see respective CV in **Table 2**). Similar results for these factors emerged in a previous study (Almeida et al., 2018). In the current study, problems in DV also appeared between TE–PSD (countries, $\varphi^2 = 0.61$; gender, $\varphi^2 = 0.62$), and minor problems between PSD and PEC because their shared variance (countries and gender, $\varphi^2 = 0.55$) was close to and lower than their AVE (see **Table 2**).

Finally, PC and MA factors were independent ($\varphi = 0.06$), verifying the assumption regarding the predictors of the regression model's gender invariance.

Regression Model Invariance Across Gender

Once the oblique model's full invariance was guaranteed across countries and gender, the regression model could be estimated and tested across gender under full invariance at measurement and factor levels to examine the slopes' invariance between predictors (PCs and MAs) and criteria (AEs).

Before examining the model slopes' invariance, we noted that model testing revealed a different pattern for PCs and MAs factors' mean comparisons by gender. Although there were gender differences in PCs, there were no differences in MAs (**Table 4**).

In **Table 4**, it can be seen that women presented a significantly higher PC factor means than men, considering that LISREL 8 fixes to zero the factor means of the first group of data, the men's

group, and estimates factor means of the second group of data, the women's group (Jöreskog and Sörbom, 1993).

The regression model with equal slopes was fully invariant across gender (**Table 5**), and its unstandardized structural estimates mostly presented very significant effects of PCs and MAs on AEs (**Figure 1**).

Figure 1 shows that the only AE not well predicted by PCs was SP, and MAs did not predict TE and PSD and only slightly predicted SM. Finally, the magnitude of the effect size of both PCs and MAs on AEs, according to Cohen's criteria (1988), was around *small* ($R^2 = 0.02$) for SM and SI, and around *medium* ($R^2 = 0.13$) for all the others.

DISCUSSION AND CONCLUSION

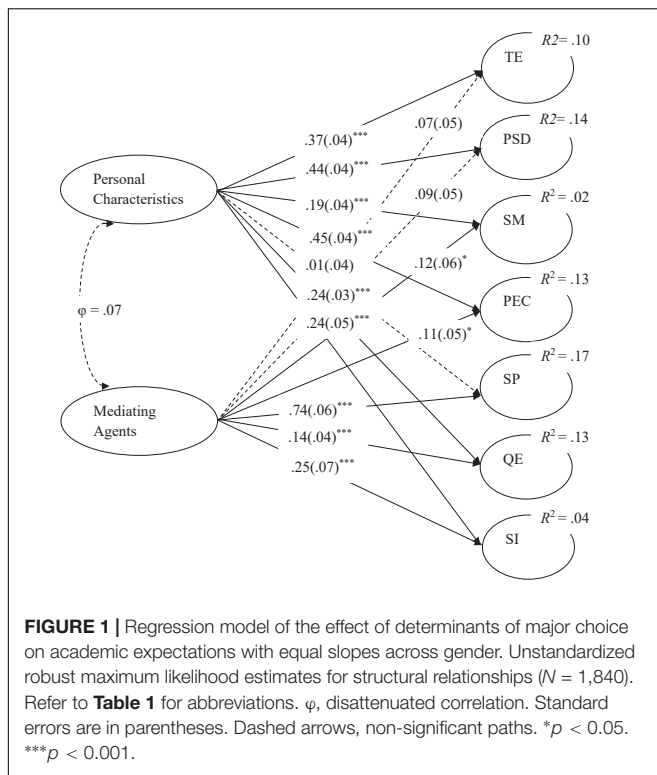
With this study, we intended to examine the predictive relationships between DMC and AEs of first-year university students across gender. To pursue this goal, we tested a multivariate regression model of the effect of PC and MA, as determinants of major choice (DMC), on APQ's 7-factors AEs with large samples of Spanish and Portuguese students.

First, we wanted to determine the model's possible equivalence across countries and gender and whether its factors would have adequate psychometric properties by country and gender. Similarly, in a two-step approach (Anderson and Gerbing, 1988; Jöreskog and Sörbom, 1993), before testing this regression model, we tested its oblique model counterpart, which was well fitted and denoted full invariance by country and gender, with all the nine factors presenting acceptable psychometric properties (Fornell and Larcker, 1981; Nunnally and Bernstein, 1994). In this process, we also observed that the two model predictors (PCs and MAs) were independent, which allows a clear interpretation of their relationships with the criteria (AEs). Concerning the gender differences in the factor means of the model predictors, we found that women attribute more importance to the PCs for a major choice than men, similar to Nadelson et al. (2013) results. By contrast, the importance of MAs for a major choice does not differentiate between women and men, contradicting the belief that women attribute more importance to MAs than men do (e.g., Päßler and Hell, 2012; Sojkin et al., 2012). This result can be a by-product of the lack of importance attributed by college students to external influences in their decision to pursue HE (Nadelson et al., 2013), reflected in the tendency for career choices in women to be increasingly

TABLE 5 | Regression model of the effect of determinants of major choice on academic expectations: invariance across gender.

Model	SB χ^2 (df)	RMSEA [90% CI]	SRMR		CFI
			Men (<i>n</i> = 793)	Women (<i>n</i> = 1,047)	
M1	7749.08 (2,045)	0.052 [0.050–0.053]	0.092	0.076	0.971
M2	7802.73 (2,259)	0.052 [0.050–0.053]	0.098	0.078	0.971
M1-M2					$\Delta CFI = 0.000$

Regression models were estimated and tested under full invariance at measurement and factor levels and with factor correlations of academic expectations freely estimated. M1, model with different slopes; M2, model with equal slopes. Refer to **Table 1** for other abbreviations.



dispersed and more similar to those of men (Montmarquette et al., 2002; Ma, 2009; Chang and ChangTzeng, 2020). Second, we conjectured that DMC would have a significant impact on the factors of expectations across gender. The invariance test of the multivariate regression model across gender indicated that it was well-fitted to data and revealed an equivalence of slopes (Jöreskog and Sörbom, 1993) between women and men. As a consequence, a unique interpretation of its predictive relationships for both genders is allowed. Traditionally and currently, numerous studies corroborate gender differences in Science, Technology, Engineering, and Mathematics (STEM) students related to career choice motives and students' educational expectations (Su and Rounds, 2015; Trusz, 2020). Nevertheless, we can speculate that this equivalence of slopes between genders occurs in our study because most of the sample is made up of students from degrees belonging to the social sciences and humanities (60%). On the other hand, it seems that family background and previous learning experiences, among others, play a much more important role in career choice and the formation of AEs than gender (Babarović, 2021). Moreover, educational institutions' increasing efforts to homogenize formative programs and methodologies minimize gender inequalities (Babarović, 2021; World Economic Forum, 2021).

The equivalent slopes for gender revealed statistically significant predictive relationships of PCs and MAs on six and five of the seven AEs, respectively. Consistent with other studies, we found that PCs predicted AEs related to PEC, PSD (Nyamwange, 2016; Eren, 2017; McLean et al., 2019), and TE (Akosah-Twumasi et al., 2018) more strongly than those related to QE, SI, and SM. Furthermore, PCs did not predict SP, implying

that accomplishing the expectations of others or finishing the degree within a given time frame seems to take a back seat to the primacy of students' self-interests (Brahm et al., 2017).

Along with evidence from some other studies, we also found that MAs' prediction of AEs, compared to PCs, was higher in SP (Guan et al., 2015) and SI (Akosah-Twumasi et al., 2018; Lorenz et al., 2020) and lower in PEC and SM. No relationship was found between TE and PSD. This may be because the content of the items of these EAs, as presented in the APQ, refer to personal evaluations of access to a job (good job, professional success) and PCs (autonomy, self-confidence). TE and PSD were better predicted by PCs, as in other studies (Navarro and Soler, 2014; Rodríguez-Muñiz et al., 2019).

Overall, we can state, as an educated guess, that students who choose HE studies according to their PCs form AEs based on the possibility of performing experiences and activities in that environment that give them the opportunity to help others to improve and enhance personal skills and achieve a successful professional future. On the other hand, students who base their choice of HE studies more on the influence of MAs expect to find experiences of social interaction in HE and to achieve an academic performance consistent with others' expectations.

Some theoretical implications can be derived from the findings of this study. First, as far as we know, there are no studies in which the DMC and AEs have been explored conjointly across gender. The multivariate regression model studied offers the opportunity to link these two key elements for academic persistence and success in HE (Bargmann et al., 2021; Eren & Rakıcıoğlu-Söylemez, 2021). It also allows interpretations in tune with Tinto (1993) attrition theory and with those made in the self-determination theory's studies on the importance of personal value-interest and expectations in career choice and student retention (Duchatelet and Donche, 2019; Eccles and Wigfield, 2020; Ryan and Deci, 2020; Schnettler et al., 2020).

Moreover, within our attempt to link career and academic psychological constructs, the finding that the DMCs' effect on AEs is equivalent across gender points to an indistinct gender role in the influence of personal and contextual variables of career choice on AEs among first-year university students.

On a practical level, educational institutions could adjust their intervention programs to favor the retention and permanence of students (Bergmark et al., 2018) who have chosen a career to some degree based on their PCs or MAs, knowing what each group expects from HE. Moreover, based on the effect size results, the majority of the effects of PCs and MAs on AEs were of *medium* size, indicating that these two DMCs are relevant in practical terms when intervening on AEs. We also note that, according to study results, this DMC relevance is independent of gender. However, it seems that no practical implications of PCs and MAs on SM and SI should be considered, as *small* effect sizes emerged for these relationships. First-year students vaguely represent these AEs because of their lack of knowledge of the new organizational structure and the opportunities to achieve them in their HE studies (Trautwein and Bosse, 2017). In contrast, we can conjecture that they are a kind of AEs that neither the students nor their significant social environment found relevant for the choice and development of an HE career.

These types of studies can help HE institutions to be more attentive to the students they welcome, to create the necessary conditions for them to have experiences matching their interests and expectations, identify and model adequately planned behaviors that promote academic adaptation and persistence (Dewberry and Jackson, 2018).

Despite the cautions to avoid threats to the study's internal validity, for example, model estimation and testing using the underlying bivariate normal approach (Jöreskog, 2005) with the robust Satorra and Bentler (1994) correction, and ensuring the model's psychometric assumptions, the external validity of the statistical conclusions presented is limited, given the non-probabilistic sampling used (Shadish et al., 2002). For example, the replication of the current study in other samples could shed light on the possible generalization of its findings. Also, it would be interesting to extend the obtained results by analyzing the invariance of the model according to the students' area or their major subject of study, or diverse cohorts of students, e.g., traditional vs. non-traditional students, or first-generation vs. continuing-generation students. We could also conduct longitudinal studies to explore how the model would behave considering the academic organization by semesters or academic years. These studies would help to better customize the interventions to the diversity of subpopulations of HE students.

Finally, this research is a contribution to the study of the significant relationships between the determinants of career choice and academic expectations, factors that are understood as

essential both for career development and student retention in HE. In addition, an integrated view of these variables is offered, considering others of a personal and contextual nature, such as gender and nationality.

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 Ethics Committee of the Ph.D. Program in Education and Behavioral Sciences, University of Vigo. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SA, AD, AC, and MG-S: literature review, material preparation, and results' interpretation. AC, MG-S, and SA: data collection. AD and SA: data analysis and writing of the manuscript. All authors read and approved the final manuscript.

REFERENCES

- Ahmed, K. A., Sharif, N., and Ahmad, N. (2017). Factors influencing students' career choices: empirical evidence from business students. *J. Southeast Asian Res.* 2017, 1–15. doi: 10.5171/2017.718849
- Akosah-Twumasi, P., Emeto, T. I., Lindsay, D., Tsey, K., and Malau-Aduli, B. S. (2018). A systematic review of factors that influence youths' career choices—the role of culture. *Front. Educ.* 3:58. doi: 10.3389/feduc.2018.00058
- Almeida, L. S., Deaño, M., Araújo, A. M., Diniz, A. M., Costa, A. R., Conde, A., et al. (2018). Equivalencia factorial de las versiones en español y portugués de un cuestionario de expectativas académicas [Factorial equivalence of the Spanish and Portuguese versions of a questionnaire of academic expectations]. *Rev. Latinoam. Psicol.* 50, 9–20. doi: 10.14349/rlp.2018.v50.n1.2
- Anders, J. (2017). The influence of socioeconomic status on changes in young people's expectations of applying to university. *Oxford Rev. Educ.* 43, 381–401. doi: 10.1080/03054985.2017.1329722
- Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychol. Bull.* 103, 411–423. doi: 10.1037/0033-2909.103.3.411
- Babarović, T. (2021). Development of STEM vocational interests during elementary and middle school: a cohort-sequential longitudinal study. *J. Career Dev.* online ahead of print. doi: 10.1177/08948453211036986
- Bargmann, C., Thiele, L., and Kauffeld, S. (2021). Motivation matters: predicting students' career decidedness and intention to drop out after the first year in higher education. *High. Educ.* online ahead of print. doi: 10.1007/s10734-021-00707-6
- Bergmark, U., Lundström, S., Manderstedt, L., and Palo, A. (2018). Why become a teacher? student teachers' perceptions of the teaching profession and motives for career choice. *Eur. J. Teach. Educ.* 41, 266–281. doi: 10.1080/02619768.2018.1448784
- Brahm, T., Jenert, T., and Wagner, D. (2017). The crucial first year: a longitudinal study of students' motivational development at a Swiss Business School. *High. Educ.* 73, 459–478.
- Braxton, J. M., Doyle, W. R., Hartley, H. V. III, Hirschy, A. S., Jones, W. A., and McLendon, M. K. (2014). *Rethinking College Student Retention*. San Francisco, CA: Jossey-Bass.
- Cabrera, A. F., Nora, A., and Castaneda, M. B. (1993). College persistence — structural equations modeling test of an integrated model of student retention. *J. High. Educ.* 64, 123–139. doi: 10.1080/00221546.1993.11778419
- Castillo-Sánchez, M., Gamboa-Araya, R., and Hidalgo-Mora, R. (2020). Factores que influyen en la deserción y reprobación de estudiantes de un curso universitario de matemáticas [Factors influencing student dropout and failure in a university mathematics course]. *Uniciencia* 34, 219–245. doi: 10.15359/ru.34-1.13
- Chang, D. F., and ChangTzeng, H. C. (2020). Patterns of gender parity in the humanities and STEM programs: the trajectory under the expanded higher education system. *Stud. High. Educ.* 45, 1108–1120. doi: 10.1080/03075079.2018.1550479
- 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
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale: Erlbaum.
- Deaño, M., Diniz, A., Almeida, L., Alfonso, S., Costa, A., García-Señorán, M., et al. (2015). Propiedades psicométricas del cuestionario de percepciones académicas para la evaluación de las expectativas de los estudiantes de primer año en enseñanza superior [Psychometric properties of the Academic Perceptions Questionnaire to assess Higher Education first-year students' expectations]. *An. Psicol.* 31, 280–289. doi: 10.6018/analesps.31.1.161641

- Dewberry, C., and Jackson, D. J. R. (2018). An application of the theory of planned behavior to student retention. *J. Vocat. Behav.* 107, 100–110. doi: 10.1016/j.jvb.2018.03.005
- Dias, D. (2011). Reasons and motivations for the option of an engineering career in Portugal. *Eur. J. Eng. Educ.* 36, 367–376. doi: 10.1080/03043797.2011.593096
- Diniz, A. M. (2008). “Escala de determinantes da escolha de curso (EDEC),” [Determinants of Major Choice Scale], in *Avaliação Psicológica: Instrumentos Validados Para a População Portuguesa*, Coords, eds L. S. Almeida, M. R. Simões, C. Machado, and M. M. Gonçalves (Coimbra: Quarteto), 75–93.
- Diniz, A. M., Alfonso, S., Araújo, A. M., Deaño, M., Costa, A. R., Conde, Á, et al. (2018). Gender differences in first-year college students’ academic expectations. *Stud. High. Educ.* 43, 689–701. doi: 10.1080/03075079.2016.1196350
- Diniz, A. M., and Almeida, L. S. (2007). Variáveis da primeira fase da integração universitária e mudança de curso [Variables of the first phase of university integration and course change]. *Psicol. Esc. Educ.* 11, 201–210.
- Duchatelet, D., and Donche, V. (2019). Fostering self-efficacy and self-regulation in higher education: a matter of autonomy support or academic motivation? *High. Educ. Res. Dev.* 38, 733–747. doi: 10.1080/07294360.2019.1581143
- Eccles, J. S., and Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: a developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* 61:e101859. doi: 10.1016/j.cedpsych.2020.101859
- Eren, A. (2017). Professional aspirations among pre-service teachers: personal responsibility, time perspectives, and career choice satisfaction. *Aust. Educ. Res.* 44, 275–297. doi: 10.1007/s13384-017-0234-y
- Eren, A., and Eren & Rakıcıoğlu-Söylemez, A. (2021). Pre-service teachers’ professional commitment, sense of efficacy, and perceptions of unethical teacher behaviours. *Aust. Educ. Res.* 48, 337–357. doi: 10.1007/s13384-020-00396-7
- Ferrando, P. J. (1996). Calibration of invariant item parameters in a continuous item response model using the extended LISREL measurement submodel. *Multivariate Behav. Res.* 31, 419–439. doi: 10.1207/s15327906mbr3104_2
- Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50.
- Fouad, N. A., Hackett, G., Smith, P. L., Kantamneni, N., Fitzpatrick, M., Haag, S., et al. (2010). Barriers and supports for continuing in mathematics and science: gender and educational level differences. *J. Vocat. Behav.* 77, 361–373.
- Gottfredson, L. S. (2002). “Gottfredson’s theory of circumscription, compromise, and self-creation,” in *Career Choice and Development*, ed. D. Brown (New York, NY: John Wiley & Sons), 85–148.
- Guan, Y., Chen, S. X., Levin, N., Bond, M. H., Luo, N., Xu, J., et al. (2015). Differences in career decision-making profiles between American and Chinese university students: the relative strength of mediating mechanisms across cultures. *J. Cross-Cult. Psychol.* 46, 856–872. doi: 10.1177/0022022115585874
- Hu, L., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparametrized model misspecification. *Psychol. Methods* 3, 424–453.
- Jöreskog, K. G. (2005). *Structural Equation Modeling with Ordinal Variables using LISREL*. https://ssicentral.com/wp-content/uploads/2021/04/lis_ordinal.pdf.
- Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International.
- Jöreskog, K. G., and Sörbom, D. (1996). *PRELIS 2: User’s Reference Guide*. Chicago, IL: Scientific Software International.
- Jöreskog, K. G., and Sörbom, D. (2006). *LISREL 8.80 for Windows [Computer software]*. Chicago, IL: Scientific Software International.
- Kuh, G. D., Gonyea, R. M., and Williams, J. M. (2005). “What students expect from college and what they get,” in *Promoting Reasonable Expectations: Aligning Student and Institutional Views of the College Experience*, eds T. E. Miller, B. E. Bender, and J. H. Schuh (San Francisco, CA: Jossey-Bass), 34–64.
- Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122. doi: 10.1006/jvb.1994.1027
- Lerkkanen, M.-K., Kiuru, N., Pakarinen, E., Viljaranta, J., Poikkeus, A.-M., Rasku-Puttonen, H., et al. (2012). The role of teaching practices in the development of children’s interest in reading and mathematics in kindergarten. *Contemp. Educ. Psychol.* 37, 266–279. doi: 10.1016/j.cedpsych.2011.03.004
- Lorenz, G., Boda, Z., Salikutluk, Z., and Jansen, M. (2020). Social influence or selection? peer effects on the development of adolescents’ educational expectations in Germany. *Br. J. Sociol. Educ.* 41:643669. doi: 10.1080/01425692.2020.1763163
- Ma, Y. (2009). Family socioeconomic status, parental involvement, and college major choices: gender, race, ethnic, and nativity patterns. *Sociol. Perspect.* 52, 211–234.
- McLean, L., Taylor, M., and Jimenez, M. (2019). Career choice motivations in teacher training as predictors of burnout and career optimism in the first year of teaching. *Teach. Teach. Educ.* 85, 204–214. doi: 10.1016/j.tate.2019.06.020
- Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors? *Econ. Educ. Rev.* 21, 543–556. doi: 10.3368/jhr.48.3.545
- Morgan, S. L. (2005). *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford, CA: Stanford University Press.
- Nadelson, L. S., Semmelroth, C., Martinez, G., Featherstone, M., Fuhrman, C. A., and Sell, A. (2013). Why did they come here? The influences and expectations of first-year students’ college experience. *High. Educ. Stud.* 3, 50–62. doi: 10.5539/hes.v3n1p50
- Navarro, P., and Soler, I. (2014). Las motivaciones de la elección de carrera por los estudiantes universitarios [Motivations for university students’ career choices]. *Rev. Asoc. Sociol. Educ.* 7, 61–81.
- Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory*, 3rd Edn. New York, NY: McGraw-Hill.
- Nyamwange, J. (2016). Influence of students’ interest on career choice among first year university students in public and private universities in Kisii County, Kenya. *J. Educ. Pract.* 7, 96–102.
- Parker, P. D., Schoon, I., Tsai, Y. M., Nagy, G., Trautwein, U., and Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Dev. Psychol.* 48, 1629–1642. doi: 10.1037/a0029167
- Päßler, K., and Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *J. Career Assess.* 20, 479–496. doi: 10.1177/1069072712450009
- Rodríguez-Muñiz, L., Areces, D., Suárez-Álvarez, J., Cueli, M., and Muñiz, J. (2019). Qué motivos tienen los estudiantes de Bachillerato para elegir una carrera universitaria? [What are the reasons for the choice of a university degree for a- level students?]. *Rev. Psicol. Educ.* 14, 1–15. doi: 10.23923/rpye2019.01.167
- 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:e101860. doi: 10.1016/j.cedpsych.2020.101860
- Satorra, A., and Bentler, P. M. (1994). “Corrections to test statistics and standard errors in covariance structure analysis,” in *Latent Variable Analysis*, eds A. von Eye and C. C. Clogg (Thousand Oaks, CA: Sage), 399–419.
- Schnettler, T., Bobe, J., Scheunemann, A., Fries, S., and Grunschel, C. (2020). Is it still worth it? applying expectancy-value theory to investigate the intraindividual motivational process of forming intentions to drop out from university. *Motiv. Emot.* 44, 491–507. doi: 10.1007/s11031-020-09822-w
- Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for generalized Causal Inference*. Boston, MA: Houghton Mifflin.
- Sojkin, B., Bartkowiak, P., and Skuza, A. (2012). Determinants of higher education choices and student satisfaction: the case of Poland. *High. Educ.* 63, 565–581. doi: 10.1007/s10734-011-9459-2
- Su, R., and Rounds, J. (2015). All STEM fields are not created equal: people and things interests explain gender disparities across STEM fields. *Front. Psychol.* 6:189. doi: 10.3389/fpsyg.2015.00189
- Tinto, V. (1993). in *Leaving College: Rethinking the Causes and Cures of Student Attrition*, 2nd Edn, ed. I. L. Chicago (Chicago University Press).
- Trautwein, C., and Bosse, E. (2017). The first year in higher education: critical requirements from the student perspective. *High. Educ.* 73, 371–387.
- Trusz, S. (2020). Why do females choose to study humanities or social sciences, while males prefer technology or science? some intrapersonal and interpersonal predictors. *Soc. Psychol. Educ.* 23, 615–639. doi: 10.1007/s11218-020-09551-5
- Whiston, S. C., and Keller, B. K. (2004). The influences of the family of origin on career development: a review and analysis. *Couns. Psychol.* 32, 493–568.
- World Economic Forum (2021). *The Global Gender Gap Report 2021*. Cologny: World Economic Forum.

- Zacher, H., Rudolph, C. W., Todorovic, T., and Ammann, D. (2019). Academic career development: a review and research agenda. *J. Vocat. Behav.* 110, 357–373. doi: 10.1016/j.jvb.2018.08.006
- Zafar, B. (2013). College major choice and the gender gap. *J. Hum. Resour.* 48, 545–595. doi: 10.3368/jhr.48.3.545
- Zimmermann, T. (2020). Social influence or rational choice? two models and their contribution to explaining class differentials in student educational aspirations. *Eur. Sociol. Rev.* 36, 65–81. doi: 10.1093/esr/jcz054

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 Alfonso, Diniz, Conde and García-Señorán. 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.



Epistemic Beliefs and Learners' Self-Efficacy as Predictors of Language Learning Strategies: Toward Testing a Model

Shaghayegh Shirzad¹, Hamed Barjesteh^{1*}, Mahmood Dehqan² and Mahboubeh Zare¹

¹ Department of English Language and Literature, Islamic Azad University, Ayatollah Amoli Branch, Amol, Iran, ² Department of English Language and Literature, University of Mazandaran, Babolsar, Iran

OPEN ACCESS

Edited by:

Lawrence Jun Zhang,
University of Auckland, New Zealand

Reviewed by:

Lili Tian,
Renmin University of China, China
Lili Qin,
Dalian University of Foreign
Languages, China

*Correspondence:

Hamed Barjesteh
ha_bar77@yahoo.com

Specialty section:

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

Received: 01 February 2022

Accepted: 11 March 2022

Published: 12 April 2022

Citation:

Shirzad S, Barjesteh H, Dehqan M
and Zare M (2022) Epistemic Beliefs
and Learners' Self-Efficacy as
Predictors of Language Learning
Strategies: Toward Testing a Model.
Front. Psychol. 13:867560.
doi: 10.3389/fpsyg.2022.867560

Understanding the beliefs held by the learners about learning a language, and the way they utilize their thoughts about knowledge and learning seem essential for planning a constructive language program. Following this line of research, this paper aims at testing a hypothetical model of the relationship between epistemic beliefs (EBs) and subscales of language-learning strategies (LLSs) through the mediating role of learners' self-efficacy (LSE). To this end, a sample of 300 Iranian high school students, taking regular courses, completed three survey questionnaires. At this stage, correlational analysis and structural equation modeling (SEM) were employed to probe the interconnections, analyze the model, and outline the conceptual framework. The results revealed that the LSE framework can adequately account for the learners' LLSs. In particular, the results indicated that efforts, persistence, and imitation (i.e., the subfactors of LSE) positively and significantly influenced LLSs. However, EBs with the mediating role of LSE were known to be a significant factor in demoting the LLSs. Notably, knowledge and learning agents were the negative predictors of LLSs. This paper suggests that LSE has higher explanatory power than EBs in predicting LLSs. The findings of this study suggest that teachers and material developers should pay serious attention to the learners' self-efficacy as they were known to influence LLSs.

Keywords: EFL learners, epistemic beliefs, language learning strategies, learners' self-efficacy, structural equation modeling

INTRODUCTION

Drawing on the second language (L2) professional literature indicates a paradigm shift from a cognitive and process-oriented approach to a beyond method-based pedagogy. Kumaravadivelu's (2006) paradigm in education potentially invites the learners with knowledge, beliefs, attitude, and autonomy necessary to foster their language learning. In line with Kumaravadivelu, voluminous studies (e.g., Hofer, 2016; Griffiths, 2018; Chamot, 2019; Lindner and Retelsdorf, 2019; Shirzad et al., 2020) have stipulated that learning conception, learners' beliefs, thinking about the essence of knowledge, and learning are connected to language learning. Recently, different studies (e.g., Morris et al., 2017; Liu et al., 2019; Takeuchi, 2019; Cheng, 2020; Mercer and Dörnyei, 2020; Razmi and Jabbari, 2021) have released evidence that learners' beliefs influence the academic achievement. They have provided strong evidence to support the predictive effect of learners' beliefs in learning

achievement and course satisfaction. The findings also indicated that there is a positive interplay between different dimensions of EBs with different fields in sociology, psychology, and education. In education, various constructs were significantly reported to connect EBs such as epistemological theories (Hofer and Pintrich, 1997), language achievement (Winberg et al., 2019), self-concepts, learning conception (Liu et al., 2019), self-efficacy and assessment (Cheng, 2020), personal beliefs (Mardiha and Alibakhshi, 2020), LLSs and motivational self-system (Shirzad et al., 2020), and perfectionism (Razmi and Jabbari, 2021) to name but a few. Moreover, the EBs have been explored in various correlational studies (e.g., Chan and Elliott, 2004; Liu et al., 2019; Winberg et al., 2019; Mardiha and Alibakhshi, 2020; Shirzad et al., 2020; Kärchner et al., 2021; Zhu et al., 2021). The findings of such empirical studies indicated that EBs correlated with different variables such as stability, contingency of self-esteem, academic achievement, regulatory focus, learning engagement, conceptions of teaching and learning, and L2 motivational self-system. The results substantiated that learners' EBs have predictive power in education, in general, and learning conception. Besides, some other studies (e.g., Schommer, 1990; Hofer and Pintrich, 1997; Hofer, 2016) endorsed that learners with a high level of EBs seemed to act differently in various aspects of language learning and learning conception.

As a complex multidimensional trait, EBs (i.e., views about the quality of knowledge and learning), and LSE (i.e., the tendency for initiating tasks, investing adequate effort to conduct activities, endurance and perseverance in facing difficulties) are among the important affective factors in educational psychology (Hofer and Pintrich, 1997; Deuling and Burns, 2017; Shirzad et al., 2020; Razmi and Jabbari, 2021). An individual's EBs depict the conceptions of his/her delineations of scientific knowledge and what it denotes. Greene et al. (2016) used the term *epistemic cognition* to imply how students gain, apprehend, justify, and utilize knowledge. They postulated that learners involve in epistemic cognition when they arouse self-beliefs about the essence of knowledge and knowing (i.e., epistemic beliefs). Hofer and Pintrich (1997) postulated that EBs are indispensable for learning conception and understanding within several domains and contexts. Another construct of the current study is self-efficacy beliefs. Bandura (1997) conceptualized *self-efficacy beliefs* as an "individual's belief in his or her own ability to organize and implement action to produce the desired achievements and results" (p. 3). It is classified as a *general* or a *specific belief*. The former concerns a general perceived ability to confront stressful conditions, while the latter deals with a particular context or situation (Bandura, 1997). This study concentrates on specific self-efficacy beliefs related to the academic field described as an individual's perceived abilities to manage various instructional areas and learning conceptions. For the current study, EBs and LSE have been used as independent constructs to predict LLSs. Various theoretical studies (e.g., Oxford, 2017; Cohen, 2018; Griffiths, 2018; Chamot, 2019) pinpointed that learning strategies are an affiliative factor in promoting language achievement. Besides, the findings of some empirical studies (e.g., Oxford, 2017; Habók and Magyar, 2018; Takeuchi, 2019; Shirzad et al., 2020; Razmi and Jabbari, 2021; Tang, 2022) corroborated that

LLSs can help the learners apply their knowledge in a real-world context, gain knowledge, and achieve higher academic results eventually. Accordingly, different taxonomies of LLSs were proposed by the authorities in educational psychology (e.g., Oxford's direct and indirect strategies; O'Malley and Chamot's socio-affective strategies; Cohen's L2 learning and use strategies). They substantiated the notion that utilizing LLSs can influence the quality of knowledge and learning (i.e., EBs) and different psychological constructs such as self-beliefs, self-concept, self-efficacy, to name but a few.

Despite the enriched literature (e.g., Hofer, 2016; Yang et al., 2019; Cheng, 2020; Kärchner et al., 2021; Lonka et al., 2021; Razmi and Jabbari, 2021; Zhu et al., 2021) on the predictive power of self-efficacy beliefs and their influence on the learning conception, the effects of EBs and LSE on the students' learning strategies are not yet clear. Accordingly, there has been no credible empirical study to support the conceptual interplay between EBs and LLSs with the mediating role of LSE. Therefore, it seems important to test a model to uncover how the learners' knowledge of EBs and their specific self-efficacy in the academic setting may predict the LLSs language learners employ in the learning process. Notably, it is significant to explore whether the dimensions of EBs promote or demote the learning strategies students utilize for language learning. Accordingly, it has been hypothesized that EBs and LSE promote students' learning strategy which in turn may foster their academic achievement. Moreover, it has been hypothesized that students' EBs and LSE positively predict high school students' LLSs. Despite sufficient evidence to support the positive effect of learners' beliefs, this paper claims that the target variables (i.e., EBs, LLSs, and LSE) have a complex and unpredictable relationship. Therefore, this study hypothesized a model based on the learners' beliefs and their sense of efficacy as the predictors of LLSs. Notably, this study was guided by the following objectives:

- (i) To identify the relationship among the students' EBs, LLSs, and LSE.
- (ii) To determine whether EBs with the mediating role of LSE positively predict the high school students' LLSs.
- (iii) To explore if LSE positively predicts the high school students' LLSs.

LITERATURE REVIEW

Theoretical Framework

The professional authorities in Epistemology (e.g., Perry, 1970; Schommer, 1990; Magolda, 1992; Hofer and Pintrich, 1997; Hofer, 2001) classified two aspects for the EBs studies. They distinguished both *developmental* and *multidimensional* facets. The first aspect (i.e., developmental) is one-dimensional. Therefore, learners move through a cycle of developmental phases (i.e., from the objectivist view of knowledge to uncertainty of knowledge and then to extreme subjectivity). Notably, individuals move through the successive stages like fashion, and they progress in a patterned sequence of developmental stages. Hofer (2001) assumed that the thinking in this model opens with the objectivist perspective of knowledge to extreme

subjectivity. Perry's model 1970 and Magolda's (1992) model of EBs are instances of the developmental model. Hofer (1994) postulated the notion of EBs as developing step by step in a linear fashion from *dualism*, *multiplism*, *relativism*, and finally *commitment*. Besides, Baxter Magolda (1992) conceptualized EBs in four stages: (a) *absolute*, (b) *transitional*, (c) *independent*, and (d) *contextual knowing*. In contrast to the developmental model, Schommer (1990) suggested a multidimensional system. Schommer claimed that EBs approximately exist in *independent beliefs*. This delineates that one dimension may be developed but another aspect may be quite immature. Notably, there are different beliefs that may (not) develop coincidentally. Schommer's model included five dimensions (e.g., *stability*, *structure*, *source*, *speed of acquisition*, and *control of acquisition*). Hofer and Pintrich (1997) criticized Schommer's model in that the model cares about the nature of learning and not the nature of knowledge and knowing. Thus, they proposed four dimensions including certainty (stability), simplicity (structure), source of knowing (authority), and judgment for knowing (evaluation of knowledge). This study, therefore, delimited its scope on the learners' EBs in the educational domain. Accordingly, Schommer's model was adopted for the study because the scope of this paper was concentrated on the nature of learning and the academic setting.

The Theoretical Connections Among EBs, LSE, and LLSs

A growing body of theoretical and empirical studies (e.g., Morris et al., 2017; Takeuchi, 2019; Cheng, 2020; Razmi and Jabbari, 2021) corroborated that the learners' beliefs and their conception of learning influence language-learning behavior. Various theoretical studies (e.g., Perry, 1970; Schommer, 1990; Hofer and Pintrich, 1997; Bandura, 2006) pinpointed that individuals' EBs and LSE play pivotal roles in the learning process. As individuals with various beliefs may adopt different learning strategies, it seems that the constructs (i.e., EBs and LSE) may affect the learning process. Therefore, students with different levels of EBs and self-efficacy may act differently in language learning. Accordingly, they may adopt various learning strategies as far as their levels of EBs and self-efficacy are concerned. Despite consensus on the implications of the beliefs held by the learners, the way they may influence language learning raised doubts among the practitioners. Accordingly, various taxonomies for LLSs (e.g., O'Malley and Chamot; Cohen; Oxford) and scales for EBs and LSE (e.g., Clarebout et al., 2001; Chan and Elliott, 2004; Rezaei, 2010) have been proposed by the authorities to conceptualize the way the targeted variables influence language learning for different cultural situations. Recently, some empirical studies have gained attention on the learners' internal factors in the learning process. They underscored the connection between LLSs in the light of voluminous affective factors like learners' self-efficacy (Cheng, 2020), learners' beliefs (Winberg et al., 2019), L2 motivational self-system (Shirzad et al., 2020), self-control depletion (Lindner and Retelsdorf, 2019), regulatory focus, and learning engagement (Liu et al., 2019). Such empirical studies developed the perspective about learning strategies, thinking

process in learning, and internal forces in education. Notably, exploring the interplay among EBs, LSE, and LLSs and the way EBs and LSE may influence language learning can foster significant pedagogical implications. Such findings formed the theoretical underpinning of this study.

Epistemological Beliefs

The term *epistemology* is an area in psychology that deals with reasoning, the essence of knowledge, and the ideas about knowledge (Hofer and Pintrich, 1997). Hofer (2016) presumed that studies in EBs concentrate on the way learners come to know, and the way students employ their thoughts about knowledge and know how to conceptualize their environment. Winberg et al. (2019) conceptualized EBs as the beliefs about the essence of knowledge, learning, and knowing. Despite the lack of agreed-upon implementation of EBs, some authorities (e.g., Hofer, 2016; Winberg et al., 2019) used a multilayered stage. In line with the different layers, Shirzad et al. (2020) introduced some terminologies (e.g., *epistemic cognition*, *epistemic cognition*, *epistemological resources*, *epistemological reflection*, *personal epistemologies*, *reflective judgment*) to refer to EBs in the L2 professional literature.

In line with Perry's (1970) *dualism* model, Schommer (1990) suggested different beliefs about the origin of knowledge. She maintained that authorities manipulate various aspects of the beliefs. To Schommer, the structure of knowledge is an isolated rather than interrelated fact. Schommer distinguished different dimensions. The first dimension was established as *simple knowledge* (i.e., isolated facts) in the L2 professional literature. The second aspect (i.e., the certainty of knowledge) considers knowledge as an *absolute* (i.e., certain) construct. Schommer (1990) coined the term *omniscient authority* for the certainty of knowledge. The third aspect (i.e., speed of acquisition) pinpoints learning as a prompt vs. a gradual process. Finally, the *control of acquisition* refers to the ability to learn as natural vs. being acquired. To Schommer, it is an *innate ability* where learners believe learning cannot be enhanced with instruction. Concerning the multidimensional nature of EBs, Hofer (2016) elucidated that EBs deal with various constructs such as the *source*, *justification*, *certainty*, and the *development of knowledge*. Hofer assumed that the *source of knowledge* is at the level of less complex beliefs that originate beyond the self and occupies an exterior authority. At more complex beliefs, it is made by the knower in interaction with the peer. The term *justification of knowledge* deals with the way individuals account for knowledge. At the lower levels, they employ authority or observation rather than experiments, data, and inquiry rules. The *certainty of knowledge* is the belief about the validity of knowledge ranging from a belief in a correct answer to complicated problems. Finally, the *development of knowledge* concerns knowledge progress. It considers science as an evolving subject.

Self-Efficacy Theory

Bandura's (1986) *self-efficacy theory* illustrates a picture of the learners' activity in which they are neither unlikely controlled by external factors nor automatically shaped by their genetic faculty. Bandura assumed that self-efficacy is an assumption

in one's ability to conquer essential life events. In his theory, Bandura (1997) defined LSE as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance" (p. 174). Notably, various cognitive, affective, and biological forces are reciprocally influential SET. Bandura proposed five human abilities (i.e., *symbolizing capabilities, forethought, self-regulatory and self-reflective potential*) as the cornerstone of social cognition. Thus, LSE affects *what learners select to do*, their level of *endeavor*, *persistence* in case of problems, and *subsequent performance* (Sherer et al., 1982). Later, Bandura (2006) conceived that learners have a self-system that helps them to manipulate their emotions, feelings, and actions. In line with this claim, Morris et al. (2017) concluded that self-reflective capability provides learners with the capacity to think and to influence their future behavior. Likewise, Cheng (2020) conceptualized the notion of LSE as the belief that individuals can optimize their learning performance by their psychological attempts besides the scaffolding received by their peers and teachers in the educational settings. Specifically, LSE serves as a *self-regulatory* function that affects the learners' cognition and actions (Liu et al., 2019).

Pajares (2007) expressed that LSE is regarded as an anticipation process within self-regulation models. Pajares believed that LSE is a personal and social construct because learners act both collectively and individually. He maintained that LSE has a proactive effect on performance and self-evaluative operations along with performance. He outlined three distinctive self-efficacy features: First, self-efficacy focuses on *perceived competence* to execute a task rather than on psychological traits. Second, LSE beliefs are *task-, domain-, and context-specific*. Third, LSE depends on the *mastery criterion* of performance instead of the normative criteria. Finally, LSE beliefs are mainly evaluated before engaging in a *specific task or activity*. Spratt et al. (2002) released a motivational construct for the term LSE which predisposes learners to autonomous behavior. In this line, Lindner and Retelsdorf (2019) postulated that LSE is a dynamic and accomplished belief system that alters in diverse tasks and situations. Bandura (1997) highlighted the impact of LSE beliefs in individual performance. Bandura maintained that "people's level of motivation, affective states, and actions are based more on what they believe than on what is objectively true" (p. 2). Therefore, the way the learners act can be anticipated by the conceptions they hold about their abilities than by what they are capable of performing. Schunk and Zimmerman (2007) believed that such a construct encompasses different facets such as *level, generality, and strength*. More precisely, the former concerns the difficulty level of a task. The latter relates the transferability of the learners' efficacy judgments on various activities such as different academic subjects.

Dimensions and Sources of Academic Self-Efficacy in the Learning Context

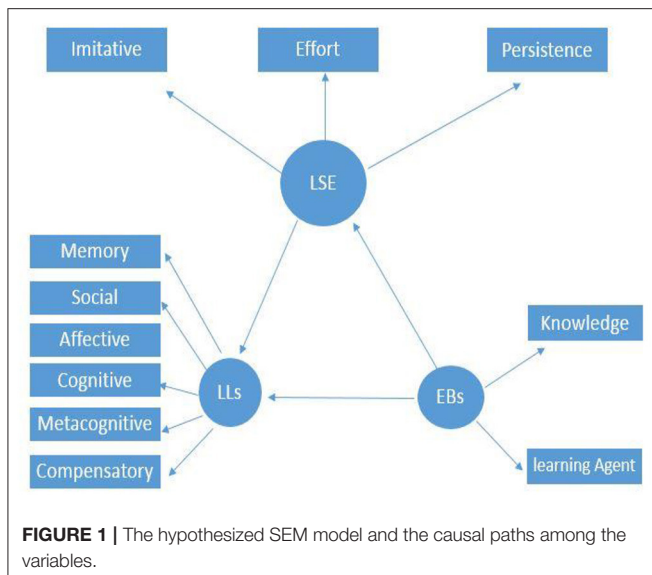
Bandura (1986) postulated that the learners assess their efficacy by analyzing information from various aspects such as *mastery experience, vicarious experience, social persuasion, and physiological and affective states*. Bandura (1986, 1997) called this sort of attainment "*performance accomplishments*"

and "*enactive attainments*". Pajares (2007) concluded that mastery experiences comprised the attainment of goals (i.e., accomplishment/attainment) *via* direct and personal action (i.e., enactive). *Vicarious experience* is deeply rooted in the social model. Bandura (1986) believed that this source of self-efficacy is influential for the development of LSE for the novel task. Bandura called this model a *coping model* that openly struggles to overcome obstacles (Cheng, 2020). Later, Bandura (2006) highlighted the role of *evaluative feedback*. To him, it is a form of social persuasion that is often conciliated by perceived knowledge. Bandura maintained that the learners' self-beliefs may be more harmed by disappointing messages than influenced by positive conviction. He highlighted the roles of the physiological and affective states as the leading sources of LSE. Cheng (2020) stated that the notion of LSE concerns various self-beliefs, such as self-esteem, self-regulation, self-concept, and self-control. However, LSE is different from other kinds of self-beliefs (i.e., self-concept, self-esteem, and self-efficacy) that are mistakenly used interchangeably.

Bandura (1986) defined self-concept as a generalized self-assessment comprising different self-feedbacks such as feelings of self-worth and general competence beliefs. On the other hand, self-efficacy beliefs are context-specific judgments of individual capacity to manipulate courses of action to achieve a specific objective (Liu et al., 2019). LSE focuses on the tasks and activities one can perform than more global assessments of self-concept. Therefore, LSE promotes academic performance both directly and indirectly by its role on a learner's self-concept (Cheng, 2020). Pajares (2007) supported that self-esteem refers to the assessment of self-worth, which relies on the way culture values the characteristics one possesses and how well one's behavior corresponds to the standards of worthiness. Kärchner et al. (2021) called self-esteem a personal judgment of worthiness. Different practitioners conceptualized the construct as a subjective experience with general, situational, and task levels to capture its multifaceted aspects.

Hypothesized Model

Drawing on previous theoretical frameworks, empirical studies, and justifying the connection among the targeted variables, this study proposed a structural model to determine the multivariate relations. The provided empirical evidence for the universal beneficial effects of EBs on academic achievement (Hofer, 2016; Peffer and Ramezani, 2019; Ongowo, 2021), and the supportive literature on the relationship between LSE and LLSs (Bandura, 1988; Pajares, 2007; Morris et al., 2017; Cheng, 2020) resulted in drawing a hypothesized path from the EBs to LSE and LSE to LLSs. For the current study, EBs and LSE are considered as independent variables, and the LLSs are regarded as dependent variables. To map the conceptual framework, a path diagram was generated based on the theoretical underpinnings to conceptualize a hypothetical model. At the theoretical phase, three constructs, measured by 11 observed variables, formulated the proposed model. Specifically, the hypothesized model predicts the path and the interconnections between EBs and LLSs with the mediating role of LSE. Following Fornell and Larcker's (1981) guidelines for generating a path diagram,



latent (i.e., circles/ovals) and observed (rectangles) variables are illustrated in **Figure 1**.

Based on the hypothesized model and the interconnections proposed in the literature from the theoretical and empirical aspects among the variables, the following research questions were addressed.

RQ1: Is there any significant relationship among EBs, LLs, and LSE?

RQ2: Do the learners' self-efficacy have a significant direct effect on their learning strategy?

RQ3: Do the students' epistemological beliefs have a significant direct effect on their learning strategy?

RQ4: Do the epistemological beliefs with the mediating role of the learners' self-efficacy have a significant indirect effect on LLs among high school students?

METHODS

Participants

To address the objectives of the current research, 300 Iranian high school students from a cluster of the entire population ($n = 500$) of Amol and Babol (i.e., two cities of Mazandaran) were recruited as the participants of this study. They were native Persian speakers who were both males ($n = 123$) and females ($n = 177$) with a similar language, social and cultural background. They were placed at the pre-intermediate level having 4 years of experience in learning English at different language institutes. To select a more representative sample, a cluster random sampling method was employed (Ary et al., 2018). The sampling multistage included two cities, five districts, and 12 state high schools. To control the bias effect, the respondents were randomly selected from two genders with different age ranges. Their ages ranged from 16 to 18 ($M = 17$, $SD = 1.7$).

Instrumentation

To collect the data, three scales were utilized regarding the target variables. Two questionnaires were translated into Persian and then back-translated by an expert translator to ensure the accuracy of the translation. To probe the translated versions, the original scales were examined by another expert holding a Ph.D. in Applied Linguistics. It was done to ensure the comprehensibility of the item, translation accuracy, and to check any ambiguities in comprehending the message. Next, the internal consistency was examined and reported. In addition, all the scales were piloted in a similar context. Specifically, to test the reliability of the scale within the EFL context of Iran, a pilot study was conducted among 120 pre-intermediated students learning English at three private English language institutes. The results enjoyed adequate reliability ($\alpha = 0.87$).

Epistemological Beliefs Questionnaire

Razmi and Jabbari (2021) EBQ has been regularly administered as a well-known scale to appraise EBs. Originally, it comprised 63 items to be completed by the respondents. However, the appropriateness of this scale has been criticized for being long with confusing items. Accordingly, different adapted versions were validated in a setting different from the original one. For this study, an adapted version of the EBQ has been validated by Rezaei (2010). Rezaei examined the reliability and factor analysis of revised EBQ among 518 Iranian students studying different majors. To test the construct validity, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were run. The EBQ scale enjoyed high reliability ($\alpha = 80.5$) and validity indices. The revised EBQ comprised 16 items in either the negative or positive extreme. It aimed to measure the knowledge ($n = 9$ items) and learning agent ($n = 7$ items). The *first* dimension aimed to assess the respondents' assumption about the nature of knowledge (e.g., *If scientists try hard enough, they can find the truth to almost anything; wisdom is not knowing the answers, but knowing how to find the answers*). The second dimension concerned the learners' assumptions about acquisition/learning (e.g., *Learning something well takes a long time or much effort; How much you get from your learning depends mostly on your effort*). Students were asked to rate the statements on a five-point Likert scale from 1 (strongly disagree), anchoring the right end to 5 (strongly agree) anchoring the left end. In addition, the revised version was piloted for this study. The questionnaire was distributed among similar subjects (e.g., 100 junior high school students) in Amol and Babol, Iran. Some of them also joined the follow-up interviews to ensure the comprehension of all items. The scale enjoyed the total reliability coefficient ($\alpha = 0.77$). The reliability coefficient for the subscale was as follows: simple/definitive knowledge ($\alpha = 0.78$) and fast/fixed learning agent ($\alpha = 0.76$).

General Self-Efficacy Questionnaire

The original GSEQ was developed by Sherer et al. (1982). It comprised 23 items with a construct of two factors (i.e., general and social self-efficacy). The general and social self-efficacy explained 26.5 and 8.5% of the variance. The alpha coefficient

TABLE 1 | Cronbach's alpha coefficients for GSEQ.

Sub-factor	Items	Alpha
Initiative	9	0.84
Persistence	5	0.82
Effort	3	0.73
Total	17	0.796

for the items is 0.86 and 0.70, suggesting that the items have a relatively high internal consistency. The GSEQ was arranged in a five-point Likert scale with 1 (not at all true) anchoring the left end and 5 (exactly true) anchoring the right end. The revised version comprised a construct of three factors: Initiative (9 items), Persistence (5 items), and effort (3 items). The GSEQ score ranged from minimum (17) to maximum (85). The score of the following items: 1, 13, 8, 9, 3, and 15, will increase from the right to the left, and the rest is *vice versa*. A higher score indicates a higher level of self-efficacy. To examine the internal consistency of the GSEQ in the EFL context of Iran, the questionnaire was piloted among high school students ($n = 210$) who were randomly selected from eight junior high schools. The estimated reliability was found to be ($\alpha = 0.796$). **Table 1** reports the reliability coefficient for each subscale.

As indicated in the table, all the subfactors (i.e., initiative, $\alpha = 0.84$; persistence, $\alpha = 0.82$; effort, $\alpha = 0.73$) enjoyed a relatively high level of reliability.

Strategy Inventory for Language Learning

Oxford's (1990) SILL (version 7) was utilized to determine the frequency of LLSs. The SILL included 50 items in six subscales: (a) memory strategies utilized for storing and retrieving data ($n = 9$ items), (b) cognitive strategy employed for comprehension and production ($n = 14$ items), (c) compensation strategy aimed to address boundaries in linguistic knowledge/ performance ($n = 6$ items), (d) metacognitive strategy aimed to plan, organize, and monitor learning ($n = 9$ items), (e) affective strategy exploited to control motivation and emotion ($n = 6$ items), and (f) social strategies applied for interactive cooperation ($n = 6$ items). It employed a five-point Likert type ranging from 5 (always or almost always true of me) to 1 (never or almost never true of me). The score in a complete SILL ranged from 50 (minimum) to 250 (maximum). Different researchers (Griffiths, 2018; Habók and Magyar, 2018; Shirzad et al., 2020) reported the reliability coefficients for the SILL ranging from 0.85 to 0.98. Tahmasebi (1999) used CFA and EFA psychometric methods for validating and norming the translated version of the questionnaire for the Iranian students. The internal consistency of the scale was $\alpha = 0.91$, suggesting that the translated version enjoys a relatively high internal consistency.

Procedure

To collect the data, three questionnaires (i.e., *EBQ*, *GSEQ*, and *SILL*) were disseminated both through the survey link and through direct contact of the students. To undertake the study, the Google Docs application was utilized as a platform to create

TABLE 2 | Skewness, kurtosis, and normality test for different variables.

Construct	Skewness		Kurtosis		Kolmogorov Smirnov*	
	SE	Statistic	SE	Statistics	Statistics	Sig.
Knowledge	0.113	0.224	-0.162	0.364	1.695	0.006
Learning agent	0.302	0.224	-0.535	0.364	2.116	0.000
EBs	0.379	0.224	-0.077	0.364	1.529	0.019
Initiative	-0.040	0.224	-0.185	0.364	1.731	0.005
Effort	-0.115	0.224	-0.116	0.364	1.521	0.02
Persistence	0.180	0.224	-0.159	0.364	1.681	0.007
LSE	-0.151	0.224	-0.217	0.364	1.814	0.003
Memory	0.275		-0.222	0.364	1.134	0.153
Cognitive	-0.46	0.224	-0.175	0.364	1.090	0.186
Compensatory	-305	0.224	-0.525	0.364	1.136	0.151
Metacognitive	0.322	0.224	-0.639	0.364	1.023	0.246
Affective	-0.311	0.224	-0.506	0.364	2.056	0.000
Social	-164	0.224	0.345	0.364	1.702	0.006
LLS	-137	0.224	-0.173	0.364	1.361	0.049

*This is a lower bound of the true significance.

TABLE 3 | Outlier detection with Mahalanobis distance.

	Minimum	Maximum	Mean	SD	N
MD	0.004	47.341	7.867	3.469	300
Leverage values	0.000	0.022	0.007	0.005	300

MD, Mahalanobis' Distance.

online questionnaires. Besides, some hard copies were distributed *via* the direct contacts of the researchers. Each questionnaire was distributed during regular class time. The students were asked to download each questionnaire and fill it out in the classroom, which took 15 min on average. They were asked to complete the questionnaires as meticulously as possible. To avoid fatigue, the instruments were administered at different intervals. The data were collected over 3 months in 12 weeks during the fall semester of 2019. A total of 240 students completed the questionnaires. All were completed to maximize the response rate (e.g., highlighting the benefits of the study, ensuring anonymity, providing reinforcement to respond). To minimize the bias effect, different high schools from three districts in two cities were randomly selected to distribute the questionnaires. After collecting the data, all the responses were screened for fact-checking to promote the veracity and correctness of reporting. Accordingly, a total of 130 questionnaires (29%) were not qualified for the analysis because they were incomplete or returned late. Specifically, 720 questionnaires (86%) met a valid response rate of 95%. Then, all the valid data were analyzed in the statistical package for the social science (SPSS) and the analysis of moment structures (AMOS) software using the SEM approach.

Data Analysis

The data were analyzed using the SPSS version 22 and the AMOS version 20. The SPSS was employed to run descriptive

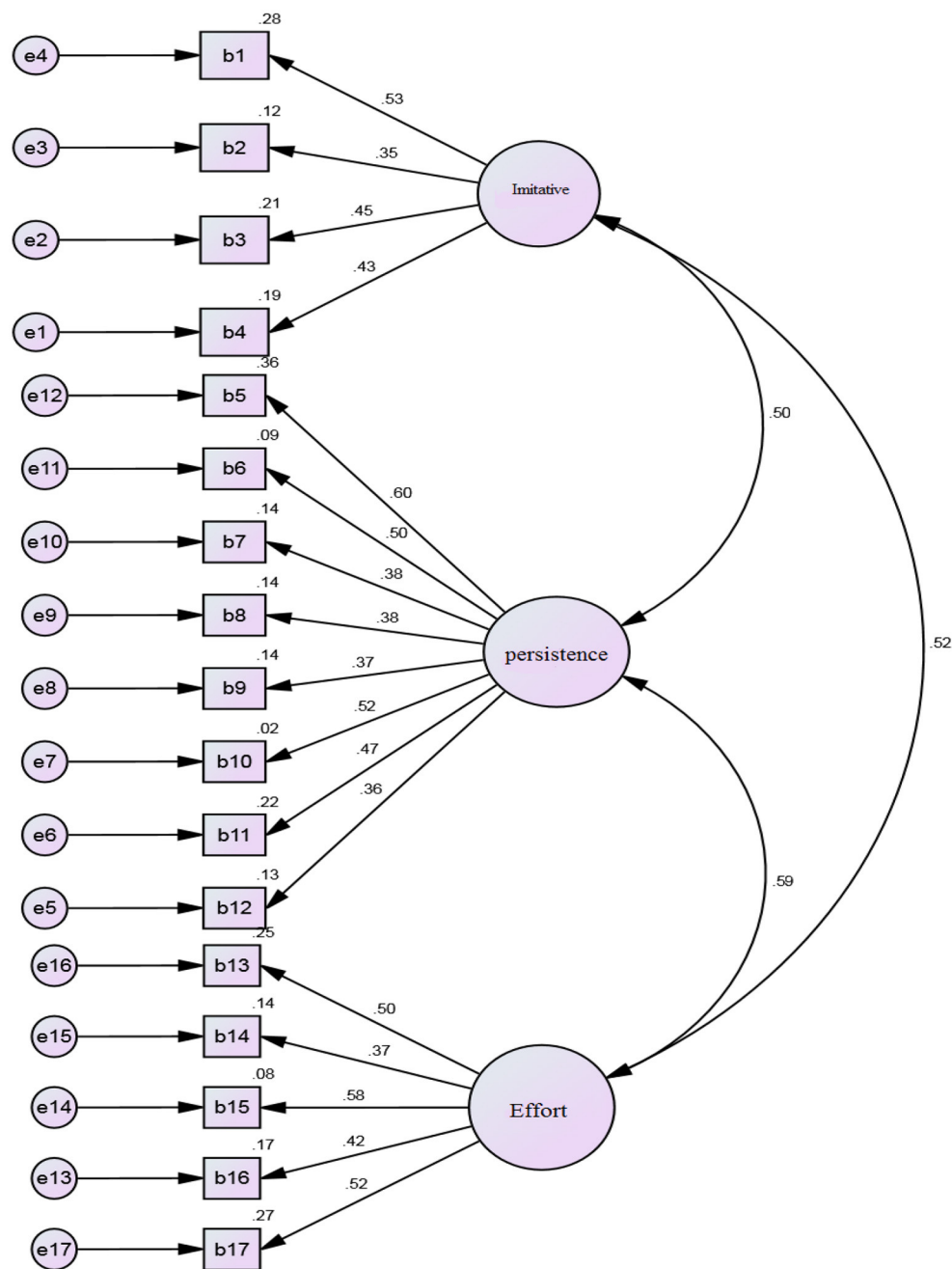


FIGURE 2 | Standardized (β) coefficients for CFA analysis and error variance of LSE.

statistics, and to check the normality of the data. The AMOS software was utilized to examine the probable structural relations between the independent variables (i.e., EBs, LSE) and the dependent variables (i.e., LLs). SEM is a theory-driven and analytic procedure that provides the capability of path analysis to examine the interplay among various latent and observed variables with the capacity of factor analysis to ensure the construct validity of the factors and subfactors (Clark-Carter,

2010; Creswell, 2014). To run the SEM, both measurement and structural models are used. The former examines the interplay between a latent variable and its indicators. The latter checks the relation between the latent variables (Kline, 2011). Thus, CFA and expectation-maximization algorithm (EMA) were used to check the validity and missing items. Following Kline's guidelines, the goodness-of-fit indices (GFI) were utilized to examine the validity of the hypothesized model. The indices included χ^2/df ,

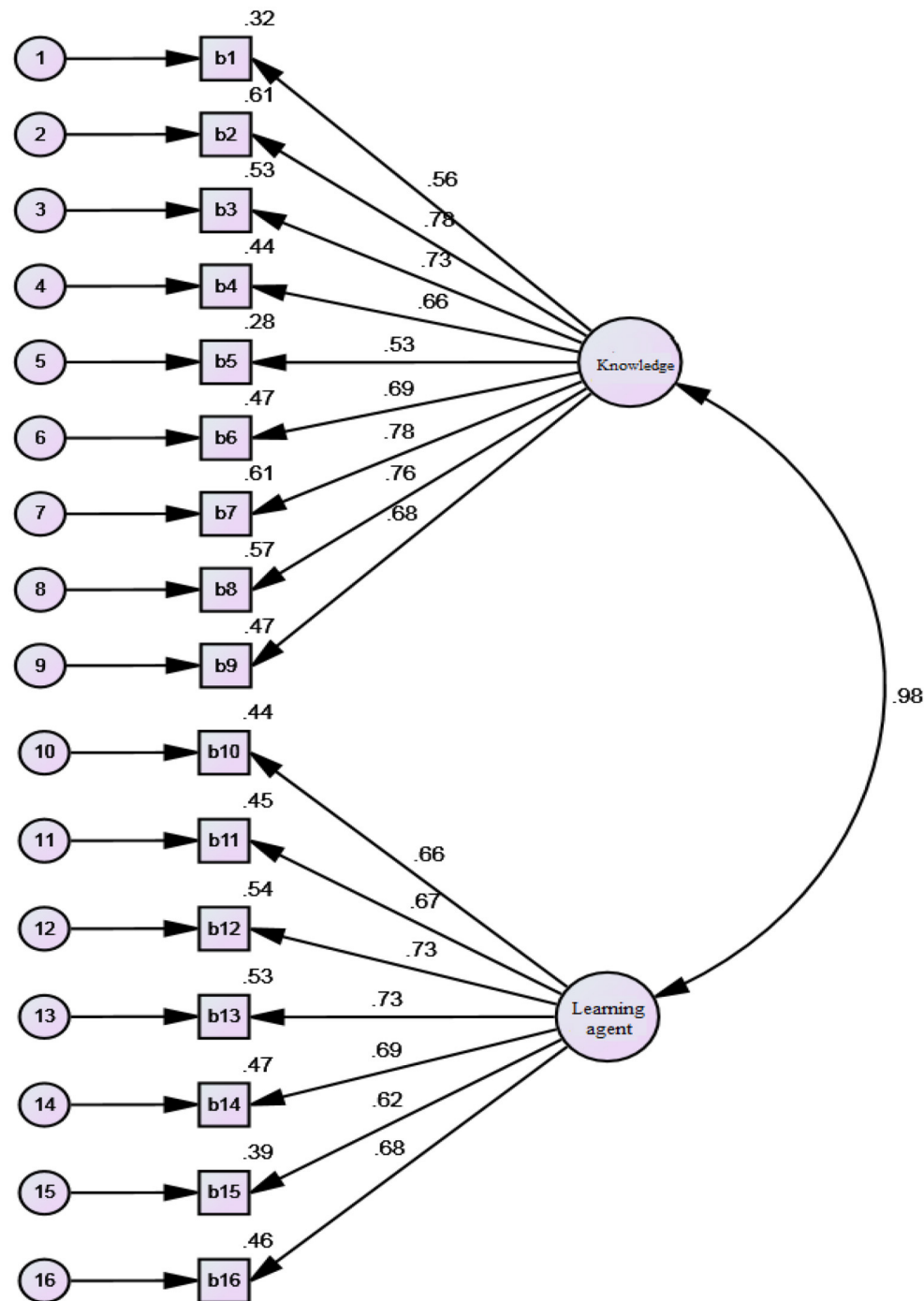


FIGURE 3 | Standardized (β) coefficients for CFA analysis and error variance of EBs.

GFI, comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA). To check the parameters of distribution by promoting a likelihood function, maximum likelihood estimation (MLE) was utilized. Kline (2011) proposed that the values of these indices are acceptable if $RMSEA < 0.06$, $\chi^2/df < 3$, $CFI > 0.95$, $TLI > 0.95$, and $GFI > 0.95$.

RESULTS

Screening the Assumptions of Normality

To answer the research questions, some preliminary steps were taken to check the assumptions for normality of EBs, LLSs, and LSE. In so doing, skewness and kurtosis analyses of the target

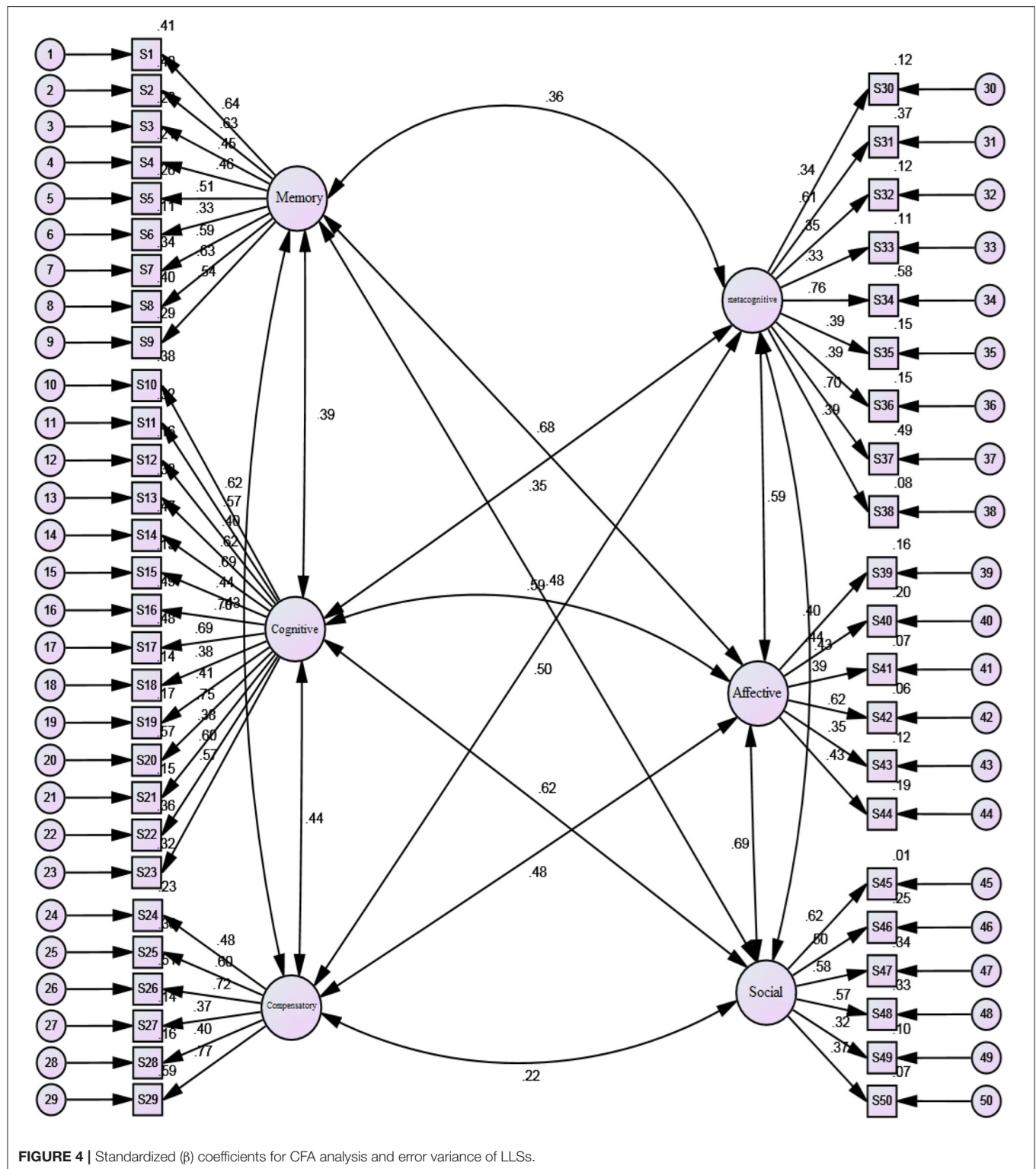


FIGURE 4 | Standardized (β) coefficients for CFA analysis and error variance of LLSs.

variables were run. Table 2 illustrates skewness, kurtosis, and normality analysis for the subscale of EBs, LSE, and LLS.

As indicated in Table 2, the distribution of data is normal, and the measure of skewness and kurtosis are at appropriate

bounds for the different constructs. It implies that the values of skewness for the subfactors fall between -3 and $+3$, and kurtosis range from -10 to $+10$. Specifically, the measure of skewness (range = -0.040 to 0.379) and kurtosis (range =

−0.077 to −0.639) are at appropriate bounds for the different subscales. In addition, the result of the KS test indicated that the data were not normally distributed ($p > 0.05$). To identify multivariate outliers, the Mahalanobis test was run. Mahalanobis distance (MD) determines the multivariate outliers. A maximum

MD larger than the critical chi-square value ($p < 0.001$) for $df = k$ (i.e., the predictor construct) shows the number of one or more multivariate outliers (Aryadoust and Raquel, 2020).

TABLE 4 | Conformity of measurement models with fitness indicators.

Constructs	CFI	X2/df	GFI	TLI	RMSEA	Sig.
LSE	0.957	104	0.967	0.972	0.048	0.000
EBs	0.954	72	0.93	0.968	0.031	0.000
LLSs	0.984	24	0.982	0.985	0.032	0.000

TABLE 5 | Composite reliability for EBs, LSE, and LLSs.

Construct	AVE	CR	Cronbach alpha
Knowledge	0.723	0.839	0.745
Learning agent	0.521	0.770	0.786
EBs	0.539	0.766	0.824
Initiative	0.567	0.793	0.811
Effort	0.555	0.751	0.754
Persistence	0.590	0.801	0.796
LSE	0.53	0.916	0.854
Memory	0.555	0.895	0.752
Cognitive	0.553	0.832	0.798
Compensatory	0.527	0.810	0.731
Metacognitive	0.502	0.734	0.751
Affective	0.723	0.839	0.769
Social	0.572	0.759	0.824
LLSs	0.555	0.726	0.846

AVE ($p > 0.5$); CR ($p > 0.7$).

TABLE 6 | Pearson correlation matrix among EBs, LSE, and LLSs.

F	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	45.97	5.87	1													
2	40.63	4.34	**0.82	1												
3	81.03	12.25	**0.62	**0.51	1											
4	11.82	2.45	**0.29	**0.33	**0.20	1										
5	21.68	2.58	**−0.18	**−0.18	**−0.18	**0.52	1									
6	13.83	3.38	**−0.17	**−0.18	**−0.17	**0.49	**0.55	1								
7	41.18	5.04	**−0.20	**−0.19	**−0.21	**0.54	**0.67	**0.62	1							
8	24.38	3.67	**−0.19	**−0.19	**−0.17	**0.21	**0.16	**0.15	**0.19	1						
9	33.17	2.89	**−0.17	**−0.18	**−0.18	**0.19	**0.17	**0.19	**0.21	**0.51	1					
10	15.96	2.32	**−0.30	**−0.27	**−0.26	**0.31	**0.19	**0.20	**0.25	**0.43	**0.43	1				
11	19.74	1.23	**−0.26	**−0.26	**−0.21	**0.25	**0.17	**0.21	**0.24	**0.35	**0.50	**0.48	1			
12	18.06	1.45	**−0.23	**−0.20	**−0.22	**0.25	**0.10	**0.17	**0.22	**0.48	**0.52	**0.68	**0.51	1		
13	15.73	1.22	**−0.22	**−0.26	**−0.37	**0.53	**0.20	**0.22	**0.22	**0.66	**0.65	**0.68	**0.40	**0.63	1	
14	127.03	8.74	**−0.21	**−0.28	**−0.33	**0.60	**0.22	**0.22	**0.23	**0.74	**0.40	**0.46	**0.51	**0.46	**0.64	1

1. Knowledge; 2. Learning agent; 3. EBs; 4. Initiative; 5. effort; 6. persistence; 7. LSE; 8. Memory; 9. Cognitive; 10. Compensatory; 11. Metacognitive; 12. Affective; 13. Social; 14. LLSs. ** $P < 0.01$.

TABLE 7 | Goodness-of-fit indices of the EBs, LSE, LLSs.

Fit index	Preference value	Obtained value before revision	Obtained value after revision
χ^2/df	<3	3.042	2.847
χ^2	–	319.41	296.088
Df	–	105	104
RESMA	<0.1	0.051	0.041
AGFI	≤0.90	0.897	0.990
NFI	≤0.90	0.909	0.982
CFI	≤0.90	0.923	0.993

TABLE 8 | Direct maximum likelihood estimation for LLSs.

Variable	Unstandardized coefficients	Standardized coefficients	R ²	T	Sig.
	B	B			
EBs	−0.482	−0.380	0.183	5.739	0.001
LSE	0.243	0.167	0.042	3.421	0.002

TABLE 9 | Bootstrap estimate of indirect effect of EBs on LLSs with mediating LSE.

Variable	B	Lower limit	Upper limit	Sig.
EBs with mediating role of LSE on LLSs	0.441	0.260	0.497	0.000

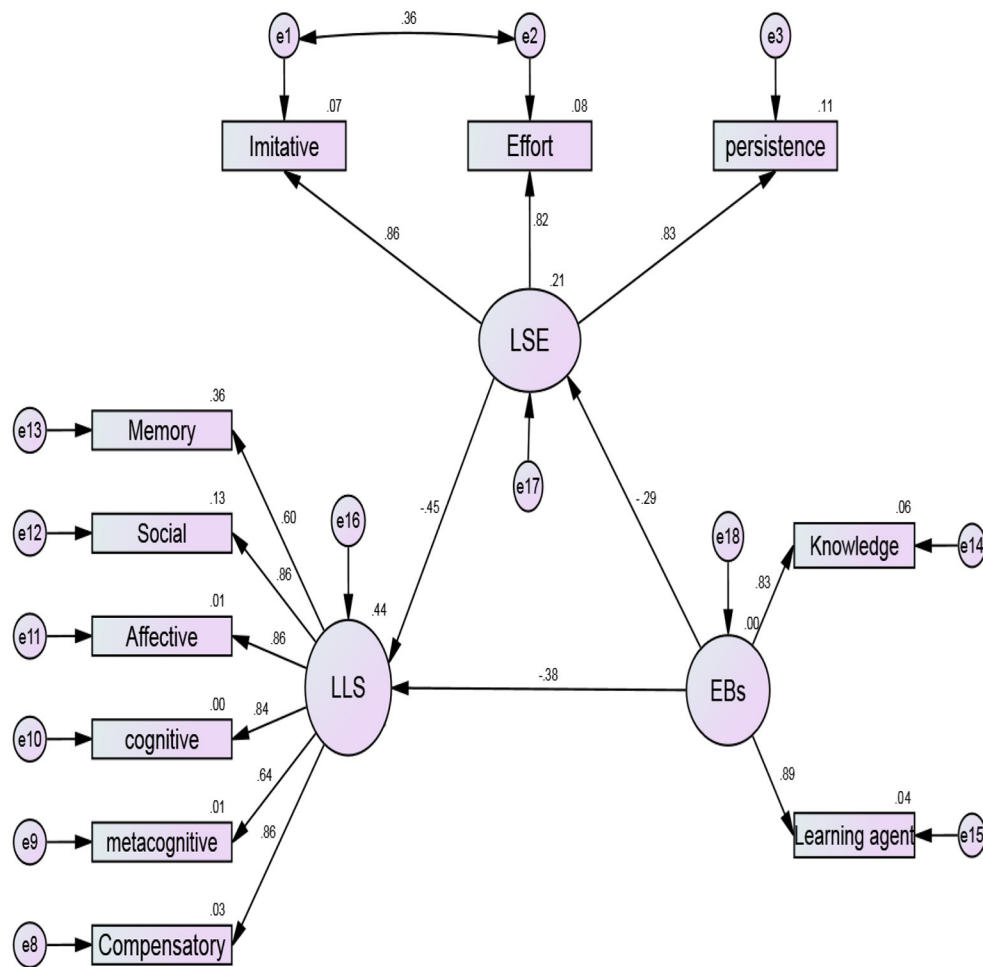


FIGURE 5 | Standardized tested model and interrelationships among the EBs, LSE, and LLSSs.

As indicated in **Table 3** ($K = 28.87$; $df = 19-1$, $p < 0.005$), the minimum and maximum MD are 0.004 and 0.47, 0.341, respectively. The MD analysis indicates that 15 multivariate outliers do not match the general character of the dataset, and the total number of students ($n = 258$) falls within the normal range.

Validation of Scales: CFA and Composite Reliability

To assure the construct validity of the instruments, CFA was run. The following models were designed and analyzed by the AMOS Graphics. Each figure schematically represents the standardized (β) coefficients for CFA analysis, different degrees of observed variables along standardized and unstandardized indices (see **Figures 2–4**).

The standardized (β) coefficients for CFA analysis indicate that all observed variables are above 0.30. **Table 4** reveals that all observed variables for the different subscales fall within the acceptable model fit for LSE (i.e., $CFI = 0.957$; $df = 104$, $GFI = 0.967$, $TLI = 0.972$, $RMSEA = 0.048$), EBs (i.e., $CFI = 0.954$; $df = 72$; $GFI = 0.93$, $TLI = 0.968$, $RMSEA = 0.031$), and LLSS

(i.e., $CFI = 0.984$; $df = 24$; $GFI = 0.982$, $TLI = 0.985$, $RMSEA = 0.032$). Thus, all the values illustrated for the model confirm that the factor loading of all the subscales for the variables are within the acceptable range and should be considered for the current study. **Table 4** summarizes the confirmatory measurement model with the indicators.

In addition, the convergent validity of the measurement model was examined by the average variance extracted (AVE) and composite reliability (CR). Following Fornell and Larcker's (1981) guidelines, the values of each indices are acceptable if AVE ($p > 0.5$) and CR ($p > 0.7$).

Table 5 indicates that the AVE and CR for all the components are above the criterion limit. Thus, it could be concluded that all questionnaires enjoy internal consistency as far as AVE and CR are concerned.

To probe the interconnections among the main constructs (i.e., EBs, LLSS, and LSE), Pearson product-moment correlation was employed. **Table 6** indicates the Pearson correlation matrix among EBs, LSE, and LLSS along with all the relevant subfactors.

The results of descriptive statistics and Pearson correlation matrix among EBs, LSE, and LLSs indicate a significant correlation between two subfactors of EBs and three subfactors of LSE and LLSs. Notably, there is a significant negative interplay between EBs and LLSs. This implies that students with a higher level of EBs employ fewer learning strategies. Besides, there is a significant positive correlation between the total LSE and total LLSs. This shows that when language self-efficacy increases, students tend to use more learning strategies.

After conducting first-order CFA, the SEM approach was conducted to uncover the causal effects in the hypothesized model and to test the significance of the effects of the main latent variables. The second phase of the study was to examine if EBs and LSE significantly predict LLSs among high school students. Therefore, different fit indices were tested to evaluate the model fit. **Table 7** indicates the GFI of the target variables.

Table 7 indicates that RMSEA (0.051) lies within the preference value ($p < 0.1$). This value represents that the mean square error of the revised model falls within the acceptable fit threshold. Following the guidelines proposed by Aryadoust and Raquel (2020), all the fit indices (i.e., $AGFI = 0.990$; $NFI = 0.982$; $CFI = 0.993$; $df = 2.847$), enjoy the acceptable fit threshold. Accordingly, the revised measurement model was considered appropriate for further analysis. **Table 8** indicates the regression analysis and coefficients for exogenous and endogenous variables. To determine the direct effect of EBs and LSE on LLSs, the MLE method was run. The MLE estimates the subfactors in distribution by maximizing a likelihood function (Richard, 2018). **Table 8** indicates the result of the MLE for LLS.

Table 8 indicates that the standardized coefficients of EBs ($\beta = -0.380$, $p < 0.001$) and LSE ($\beta = 0.167$, $p < 0.001$) are significantly predicted by LLSs. In addition, R^2 for the EBs ($R^2 = 0.183$) and LSE ($R^2 = 0.042$) reveals that the conceptual model is statistically significant. To examine the indirect effects of EBs on LLSs with the mediating role of LSE, bootstrapping regression model was run. **Table 9** reveals the bootstrap estimate of the indirect effect.

Table 9 indicates the standardized beta coefficients ($\beta = 0.441$, the lower limit = 0.260, upper limit = 0.422; $p < 0.05$). The result of the bootstrap for testing the indirect effect shows a significant level. In short, the path analysis for all direct and indirect effects predicts 34% of LLSs. The results show that EBs reduce LLS by 38% and LSE increases LLSs by 167%. **Figure 5** indicates the interrelationship among the EBs, LSE, and LLSs.

DISCUSSION

The purpose of this study was to investigate the significance of the EFL learners' EBs and LSE in predicting their LLSs. To pursue the objectives, the interplay among EBs, LLSs, and junior high school LSE along with their components was examined using the SEM approach. The results of the SEM analyses indicated that the constructs (i.e., EBs and LSE) had a different contribution to LLSs. The standardized paths after correction for direct and

indirect analyses revealed a direct and indirect effect on the EFL learners' LLSs concerning the exogenous factors (i.e., EBs and LSE). Notably, the results indicated that the subfactors of EBs and LSE significantly affect the subfactors of LLSs. The correlation coefficients among the two general variables of EBs and LSE were found to have significant positive and negative effects on LLSs.

The primary focus of this study was to probe whether EBs positively predict learners' LLSs. The path analysis of the hypothesized model revealed that EBs have a direct significant effect on LLSs. It has been hypothesized that EBs promote students' learning strategy which in turn may affect their academic achievement. The findings revealed that EBs with the mediating role of LSE on LLSs had a significant negative correlation with the subscales of LLSs. Notably, the findings suggested that EBs had an indirect effect on the types of strategy L2 learners may employ in the learning process. This finding implies that when students feel competent in their knowledge, they employ less learning strategy. The result echoes some theoretical studies (Bandura, 1997; Hofer and Pintrich, 1997; Hofer, 2016) that released evidence for the different aspects of EBs. The findings of such theoretical studies indicated a significantly strong correlation between the learners' LLSs and EBs. They postulated that some dimensions of EBs are negative predictors of academic achievement. They underscored that students are more unlikely to use different strategies when they show stronger beliefs in their knowledge and learning agent. Similarly, the findings corroborated the previous claims by L2 practitioners (e.g., Oxford, 2017; Griffiths, 2018; Chamot, 2019; Winberg et al., 2019; Tang, 2022). They corroborated that learning strategies are consistent with the beliefs held by learners. Different qualitative studies discussed in the literature review had also confirmed that the learners' beliefs (i.e., the beliefs about knowledge and learning) can influence the learning strategies their academic achievement (Habók and Magyar, 2018; Liu et al., 2019; Yang et al., 2019; Cheng, 2020; Mercer and Dörnyei, 2020). Such studies provided evidence that the level of the learners' beliefs affected their academic language achievement. Similarly, some authorities (Duell and Schommer-Aikins, 2001; Hofer, 2016; Griffiths, 2018) in the L2 professional literature highlighted that learning conception, thinking on the essence of knowledge, and learning strategies are interrelated with each other. Accordingly, they might have different positive and negative impacts on language learning. This study identified that the EBs influenced LLSs in a negative direction. The finding is consistent with Shirzad et al. (2020) corroborating that the level of the learners' EBs influence the type of strategies they utilize. Similarly, they concluded that the students who evaluate themselves as competent in their knowledge did not use strategy in terms of language learning. Therefore, it could be suggested that the higher the EBs, the fewer LLS would be employed in the learning process by language learners. In addition, the finding of the current study echoes the claim made by some practitioners (e.g., Sherer et al., 1982; Ekinci, 2017) that the learners' beliefs about their efficacy in language learning might affect their imitation, effort, and persistence. In other words, the high school students' beliefs in their competencies in employing appropriate

strategies can predict their likelihood of effort, persistence, and imitation. To simply put, the EFL students who have stronger beliefs in their knowledge and learning agents are less likely to experience LLSs.

The findings of this study added to the claim in the L2 professional literature that when the learners imagine themselves as competent learners, they can use less learning strategy. The results echo Cohen (2018) and Hofer (2016) who highlighted the learners' beliefs and learning strategies. They underscored that the beliefs held by the learners and learning strategies play a pivotal role in the learning process. Specifically, the findings proved the theoretical underpinning that EBs can affect academic success. Accordingly, they corroborated Razmi and Jabbari (2021) model for the learners' beliefs in that EBs of the learners are in line with the learners' cognitive and affective factors. Notably, this finding supports Razmi and Jabbari (2021) theoretical claim that different aspects of the beliefs about the structure and source of knowledge affect the learners' academic achievement and psychological factors. Moreover, several qualitative studies, discussed in the literature review released evidence that the learners' beliefs promoted academic success (e.g., Morris et al., 2017; Lindner and Retelsdorf, 2019; Takeuchi, 2019; Mercer and Dörnyei, 2020). Similarly, some practitioners (e.g., Liu et al., 2019; Yang et al., 2019; Cheng, 2020) substantiated the predictive role of the learners' beliefs in the learning outcome and course satisfaction. They postulated that the beliefs held by the learners about knowledge and learning could affect their academic language achievement positively.

Another aspect of this study was to probe if LSE positively predicts their LLSs. Despite sufficient evidence to justify the positive effect of the learners' beliefs, this paper hypothesized that LSE may have a complex and unpredictable effect. Thus, it has been hypothesized that LSE promotes their learning strategy which in turn may foster their academic achievement. The path analysis of the hypothesized model revealed that LSE has a direct significant effect on LLSs. The analysis verified that all direct and indirect effects could account for 34% of the LLSs. The standardized tested model and correlation among the components of LSE and LLSs indicated that the two constructs (i.e., LSE and LLSs) correlated in a positive direction. Therefore, it could be postulated that when the students have a high level of academic self-efficacy, they use more learning strategies. It implies that the students who had a higher level of LSE seemed to employ greater LLSs. Moreover, some qualitative studies highlighted that self-efficacy positively correlated with the overall strategy use. Different studies (e.g., Bandura, 2006; Pajares, 2007; Lindner and Retelsdorf, 2019; Liu et al., 2019; Cheng, 2020) discussed in the L2 professional literature have also acknowledged a positive interplay between these two variables. The findings of the present study echoes different bodies of studies (i.e., Schunk and Zimmerman, 2007; Hofer, 2016; Osiochru, 2018; Winberg et al., 2019) and suggests that LSE was the robust predictor of LLSs. Moreover, the findings supported Bråten and Olaussen (2005, cited in Bandura, 2006) who specified that learners with higher levels of LSE seem to have higher knowledge beliefs. The result of the present study affirmed previous studies that learners who had a higher level

of LSE also reported greater use of LLSs (e.g., Bandura, 2006; Pajares, 2007). The findings of this study added the claim in the literature that LSE might be increased teaching how to learn LLSs. Specifically, the findings reinforced the claim made by some authorities (e.g., Oxford, 2017; Cohen, 2018; Takeuchi, 2019) that learning strategies can promote different self-learning (e.g., Selfefficacy, self-regulated learning strategies, self-directed learning). Likewise, the findings support Cheng (2020) and Liu et al. (2019) who claimed that LSE serves a *self-regulatory* function. They postulated that LSE provides students with the ability to affect their cognitive processes and actions. Overall, the findings disclosed that LSE could affect LLSs in a positive direction and it could foster academic success in general. This finding might be a direction for future research.

CONCLUSION

Considering the findings of the study, some pedagogical implications were made for those individuals who work as educational planners, teachers, and learners in the educational contexts. Furthermore, the results can be helpful for educational policymakers to review their educational policies and programs for teacher training according to the proposed model. Specifically, positive interconnections were observed between two subfactors of EBs and three subfactors of LSE concerning LLSs. Accordingly, a general conclusion for the current study is that incorporating a focus on learners' beliefs into L2 language learning can promote EFL learners' pedagogical efficacy in general. A straightforward implication of the study is that LSE should receive more attention from the language teachers and language policymakers. Notably, more attention should be given to fostering learners' beliefs, in general, and promoting LSE skills. Thus, L2 learners will not only get higher academic achievements, but may also be motivated in learning, develop autonomous learning, and self-regulate their academic activities. To put it simply, a distinct conclusion for this study is that the EFL students' EBs and LSE can affect their choice and application of LLSs. The findings illustrate that there is a negative correlation between EBs and LLS, and there is a positive causal relationship between LSE and LLSs. Notably, the model proposed in this study suggests that the higher the students have epistemic beliefs, the less likely they will adopt a wide range of LLSs. Besides, the more LSE they have, the more likely they will use LLSs. This reveals that the beliefs held by the learners and the level of LSE influence the type of strategies they adopt. The proposed model encourages material developers and school managers to consider learners' beliefs seriously to help students promote sophisticated beliefs about knowledge and learning agent. However, due to the limitations we encountered for collecting the data and selecting the subjects at the high school level, we may not generalize the findings to other contexts. Therefore, the current study could be replicated to investigate the level of EFL learners' EBs, reflective thinking, learning strategy use, and the contribution of EBs to their learning strategy use (i.e., cognitive, socio-affective, and metacognitive strategies). Besides, future studies may be directed if qualitative or mixed-method research designs with different validated scales are adopted to generalize the findings.

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

REFERENCES

- Ary, D., Jacobs, L. C., Irvine, C. K. S., and Walker, D. (2018). *Introduction to Research in Education*. Boston, MA: Cengage Learning.
- Aryadoust, V., and Raquel, M. (2020). *Quantitative Data Analysis for Language Assessment Volume II: Advanced Methods*. New York, NY: Routledge. doi: 10.4324/9781315187808
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive View*. Eaglewood Cliffs: Prentice-Hall.
- Bandura, A. (1988). Self-efficacy conception of anxiety. *Anxiety Res.* 1, 77–98. doi: 10.1080/10615808808248222
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W. H. Freeman.
- Bandura, A. (2006). “Guide to constructing self-efficacy scales,” in *Self-Efficacy Beliefs of Adolescents*, eds F. Pajares and T. Urdan (Greenwich, CT: Information Age Publishing), 307–337.
- Chamot, A. U. (2019). “Differentiation in language learning strategy instruction on,” in *Learning Strategy Instruction in the Language Classroom: Issues and Implementation*, eds A. U. Chamot and V. Harris (Bristol: Multilingual Matters), 115–132. doi: 10.21832/9781788923415
- Chan, K. W., and Elliott, R. G. (2004). Relational analysis of personal epistemology and conceptions about teaching and learning. *Teach. Teacher Educ.* 20, 817–831. doi: 10.1016/j.tate.2004.09.002
- Cheng, Y.-Y. (2020). Academic self-efficacy and assessment. *Educ. Psychol.* 40, 389–391. doi: 10.1080/01443410.2020.1755501
- Clarebout, G., Elen, J., Luyten, L., and Bamps, H. (2001). Assessing epistemological beliefs: Schommer’s questionnaire revisited. *Educ. Res. Educ.* 7, 53–77. doi: 10.1076/edre.7.1.53.6927
- Clark-Carter, D. (2010). *Quantitative Psychological Research: The Complete Student’s Companion*. Sussex: Psychology Press-Taylor & Francis.
- Cohen, A. D. (2018). “Moving from theory to practice: a closer look at language learner strategies,” in *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*, eds R. L. Oxford and C. M. Amerstorfer (London: Bloomsbury), 31–53.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches, 4th Edn.* Thousand Oaks, CA: Sage.
- Deuling, J. K., and Burns, L. (2017). Perfectionism and work-family conflict: self-esteem and self-efficacy as mediator. *Personal. Individ. Differ.* 116, 326–330. doi: 10.1016/j.paid.2017.05.013
- Duell, O. K., and Schommer-Aikins, M. (2001). Measures of people’s beliefs about knowledge and learning. *Educ. Psychol. Rev.* 13, 419–444. doi: 10.1023/A:1011969931594
- Ekinci, N. (2017). Examining the relationships between epistemological beliefs and teaching and learning conceptions of lower-secondary teachers. *Inonu Univ. J. Faculty Educ.* 18, 344–358. doi: 10.17679/inuefd.307065
- Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Market. Res.* 18, 382–388. doi: 10.1177/002224378101800313
- Greene, J. A., Sandoval, W. A., and Bråten, I. (2016). *Handbook of Epistemic Cognition*. New York, NY: Routledge. doi: 10.4324/9781315795225
- Griffiths, C. (2018). *Strategy Factor in Successful Language Learning: The Tornado Effect, 2nd Edn.* Bristol, UK: Multilingual Matters. doi: 10.21832/9781783099757
- Habók, A., and Magyar, A. (2018). The effect of language learning strategies on proficiency, attitudes, and school achievement. *Front. Psychol.* 8, 23–58. doi: 10.3389/fpsyg.2017.02358
- Hofer, B. (1994). “Epistemological beliefs and first-year college students: Motivation and cognition in different instructional contexts,” in *Paper Presented at the Annual Meeting of the American Psychological Association* (Los Angeles, CA: American Psychological Association).
- Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educ. Psychol. Rev.* 13, 353–383. doi: 10.1023/A:1011965830686
- Hofer, B. K. (2016). “Epistemic cognition as a psychological construct: advancements and challenges,” in *Handbook of Epistemic Cognition*, eds J. A. Greene, W. A. Sandoval, and I. Braten (New York, NY: Routledge), 19–38.
- Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67, 88–140. doi: 10.3102/00346543067001088
- Kärchner, H., Schöne, C., and Schwinger, M. (2021). Beyond the level of self-esteem: exploring the interplay of level, stability, and contingency of self-esteem, mediating factors, and academic achievement. *Soc. Psychol. Educ.* 24, 319–341. doi: 10.1007/s11218-021-09610-5
- Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.
- Kumaravadivelu, B. (2006). *Understanding Language Teaching From Method to Postmethod*. Mahwah, NJ: Lawrence Erlbaum.
- Lindner, C., and Retelsdorf, J. (2019). Perceived—and not manipulated—self-control depletion predicts students’ achievement outcomes in foreign language assessments. *Educ. Psychol.* 40, 490–508. doi: 10.1080/01443410.2019.1661975
- Liu, H., Yao, M., Li, R., and Zhang, L. (2019). The relationship between regulatory focus and learning engagement among Chinese adolescents. *Educ. Psychol.* 40, 430–447. doi: 10.1080/01443410.2019.1618441
- Lonka, K., and Elina Ketonen, E., and Vermunt, J. D. (2021). University students’ epistemic profiles, conceptions of learning, and academic performance. *Higher Educ.* 81, 775–793. doi: 10.1007/s10734-020-00575-6
- Magolda, M. B. B. (1992). *Knowing and Reasoning in College: Gender-Related Patterns in Students’ Intellectual Development*. San Francisco, CA: Jossey Bass.
- Mardiha, S. M., and Alibakhshi, G. (2020). Teachers’ personal epistemological beliefs and their conceptions of teaching and learning: a correlational study. *Cogent Educ.* 7, 1–14. doi: 10.1080/2331186X.2020.1763230
- Mercer, S., and Dörnyei, Z. (2020). *Engaging Language Learners in Contemporary Classrooms*. Cambridge: Cambridge University Press.
- Morris, D. B., Usher, E. L., and Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: a critical review of emerging literature. *Educ. Psychol. Rev.* 29, 795–833. doi: 10.1007/s10648-016-9378-y
- Ongowo, R. O. (2021). Students’ epistemological beliefs from grade level perspective and relationship with science achievement in Kenya. *Educ. Inq.* 12, 1–17. doi: 10.1080/20004508.2021.1892917
- Osiocru, C. (2018). Can the study of students’ epistemological beliefs and epistemic match help us to explore the disciplinary nature of education studies? *Educ. Stud.* 54, 319–334. doi: 10.1080/00131946.2017.1413369
- Oxford, R. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York, NY: Newbury House/Harper & Row.
- Oxford, R. (2017). *Teaching and Researching Language Learning Strategies: Self-Regulation in Context, 2nd Edn.* New York, NY: Routledge.

- Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Measur. Eval. Counsel. Dev.* 39, 239–249. doi: 10.1080/07481756.2007.11909801
- Peffer, M. E., and Ramezani, N. (2019). Assessing epistemological beliefs of experts and novices via practices in authentic science inquiry. *Int. J. STEM Educ.* 6, 1–2. doi: 10.1186/s40594-018-0157-9
- Perry, W. G. (1970). *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. New York, NY: Holt; Rinehart and Winston.
- Razmi, M. H., and Jabbari, A. A. (2021). An investigation into perfectionism, self-efficacy, metacognitive strategy use, and second language listening comprehension: a path analysis approach. *J. Teach. Lang. Skills.* 40, 119–159. doi: 10.22099/jtls.2021.40272.2977
- Rezaei, A. (2010). Reliability and factor analysis of epistemic beliefs questionnaire. *Quart. Educ. Measur.* 1, 1–16.
- Richard, J. (2018). *Mathematical Statistics: An Introduction to Likelihood Based Inference*. New York, NY: John Wiley & Sons.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psychol.* 82, 498–504. doi: 10.1037/0022-0663.82.3.498
- Schunk, D. H., and Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Read. Writ. Quart.* 23, 7–25. doi: 10.1080/10573560600837578
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., and Rogers, R. W. (1982). The self-efficacy scale: construction and validation. *Psychol. Rep.* 51, 663–671. doi: 10.2466/pr0.1982.51.2.663
- Shirzad, S., Barjesteh, H., Dehqan, M., and Zare, M. (2020). The interplay among EFL students' epistemic beliefs, language learning strategies, and L2 motivational self system: a structural equation modeling approach. *J. Modern Res. Engl. Lang. Stud.* 8, 169–195. doi: 10.30479/jmrrels.2020.12621.1562
- Spratt, M., Humphreys, G., and Chan, V. (2002). Autonomy and motivation: which comes first? *Lang. Teach. Res.* 6, 245–266. doi: 10.1191/1362168802lr1060a
- Tahmasebi, A. (1999). *Vocabulary Learning Strategies and Language Proficiency*. (Unpublished MA thesis), University of Tehran, Tehran, Iran.
- Takeuchi, O. (2019). "Language learning strategies: insights from the past and directions for the future," in *Second Handbook of English Language Teaching*, ed X. Gao (Cham: Springer), 1–20. doi: 10.1007/978-3-319-58542-0_37-1
- Tang, M. (2022). The current situation and learning strategies of foreign students in Chinese learning following entrepreneurial psychology. *Front. Psychol.* 12:746043. doi: 10.3389/fpsyg.2021.746043
- Winberg, T. M., Hofverberg, A., and Lindfors, M. (2019). Relationships between epistemic beliefs and achievement goals: developmental trends over grades 5–11. *Eur. J. Psychol. Educ.* 34, 295–315. doi: 10.1007/s10212-018-0391-z
- Yang, F. Y., Bhagat, K. K., and Cheng, C. H. (2019). Associations of epistemic beliefs in science and scientific reasoning in university students from Taiwan and India. *Int. J. Sci. Educ.* 41, 1347–1365. doi: 10.1080/09500693.2019.1606960
- Zhu, X., Li, G. Y., Cheong, C. M., Yu, G., and Liao, X. (2021). Secondary school students' discourse synthesis performance on Chinese (L1) and English (L2) integrated writing assessments. *Read. Writ.* 34, 49–78. doi: 10.1007/s11145-020-10065-x

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 Shirzad, Barjesteh, Dehqan and Zare. 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.



Exploring the Relations Among Teachers' Epistemic Theories, Work Engagement, Burnout and the Contemporary Challenges of the Teacher Profession

Heidi Lammassaari^{1*}, Lauri Hietajärvi¹, Katariina Salmela-Aro¹, Kai Hakkarainen¹ and Kirsti Lonka^{1,2}

OPEN ACCESS

Edited by:

Min-Hsien Lee,
National Taiwan Normal University,
Taiwan

Reviewed by:

C. S. Chai,
The Chinese University of Hong Kong,
Hong Kong SAR, China
Tzung-Jin Lin,
National Taiwan Normal University,
Taiwan

*Correspondence:

Heidi Lammassaari
heidi.lammassaari@helsinki.fi

Specialty section:

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

Received: 24 January 2022

Accepted: 14 February 2022

Published: 25 April 2022

Citation:

Lammassaari H, Hietajärvi L,
Salmela-Aro K, Hakkarainen K and
Lonka K (2022) Exploring
the Relations Among Teachers'
Epistemic Theories, Work
Engagement, Burnout
and the Contemporary Challenges
of the Teacher Profession.
Front. Psychol. 13:861437.
doi: 10.3389/fpsyg.2022.861437

¹ Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, ² Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

Current educational reforms concerning curricula and digitalization challenge educators to meet new demands for learning and schooling. What is common for current educational reforms is that they tend to emphasize competencies that are not related to the traditional subject-matters and reflect a stance which presents learning as a naturally reflective and collaborative act. It is often assumed that teachers are automatically ready to implement ideas of this kind in practice. In this study, we propose that teachers' theories about knowledge, knowing and learning, particularly their epistemic theories, may be related to how teachers approach these reforms which challenge their previous ways of working and how they perceive their wellbeing at work. To examine these matters, we explored the dynamic interrelations between teachers' epistemic theories, conformity with the novel curricular and digital reforms (ideas behind the new curriculum and digitalization program), perceptions of the school leadership, work engagement and burnout. Participants (Study 1 $n = 228$; Study 2 $n = 200$) were Finnish class teachers and subject-matter teachers. Both data sets were collected before the COVID-19 pandemic. For data analysis, we plotted correlation network figures. Results showed that if teachers' epistemic theory was in harmony with the curricular or digital reforms, there is a positive association with work engagement and negative association with burnout. In sum, results of this provided a hint of the phenomenon suggesting that teachers' epistemic theories may be a factor which buffers teachers to meet the current epistemic and developmental challenges of teachers' profession, and furthermore, serve as grounds for a positive association for teachers to feel adequate and satisfied in their work.

Keywords: epistemic theory, teacher, work engagement, burnout, curriculum, digitalization

INTRODUCTION

While current policy documents concerning education have started to reflect exceedingly complex theories on the nature of knowledge and learning, new national curricula have been introduced during the last 5 years in several parts of the world followed by the Organisation for Economic Co-operation and Development (OECD) guidelines for 21st century competencies (OECD, 2018). Simultaneously, the overall globalization megatrend has guided schools to take a digital leap and therefore to reconsider the overall conceptions and contexts of learning (e.g., Chiu et al., 2021). In implementing these reforms, teachers play a central role and that is not an effortless task.

Various teacher-related factors may foster or hinder the implementation of curricular requirements, but teachers' beliefs are the filters, frames and underlying states of expectancy which color teachers' work, and from a wider perspective, schools' organizational change (Rokeach, 1968; Fives and Buehl, 2012). Especially teachers' *epistemic beliefs* (beliefs about knowledge and processes of knowing) are of interest when looking at the ideas teachers put into action (Fives and Buehl, 2012, 2016). That presumption is based on evidence showing that epistemic beliefs may shape and predict which teaching practices and pedagogy are applied (see e.g., Nespor, 1987; Pajares, 1992; Richardson, 1996; Fives and Buehl, 2012, 2016; Buehl and Beck, 2015). It is also indicated that teachers' epistemic beliefs (or personal epistemologies or conceptions of knowledge and learning, as they were referred to at the time) were related to their ideas about how to foster learning of their students (Lonka et al., 1996; Hofer and Pintrich, 1997). It has been also pointed out that teachers' beliefs about learning and teaching could even become obstacles to instruction, when teachers are overly relying on their intuitive laypeople theories, for instance, teaching and learning as non-problematic processes that can be learned by experience only (Joram and Gabriele, 1998). However, epistemic beliefs are not only separate beliefs about the nature of knowledge and knowing, but together they construct coherent *epistemic theories* (Hofer, 2016; Lonka et al., 2021).

In the present article, we explore Finnish teachers' epistemic theories and their perceptions of the Finnish National Core Curriculum (Finnish National Agency for Education, 2014) which underlines the development of broad-based competencies and interdisciplinary projects that cross the boundaries of subject-matter learning (Lonka, 2018). In general, implementing the new national curriculum reflected the growing trend of emphasizing 21st century competencies and active citizenship, as well as calling for metacognitive skills and collaborative knowledge construction was something new, even reformative (European Parliament, 2015). After this curricular renewal there was a pressure to implement a more detailed digital strategy to implement digital tools and digital learning practices better into schooling in 2019. Related to this, we then studied how teachers' epistemic theories are associated with their approaches to new digital demands.

Adopting pedagogical ideas that are not based on traditional subject-matter teaching and schooling practices and bringing

them into the classroom may fundamentally challenge teachers' intuitive epistemic theories, as they need to rethink the basis of their ideas about what learning is and how it should be promoted. Therefore, we are also interested in how challenging teachers' epistemic theories are related to teachers' wellbeing at work, especially work engagement and burnout. There is evidence indicating that shared and congruent beliefs about the importance of desirable end states or behaviors relate to job satisfaction and organizational engagement (Edwards and Cable, 2009). We focused on work engagement that typically manifests itself as high dedication, vigor, and absorption at work (Schaufeli et al., 2006). It is important that the work is experienced as meaningful and aligned with the teachers' state of expectations and values. If there is a match between teachers' epistemic theories and the ideas behind the reforms, it may promote work engagement.

In turn, if occupational stress is prolonged and gradually becomes chronic, it may lead to burnout (Maslach et al., 2001; Schaufeli et al., 2002; Salmela-Aro et al., 2011). There also is evidence that a lack of social or administrative support or a lack of coherence between the teacher and the working environment may be the kinds of stressors that pose a risk of burnout (Verquer et al., 2003; Cable and Edwards, 2004; Pillay et al., 2005; Kokkinos, 2007; Sharplin et al., 2011). Recently, there have been many other challenges in schools even before the advent of the COVID-19 pandemic, such as financial cuts, a refugee crisis and an ambitious inclusive approach in Finnish schools: there can be students from a range of socio-economic and ethnic backgrounds as well as special needs students in the same class (Lonka, 2018).

To examine these matters, we studied the dynamic interrelationship between teachers' epistemic theories, conformity with the new curriculum and digital demands (ideas behind the new curriculum and digitalization program), perceptions about pedagogical leadership, work engagement and burnout. We investigated the correlation networks in two different datasets that helped us to point out the connections between these variables. These matters were first studied at the time when the new curriculum was introduced (2016; Study 1). Then we studied the same matters 4 years later, except the main challenge was no longer implementation of the new curriculum, but the new strategy for successful digitalization process in Finnish schools (2019; Study 2). We assumed that if teachers' epistemic theories were in line with the ideas behind the curriculum reforms, it would be manifested as conformity in implementing the new practices, and moreover, it would relate to higher work engagement and lower rates of burnout. Moreover, it was presupposed that major changes in teachers' work call for an empowering or servant kind of leadership (Upadaya et al., 2016).

Teachers' Epistemic Beliefs Constitute Epistemic Theories

Teachers maintain a range of beliefs about knowledge and knowing. These beliefs are considered to be *epistemic beliefs*, and they are independent beliefs about knowledge, especially about the nature of knowledge and knowing, and how knowledge

can be acquired and justified (see e.g., Schommer, 1990; Hofer, 2000, 2004). Epistemic beliefs have links to several facets of knowledge-related functioning. Schommer (1990) and Lonka and Lindblom-Ylänne (1996) showed that students' epistemic beliefs were closely related to their conceptions of learning (see also Lonka et al., 2021). Further, when investigating teachers' epistemic beliefs, they have shown to be closely connected with teachers' beliefs about teaching and learning (Maggioni et al., 2009; Olafson and Schraw, 2010). Epistemic beliefs also have a practical dimension: among teachers, epistemic beliefs are related to their instructional practices, such as how they help students in acquiring knowledge, building understanding on learning, and seeking for the truth (Nespor, 1987; Pajares, 1992; Hofer, 2004; Tsai, 2007; Maggioni and Parkinson, 2008; Vedenpää and Lonka, 2014; Buehl and Fives, 2016).

Epistemic beliefs constitute systematic dimensional constructs, namely, epistemic theories (Hofer, 2016; Muis et al., 2016; Lonka et al., 2021) that are also as an essential part of broader cognitive construction referred to as *epistemic cognition* (Chinn et al., 2011; Hofer, 2016; Lonka et al., 2021). This term is holding a position as an umbrella term for various kinds of cognitive processing related to epistemic matters (e.g., Greene et al., 2016; Hofer, 2016). The term *personal epistemology* has also been used (Hofer and Pintrich, 1997), but as terms, epistemic cognition and epistemic beliefs have recently been the most predominant (Sinatra, 2016), especially epistemic cognition holding a position as an umbrella term for various kinds of cognitive processing related to epistemic matters (e.g., Greene et al., 2016; Hofer, 2016). Common in epistemic beliefs and epistemic theories is their multidimensional nature meaning that separate beliefs constitute belief systems which may or may not develop in a synchronized way (Schommer, 1990, 1993; Hofer and Pintrich, 1997; Hofer, 2004). This multidimensional approach for epistemic beliefs can be distinguished from other approaches such as developmental or contextual approach which represent a slightly different perspective on epistemic beliefs (see e.g., Hofer, 2016).

On the basis of work by Lammassaari et al. (2021) and Lonka et al. (2021), the present article explores two quite general dimensions of teachers' epistemic theories that were confirmed both in Finnish and in Taiwanese context: *the knowledge transmission theory* and *the reflective-collaborative theory*. The former sees knowledge as certain and simple in nature, referring to the process of knowing as mere transmission, and the justification of knowledge being based mostly on certain facts delivered by the teacher. The latter epistemic theory presents knowledge as complex, relativistic and integrated in nature, and the process of knowing is creative, constructive and collaborative, and the source of knowledge is mostly based on scientific references and methods, which can be critically reflected by applying metacognitive processes. Whereas the knowledge transmission theory is often manifested as passive reproduction [e.g., the review by Fives et al. (2015)], the reflective-collaborative theory could be seen more as active knowledge construction (Lonka, 1997; Deng et al., 2014). As these two kinds of theories may occur rather clearly, they are not necessarily opposite (Lammassaari et al., 2021). In higher

education, such theories have also been referred to as *the student-centered approach*, aimed at facilitating students' learning processes and knowledge construction, and *the teacher-centered approach* which emphasizes a way of teaching in which students are more like passive recipients of information that transmitted from the teacher to the students (e.g., Trigwell and Prosser, 1996a,b; Kember and Kwan, 2000; Postareff and Lindblom-Ylänne, 2008). These approaches may be field-related: recently Lonka et al. (2021) showed that pre-service teachers were more likely to express the reflective-collaborative epistemic theory than higher education students from other fields. Less is known about how teachers' epistemic theories are in line with the novel challenges of their profession (e.g., implementing new curricula or demands related to digitalization), and how it is related to teachers' wellbeing at work.

Curricular and Digital Reforms in Schools

In addition to the traditional subject-matter content, new, modernized curricula around the globe reflect new learning-related aims as they have begun to introduce content such as transversal competencies and interdisciplinary learning to be implemented in education at all levels (Lonka, 2018; OECD, 2018). Alongside the nation-specified curricular work, the OECD as a global, widely acknowledged policy forum has notably had its impact on the future of education. As an example, the OECD's Learning Compass (OECD, 2019) is a framework which presents not only shared language for international education-related discussion, but also a broad vision for the future of learning and desired competencies. It expresses "the need for students to learn to navigate by themselves through unfamiliar contexts and find their direction in a meaningful and responsible way, instead of simply receiving fixed instructions or directions from their teachers" (OECD, 2019, p. 24). These agency-related and competence-related ideas come alive in schools when constructivist, collaborative and reflective learning activities are required. However, it seems that implementing such inherently complex ideas into teaching is more difficult than the educational community tend to expect (Windschitl, 2002). Moreover, it has been shown that although teachers may express the desire to maintain nuanced epistemic beliefs and theories in harmony with the epistemic ideas of the new curricula, their actual classroom practices could continue to be quite traditional (Hofer, 2001; Vedenpää and Lonka, 2014; Buehl and Fives, 2016).

In the ideal situation, technology use should co-evolve hand-in-hand with novel learning and teaching practices (Hakkarainen, 2009; Lonka, 2018). Before the COVID-19 pandemic, the overall digital transformation in schools tended to be inert, the reason probably being that using digital technologies successfully in education call for fundamental changes in the knowledge practices in schools and other institutions (Hakkarainen, 2009). At present, collaborative tools and data-enriched technologies are increasingly about to be adopted in education (OECD, 2018), but that still requires both strong developmental upbeat from teachers, and

readiness to review critically one's own beliefs, expectations and conceptions about learning.

Becoming a modern learning environment, various practices to enable both individual and collaborative professional growth must be adopted. According to Vaara and Lonka (2014), this requires dynamic pedagogical leadership which occurs in schools mostly at the interface between the traditional domains of classroom activities and formal administration. Such leadership is often characterized by empowerment and accountability, namely servant leadership (Russell and Stone, 2002), and it creates opportunities for employees, helping them to grow. This kind of leadership promotes work-specific and general wellbeing but also seems to buffer against the negative impact of workload (Luthans, 2002; Luthans and Avolio, 2003; Upadaya et al., 2016) which might appear especially during reform situations which fundamentally change the basic principles of one's work.

Wellbeing at Work

In the present study, wellbeing at work was approached through work engagement and burnout. Work engagement is a positive state of mind that reflects one's experience about one's work (Schaufeli et al., 2002; Bakker et al., 2011). According to Schaufeli and Bakker (2004), engagement is a multidimensional factor which reveals to how vigorous, dedicated and absorbed teachers felt while working. Energy means having the will to invest mental resources in work, including persistence, which also helps to work resiliently in challenging situations and environments. Dedication is a dynamic combination of enthusiasm, sense of meaning, and deep devotion about work. For its part, absorption reflects deep concentration and attraction to work. Energy and dedication are often considered to be the key elements of work engagement, whereas it is disputed whether absorption is more like the third key dimension of work engagement or rather a consequence of it (Bakker et al., 2008).

While experiencing work engagement, there appears to be a positive individual-workplace relationship which is beneficial for both the organization and the individual (see e.g., Macey and Schneider, 2008; Bakker, 2011; Eldor, 2016). There is evidence that work engagement is related to not only individual proactivity but also how innovative the working community is (see e.g., Hakanen et al., 2008). Work engagement is typically a stable state which makes the difference to the experience of flow that is linked to a certain situation or context, for instance (Schaufeli et al., 2002; Bakker, 2011). Still, work engagement doesn't appear in a vacuum: both work-related resources and personal resources are factors that relate to how the individual experiences their work (Bakker and Demerouti, 2007; Luthans et al., 2008; Hakanen et al., 2011). Work-related resources might be physical resources, social environment and support for personal professional growth, and personal resources such as individual abilities, competencies and resilience. According to Bakker and Demerouti (2008), these two types of resource, independently or in combination, predict work engagement.

In teachers' work, its independent nature, sense of competence and support from colleagues and school leaders are found to be such organizational resources that are positively related to work engagement (Rosenholtz and Simpson, 1990; Coladarci, 1992).

While personal resources are presented as being positive self-evaluations linked to sense of ability to control and impact on one's professional environment (Xanthopoulou et al., 2009), we propose that one personal resource in teachers' work might be their approach to learning in general. First, referring to processes and outcomes related to teachers' own learning, teacher learning has been found to be a factor that is valuable, especially during school reforms (Bakkenes et al., 2010). In contrast, if teachers' intentions and epistemic theories are not in line with the demands of the new curriculum, their readiness to change their thinking and practices may hinder. Second, there is evidence in the student-context that appreciating a certain kind of learning (in this case, *reflective* learning) was positively related to study engagement (Heiskanen and Lonka, 2012). Lonka et al. (2021) also found that students who maintained collaborative-reflective epistemic theory expressed higher levels of study engagement, and furthermore, they were more likely to be student teachers than, e.g., science or engineering students. Based on these findings, it seems that in different academic cultures, a certain type of epistemic theory may create a positive relationship to engagement, especially if it is shared institutionally and in the community at hand (see also Lonka et al., 2019).

Job-related burnout, in turn, is a psychological, gradually developing syndrome in response to prolonged stress on the job (Maslach and Leiter, 1999; Maslach et al., 2001). It manifests itself as symptoms of exhaustion, cynicism, and professional inefficacy (see e.g., Maslach et al., 2001; Maslach and Leiter, 2005; Hakanen et al., 2006) or alternatively sense of inadequacy (Salmela-Aro et al., 2011). Exhaustion refers to lack of emotional energy, feelings of tension and particularly chronic fatigue caused by overtraining work while the second symptom, cynicism, is a detached or a distant attitude about work and a disaffected attitude to the people with whom one works, leading to low organizational commitment and feeling that work has lost its meaning (Maslach et al., 2001; Schaufeli and Buunk, 2003; Hakanen et al., 2006). The third symptom, lack of professional efficacy or sense of inadequacy, refers to a sense of incompetence and feeling of lacking successful achievement and productivity at work (Hakanen et al., 2006; Salmela-Aro et al., 2011).

How burnout evolves has been widely discussed. Previous studies in work contexts have shown that time pressure, work overload, emotional burden and lack of social or administrative support are all positively correlated with burnout symptoms (Schaufeli and Bakker, 2004; Hakanen et al., 2006; Kokkinos, 2007). Moreover, the lack of congruence and accordance between an employee and the working environment have been found to increase stress and job dissatisfaction, and therefore produce an increased risk of burnout (Pillay et al., 2005; Sharplin et al., 2011).

In comparison to other academic client-related professions, the average level of experiencing burnout symptoms among teachers has been found to be comparably high (Maslach et al., 1996; Schaufeli and Enzmann, 1998). Continuous challenging interaction situations, experiencing an intense work pace, heavy workload and an increasing number of administrative assignments are examples of such factors that pose a burnout risk for teachers (Leithwood et al., 1999; Hargreaves, 2003; Lindqvist and Nordänger, 2006; Skaalvik and Skaalvik, 2010). Also, it has

been found that prolonged unsolved problems and perceived destructive frictions within the professional community play a role in the stage of burden that is experienced, gradually leading to the development of teacher burnout (Pyhältö et al., 2011).

Teacher burnout is not only an individual matter, but it might also have its dynamic and complex reflections to the working context in hand. Retelsdorf et al. (2010) found that teacher burnout is related to an increased use of performance-oriented teaching practices. In turn, they are related to lower improvement in the students' conceptual application skills (Fraser, 1998; Gillies and Ashman, 2003). Moreover, Smylie (1999) showed that teachers' burnout was inversely related to schools that have been classified as innovative and learning-oriented working communities. In contrast, if individuals who should adopt something new and simultaneously have a strong belief in certain and stable facts, they may cling to their prior knowledge and focus only on things that support what they already know (Schommer-Aikins, 2011). In the teacher context, this could lead to being incapable of the cognitive flexibility that educational reforms inevitably call for. In this study, we assume that teachers' epistemic theories may color the way they perceive the novel epistemic aims of obligatory curricula and new digital demands, and therefore how much burden teachers express in their work – manifesting as burnout.

Context of the Study

This study was carried out in Finland, where the National Core Curriculum for Basic Education (Finnish National Agency for Education, 2014) is an obligatory document for “basic education” (classes 1–9) nation-wide. The National Curriculum of Basic Education (Finnish National Agency for Education, 2014) came into force in 2016, and it introduced a requirement to implement a set of broad-based competencies, as well as requiring regular projects that bridge across the borders of subject-matter domains to be carried out. Simultaneously, the Finnish National Agency for Education (2015), which included similar ideas about learning and competencies was about to be implemented. This curricular content and requirements are created by a group of municipal officials, scholars and experts, and they also guide local curricula development. The core goal for this curricular reform was to strengthen students' own activity and to increase the sense of meaning of content overall (Finnish National Agency for Education, 2014, 2015). Both the learning goals and learning content were revised to bring them into line with the competence and knowledge requirements of a modern society. In the Finnish National Agency for Education (2014) and for Finnish National Agency for Education (2015) this was manifested as referring to the importance of learning to learn, and as a demand to apply collaborative practices, creative and critical thinking, self-regulatory skills, continuous evaluation of one's own learning and meaningful use of information technology in instruction. In this study, we operationalized the new curricular demands to call for reflection and collaboration, whereas more traditional ways of instruction based on teaching certain and simple facts would be an aim that is contradictory to the ideals of the current curriculum. Our first data set was collected in 2016, just after the new curriculum implementation process was started.

Technology-enhanced pedagogy has been actively put into practice hand in hand with curricular reform of the Finnish basic education. The prevalent basic education curriculum (Finnish National Agency for Education, 2014, p. 29) makes it imperative for information technology to be “an essential part of multifaceted learning environments” and “novel solutions related to information technology will be adopted to enhance and support learning.” Our second data set was collected in 2019, when the new, municipality-specified digitalization program aiming at bringing digitalization into the core of learning and education was supposed to be implemented and applied in schools that are represented in our data (City of Helsinki Education Division, 2015). Besides enhancing students' future skills and updating the concept of learning environment, one of the key goals of the digitalization program was to develop teachers' pedagogical and digital competencies. Another was to bring change management to school principals (City of Helsinki Education Division, 2015, 2016). In this study, the new digital demands were operationalized to call for teachers' overall active role in developing their digital pedagogy and use digital tools in teaching.

Aims

In this study, we explored the relations among teachers' epistemic theories, contemporary challenges of teachers' profession and wellbeing. We aimed to find out how teachers' epistemic theories are connected to the novel ideas and requirements of modern curricular and digital demands, perceptions of pedagogical leadership work engagement and burnout. On this basis we posed the following research questions:

What are the dynamic interrelationships between teachers' epistemic theories, work engagement, burnout, and perceptions of pedagogical leadership:

1. In relation to the new curriculum demands, when was the new curriculum introduced in 2016 (Study 1)?
2. With approach to digital pedagogy and demands set for it when new digital strategy was introduced in 2019 (Study 2)?

We expected that teachers who entertained reflective-collaborative theory, coherent with the new curricular and digital demands, would report higher work engagement and lower burnout. In contrast, it was assumed that teachers who entertained an epistemic theory that reflects knowledge transmission would report lower engagement and higher burnout. Based on previous research we assumed that implementing the curricular reforms described above would require dynamic pedagogical leadership (Vaara and Lonka, 2014).

MATERIALS AND METHODS

Participants

Study 1

The participants ($n = 228$) were teachers from two major cities and one town in North-East area of Finland. Participants were

teaching in comprehensive schools as class teachers or subject-matter teachers in middle school or upper secondary school. The participants represented a range of subject-matter domains and were at various stages of their career.

Study 2

The participants ($n = 200$) were class teachers, subject-matter teachers, and special education teachers in a major city in Finland. They were working either in a primary school or middle school. The participants represented various subject-matter domains and stages of their career.

Procedure

Study 1

The data for the Study 1 were collected by a teacher questionnaire in 2016 as a part of the Mind the Gap study (Academy of Finland, 265528). This convenience sample represented selected schools which participated in the research project (2013–2016). The overall project aim was to examine prevailing gaps between the personal and social practices of students and those of their schools and the educational institutions.

Study 2

The data for Study 2 were collected by a teacher wellbeing questionnaire in 2019 in selected six schools as a part of the ongoing Growing Mind study (Strategic Research Council of the Academy of Finland, 312527). These six schools were chosen to be a part of this cooperative project aiming at producing developmental activities related to the aims of the new core curriculum and teachers' professional development.

In both studies, the participating schools were contacted by the principal investigators of the project and the teachers at these schools were asked to complete the questionnaire. Both questionnaires were conducted electronically.

Data collections were carried out in line with the ethical guidelines for research (National Advisory Board on Research Ethics, 2019), basing the participation on volition, ensuring participants' anonymity of participants, as well as storing and handling the data according to the ethical permission obtained from the University of Helsinki.

Instruments

Instruments for Study 1 and Study 2

Teachers' epistemic theories were measured with 12 questions that first posed the same item at the idea level (A), and then at the practical level (B) (Vedenpää and Lonka, 2014; see an example in **Table 1**) on scale 1–6 (1 = totally disagree; 6 = totally agree). Their instrument was based on the MED NORD questionnaire (Lonka et al., 2008), which was originally developed for assessing medical students' conceptions and orientations to knowledge and learning. The original version consisted of 93 items, and later the briefer versions were formulated based on the original version. In this study, four scales were used valuing metacognition, collaborative knowledge construction, certain knowledge, and simple/surface learning. Each scale consisted of three items. As in our more recent study, the first two scales formed *reflective-collaborative theory* and the last two formed *knowledge transmission theory* (Lammassaari et al., 2021). Reflective-collaborative theory emphasized teaching as a process to promote metacognition and collaborative knowledge construction, whereas knowledge transmission theory brought together the dualistic epistemic beliefs emphasizing the certainty of knowledge and seeing teaching as a knowledge transmission process.

Work engagement was measured by the short Utrecht Work Engagement Scale with nine items (UWES-S; Schaufeli et al., 2002, 2006). The nine items included the following three factors: vigor, dedication, and absorption. The responses were rated on a 7-point scale (0 = never; 7 = daily). Previous research had suggested the use of an overall measure of work engagement (Schaufeli et al., 2006), so we measured the participants' overall engagement at work in both studies.

Burnout was measured by nine items from the Bergen Burnout Inventory (BBI-9; Näätänen et al., 2003; Salmela-Aro et al., 2011) including the following three factors: exhaustion at work, cynicism about the meaning of work and sense of inadequacy. The responses were rated on a 6-point scale (1 = strongly disagree; 6 = strongly agree). Previous research has supported the use of an overall burnout indicator (Salmela-Aro et al., 2011) therefore, we measured the participants' overall burnout in both studies.

TABLE 1 | Instruments.

Scale name	Scale	No. of items	Abbreviation	Example item	Study 1	Study 2
Reflective-collaborative theory	1–6	6	RCT	(A) In my opinion, it is essential that students generate new ideas and thoughts together (B) I reserve a lot of time for this in my work as a teacher	x	x
Knowledge transmission theory	1–6	6	KTT	(A) The main aim of teaching is to offer certain facts about the subject of study (B) This is essential in my teaching	x	x
Work engagement	1–7	9	EDA	I find the work that I do full of meaning and purpose	x	x
Burnout	1–6	9	BBI	My expectations of my job and of my performance have been reduced	x	x
New curriculum demands	1–6	6	NCD	In my opinion, it is essential to develop students' critical thinking	x	
New digital demands	1–5	8	NDD	I strive to develop myself professionally in digital pedagogy skills		x
Empowering leadership	1–7	7	LEAD	Leadership offers me abundant opportunities to learn new skills	x	
Servant leadership*	1–5	4	LEAD	The principal encourages me to come up with new ideas		x

*This construct consists of four dimensions of servant leadership.

Contextual Instruments: Study 1

New curriculum demands (Study 1) were measured by six items considering the emphasis of teaching critical and creative thinking and providing constant constructive feedback to the pupils (Finnish National Agency for Education, 2014; Lonka, 2018). The responses were rated on a 6-point scale (1 = totally disagree; 6 = totally agree).

Empowering leadership (Study 1) was measured by seven items. The scale was based on Van Dierendonck and Nuijten (2011) scale, and the responses were rated on a 7-point scale (1 = strongly disagree; 7 = strongly agree). The items considered the encouragement, resources, and autonomous approach for one's work that the school administration provided and favored.

Contextual Instruments: Study 2

New digital demands were measured by eight items rated on a 5-point scale (1 = never; 5 = very often). The scale was developed as a part of the Growing Mind study (2019–2023) and its questionnaire targeted at teachers. Items consisted of claims related to the participants' initiative role in digital development projects in their schools and willingness to learn and adopt new digital tools, methods, and practices to develop their pedagogy. The items were formed to be in line with aims of the new digital program implemented in 2019.

Servant leadership was measured by four items (Upadaya and Salmela-Aro, 2020), and the responses were rated on a 5-point scale (1 = strongly disagree; 5 = strongly agree). The scale was based on Van Dierendonck and Nuijten (2011) scale. The items considered the encouragement, resources, and autonomous approach for one's work that the school administration provided and favored.

Data Analysis

Identical procedure of data analysis was applied for both Study 1 and Study 2. First, we specified and tested the structure of the model using item parcels of the A and B parts of the items for the epistemic theories and the raw items for the other constructs. We followed the next fit indices: chi-square, root mean square error of approximation (RMSEA) with an approximate cutoff value for a good fit of less than 0.05, SRMR with a cut-off value <0.08, comparative fit index (CFI) with a cutoff value of greater than 0.96 as well as the Tucker-Lewis index (TLI) with a cutoff value of greater than 0.95 (Hu and Bentler, 1999; Yu, 2002).

The model was specified as a confirmatory factor analysis model (CFA) with ordered items. All items were allowed to load on their corresponding factor only. The analyses were conducted using "lavaan" (Rosseel, 2012) in R and RStudio (RStudio Team, 2020). We used robust weighted least squares (WLSMV) as the estimator. Second, we used a method of visualizing the correlations between the constructs in the model (see e.g., Sjöblom et al., 2020).

To describe the method more specifically, the correlations among the epistemic theories, new curriculum demands, and approach to digital pedagogy, pedagogical leadership as well as work engagement and burnout were visualized and examined. We did this by exporting the latent variable correlation matrix of the model and visualizing the cross-sectional correlations by

plotting the latent variables as nodes in a correlation network (Epskamp and Fried, 2018). We used R-package "qgraph" (Epskamp et al., 2012), similarly to a latent variable network model (see e.g., Epskamp et al., 2017).

The edges in the latent correlation network represent simple bivariate correlations. The figures display the strength of correlations between the different constructs as well as whether it is negative or positive. The strength of this modeling is that it allows for powerful measurement error-corrected modeling of undirected structural relationships between latent variables (Guyon et al., 2017).

RESULTS

Table 2 shows means, standard deviations and Cronbach's alphas for variables used in this study.

The model in Study 1 was a good fit for the data [chi-square (725) = 1214.32, scaling = 1.837, $p < 0.001$, RMSEA = 0.055, SRMR = 0.081, CFI = 0.947, TLI = 0.942]. The model in Study 2 also fit the data well enough [chi-square (804) = 1324.53, scaling = 2.366, $p < 0.001$, RMSEA = 0.065, SRMR = 0.095, CFI = 0.937, TLI = 0.932], and no *post hoc* modifications were considered for the sake of parsimony and replicability.

Study 1

We studied the relationships between the epistemic theories and new curriculum demands, empowering leadership, work engagement and burnout. The model seen in **Figure 1** indicated that there were direct positive relationships between reflective-collaborative theory, conformity with the new curriculum demands (0.75), work engagement (0.40), and empowering leadership (0.15). This was in line with our expectations. Reflective-collaborative theory also mediated a negative relationship between knowledge transmission theory and burnout. In turn, teachers' knowledge transmission theory was positively related to empowering leadership (0.09) and burnout (0.07), but negatively to conformity with the new curriculum demands (−0.08) and work engagement (−0.03). The relationships between knowledge transmission theory and

TABLE 2 | Raw M, SD, and α for Study 1 and Study 2.

	Study 1			Study 2		
	Raw M	SD	α	Raw M	SD	α
Reflective-collaborative theory	4.8	0.619	0.887	4.9	0.494	0.828
Knowledge transmission theory	3.3	0.903	0.901	3.5	0.656	0.832
Work engagement	6.1	0.854	0.898	5.4	0.892	0.915
Burnout	2.4	0.969	0.861	2.6	0.887	0.860
Contextual measures						
New curriculum demands	4.5	0.691	0.776			
New digital demands				3.3	0.720	0.879
Empowering leadership	5.0	1.43	0.929			
Servant leadership				3.5	0.662	0.608

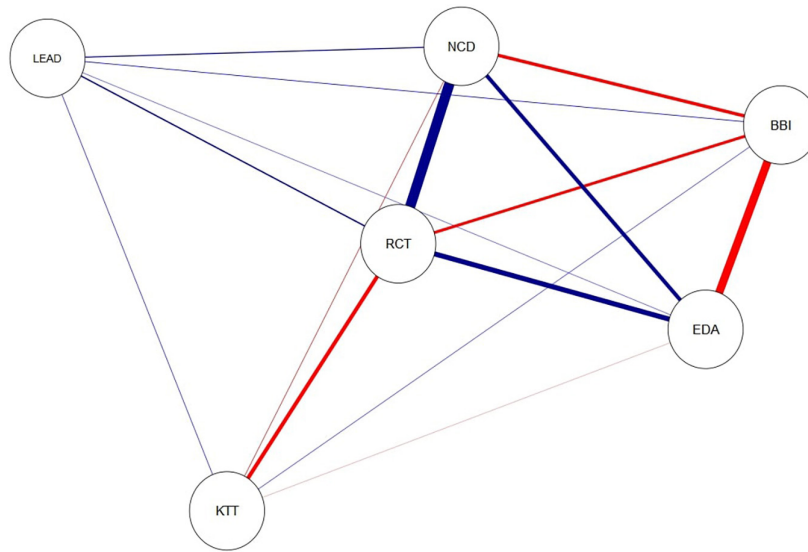


FIGURE 1 | Latent variable correlation network Study 1. Abbreviation RCT refers to reflective-collaborative theory, KTT to knowledge transmission theory. Abbreviation NCD (NCD, new curriculum demands) refers to items measuring the conformity with the new curriculum demands, LEAD to empowering leadership, EDA to work engagement, and BBI to burnout symptoms. Blue edges in the figure refer to positive relations, and red edges to negative relations. The width of the edges corresponds to the absolute value of the correlations: the higher the correlation, the thicker the edge (see Epskamp et al., 2012).

the other variables were mostly expected; however, except for a negative relationship between the two epistemic theories, the relationships were quite modest, as the width and frail edges between these variables indicate in **Figure 1**.

Then we looked at the results from the perspective of wellbeing variables. **Figure 1** shows that teachers' work engagement had a direct positive association to reflective-collaborative theory (0.40), conformity with the new curriculum demands (0.75), and empowering leadership (0.15). In contrast, work engagement was negatively related to knowledge transmission theory (-0.03) and burnout (-0.60). When observing teachers' burnout, results show that it was positively related to knowledge transmission theory (0.07) and empowering leadership (0.10), but negatively related to reflective-collaborative theory (-0.25), new curriculum demands (-0.28) and work engagement (-0.60). These findings were mostly in line with our presumptions regarding the coherence between a certain kind of epistemic theory and new curriculum demands, and its reflection on better occupational health.

Study 2

We examined the relationships between epistemic theories and new digital demands, servant leadership, work engagement and burnout. As **Figure 2** shows, there were expected positive relationships between reflective-collaborative theory and new digital demands (0.17), servant leadership (0.44), and work engagement (0.14). Contrariwise to Study 1, reflective-collaborative theory was also positively related to knowledge transmission theory (0.16). In turn, knowledge transmission theory was positively related to burnout (0.07), but negatively related to new digital demands (-0.13) and servant leadership (-0.06). These relationships were in line

with our expectations although especially relationships between knowledge transmission theory and the rest of the variables were quite modest as the width and thickness of the edges in **Figure 2** indicate.

When taking a specific look at the wellbeing variables, **Figure 2** shows that the associations between work engagement, burnout and the rest of the variables were mostly in line with our expectations: Work engagement had a positive association with reflective-collaborative theory (0.14), new digital demands (0.25), and servant leadership (0.43). Work engagement had an only direct negative association to burnout (-0.65). In contrast, burnout was negatively related to reflective collaborative theory (-0.16), new digital demands (-0.22), and servant leadership (-0.35), but positively to knowledge transmission theory (0.07). In contrast to Study 1, in Study 2, servant leadership was also mediating the associations between reflective-collaborative theory, and work engagement as well as new digital demands. In these two mediated cases, the relationships were more apparent as thicker edges in **Figure 2** indicate.

DISCUSSION

Discussion of the Results

In the present study, we aimed to find out what kinds of interrelations can be found out between teachers' epistemic theories, contemporary professional challenges related to novel curricula and digitalization, perceptions of work-related wellbeing and pedagogical leadership. Overall, the findings of this study extend the understanding of the factors which may be associated with the current institutional renewals concerning teachers' work. In Study 1 (data collected in 2016), we expected

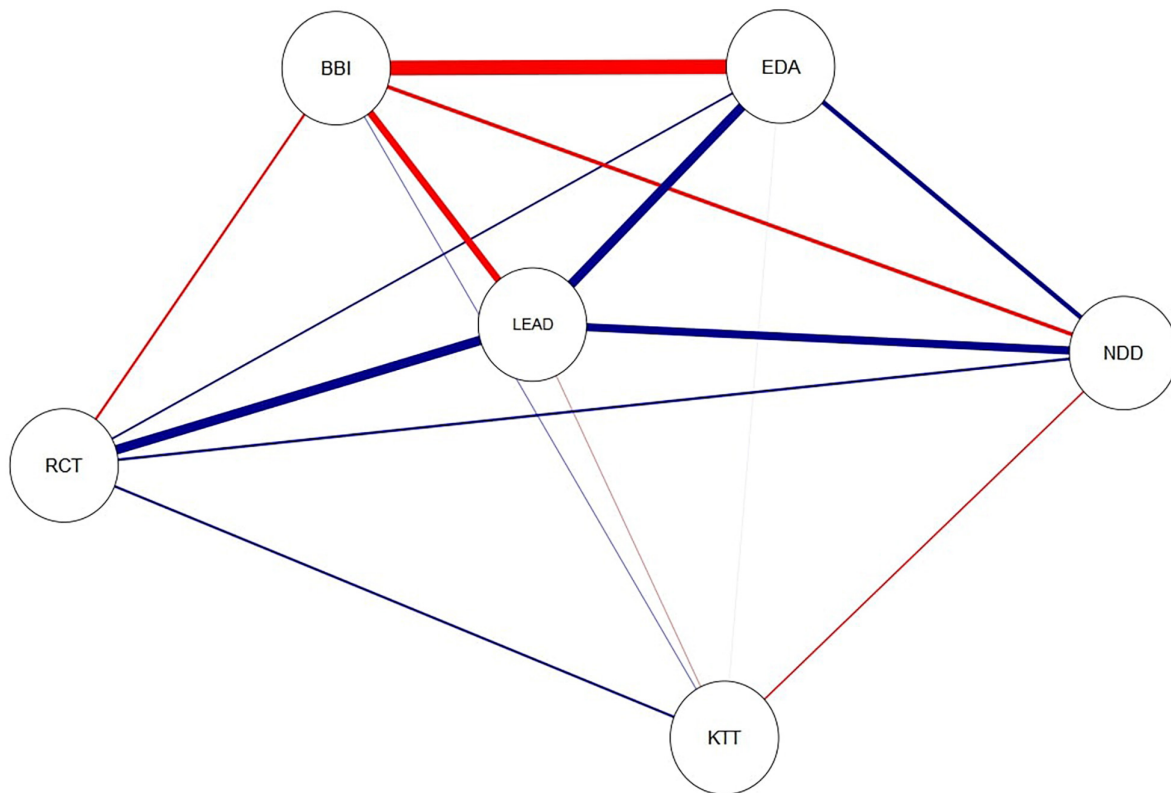


FIGURE 2 | Latent variable correlation network Study 2. Abbreviation RCT refers to reflective-collaborative theory, KTT to knowledge transmission theory, NDD (New digital demands) to developmental orientation for digital pedagogy, LEAD to servant leadership, EDA to work engagement, and BBI to burnout symptoms. Blue edges in the figure refer to positive relations, and red edges to negative relations. The width of the edges corresponds to the absolute value of the correlations: the higher the correlation, the thicker the edge (see Epskamp et al., 2012).

that reflective-collaborative epistemic theory would be positively associated with work engagement and to the new curriculum demands. This was due to the novel ideas and practices presented in the new national curriculum which emphasized reflection and collaboration as an essential part of learning and classroom activities. In Study 2 (data collected 2019), we expected that teachers who expressed this same reflective-collaborative epistemic theory would be more willing to develop their digital pedagogy by using modern digital tools and would therefore also express higher work engagement and lower burnout. During Study 2's data collection the new digitalization program reflecting the background ideas of the national curriculum reforms was about to be fully implemented.

The results in both studies were aligned with these expectations, however, in Study 2 servant leadership seemed to play its role as a strengthener or mediator of associations between reflective-collaborative epistemic theory, approach to new digital demands, and wellbeing. Both studies pointed out that if teachers' epistemic theory was in harmony with the reforms, there would appear a positive association with work engagement and negative to burnout. This finding is in line with several previous studies (see e.g., Hakanen et al., 2006; Upadaya et al., 2016) which have suggested that there occur various person-related factors which might be positively associated with teachers' wellbeing at work.

In sum, results of these studies provided a hint of the phenomenon suggesting that teachers' epistemic beliefs could be a resource which buffer teachers to meet the current epistemic and developmental demands of teachers' profession, and furthermore, serve as grounds for a positive association for teachers to feel adequate and satisfied in their work. As epistemic beliefs or theories are suggested to be socially shared, so that the origin of knowledge does not lie only within an individual mind (see e.g., Packer and Goicoechea, 2000), this resource might not occur only in the individual level but even more in a group or community level. In the case of teachers, this community could be for instance the school where they work.

In contrast, teachers' knowledge transmission theory was assumed to be less likely to agree with ideas of these same curricular and digital reforms. Knowledge transmission theory was also expected to be associated with lower work engagement and higher rates of burnout. Especially in the case of Study 2, when the focus was on digitalization, the expected associations were not that obvious.

However, what was notable were the associations between reflective-collaborative theory and knowledge transmission theory: in Study 1 the association between these two epistemic theories was negative whereas in Study 2 the association was positive. Due to a cross-sectional nature of this study, there

is limited capacity to find an explanation for this difference, nevertheless, in 2016, there might have been a more prominent need to make a shift toward the new national curriculum and its comparably sophisticated *epistemic climate* (meaning how the nature of knowledge and knowing is portrayed and perceived in a certain context, such as in classroom practices or curriculum discourses; see e.g., Feucht, 2010; Muis and Duffy, 2013). That shift might have reflected in our Study 1 results as a more observable juxtaposition between the two epistemic theories and their relationships to other variables under this exploration.

In Study 2 (2019), results hinted that the focus has changed from purely epistemic issues toward the actual challenges of digitalization. This may be due to the fact that a profound curriculum reform is more closely linked to epistemic theories than a digital reform, which may be based on various theories. This could also indicate the growing emphasis on *epistemic fluency* meaning that standing behind a certain epistemic stance is more important than one being able to recognize and use several culturally shared ideas about what knowing is and how knowledge should be constructed (Morrison and Collins, 1995): there is a place for details and facts, but also for a comprehensive approach integrating reflective and collaborative learning practices. In this respect, it is appropriate to remark that these two epistemic theories are not automatically opposite by nature, but they may exist side-by-side as Lammassaari et al. (2021) pointed out.

Methodological Limitations

The present study has some limitations. First, the study was based on teachers' self-reports, and the sample size in both studies was quite small. Especially Study 1 represented a convenience sample which is a consequence of the wider challenges in data collection related to teacher questionnaires. Over time, it has been a challenge to collect comprehensive teacher data which may be related to the considerable number of questionnaires constantly addressed to teachers in the profession. Moreover, participants in these two studies were different, making the studies present in this study cross-sectional as a design. Studies in the future call for being carried out with larger samples and with longitudinal design. Longitudinal design could reveal a more precise perception of how new institutional demands such as curricular renewals or novel digital programs and their associations to teachers' epistemic theories are reflected and evolved over time which was not possible in this study's setting.

Some scale and overall methodological developments are also required. As an example, our results did not indicate yet the directions of the relationships which would have offered another angle to overview this complex phenomenon of teachers' epistemic theories and their associations with a range of variables. In this study, we were not assuming causality, merely exploring the complex relationships between epistemic theories, work engagement and the epistemic demands of teachers' contemporary work. The scales used in this study were mostly robust, therefore more detailed perceptions related to teachers' wellbeing and as mentioned, more refined associations between used variables might have been left out of the network. Despite these promising results about teachers' epistemic theories and their relationships to current challenges of teachers' work and

occupational health, a more subtle approach should be applied to explore the wider spectrum of the epistemic theories and overlap that potentially exists between them. This may also require a mixed-method approach. Overall, this is a worthwhile working hypothesis and in the future, we will look more closely into these relations and triangulate the results with several representative samples.

Future Implications

This variable-oriented study identified associations between teachers' epistemic theories, new curricular and digital demands, school leadership and teachers' occupational health. Our research revealed that teachers' epistemic theories, which consist of teachers' epistemic beliefs and how they report putting these ideas into practice, play a role when looking at dynamic interrelationships between these theories, new demands set for schooling, work engagement and burnout. Previous research has shown that when taking a closer look at teachers' wellbeing at work, two profiles of Finnish teachers have been identified: engaged (30%) and engaged-burnout (70%) profiles, the latter being still engaged in their work, but already starting to show some symptoms of burnout (Salmela-Aro et al., 2019). Simultaneously, it has been shown that teachers may express quite complex and constructivist epistemic beliefs and theories, but their actual classroom practices may contradict them (Hofer, 2001; Vedenpää and Lonka, 2014; Buehl and Fives, 2016). Some researchers even propose that theories and beliefs about learning should be distinguished more clearly from theories about teaching (Richardson and Placier, 2001; Fives et al., 2015). In the context of this discussion, we suggest that as a combination, the possible discontinuity between teachers' epistemic theories, novel epistemic aims and demands set for schooling might show up overwhelming for some teachers and manifest, in some timespan, even as symptoms of burnout.

In this sense, novel approaches are needed to offer all teachers an equal opportunity to adopt new curriculum-related and developmental ideas as well as to implement these changes in their practice. As this study indicated, this is not only an individual matter but also a community-level matter related to, e.g., school leadership that, in its best, encourages and supports teachers to evolve and even transform in their thinking and profession to meet the new requirements. However, epistemic change, if looked for, is not a simple mission since unless individuals have a good reason to abandon their beliefs, they will be unlikely do so (see e.g., Bendixen, 2002; Muis and Duffy, 2013). In the teacher context this might be manifested as clinging onto the traditional pedagogical ideas and practices which have governed schools over time and might still offer a sense of something that surely works. For this reason, it is important to note that a prerequisite for epistemic change is that individuals must be able to understand the new beliefs, and consider them to be plausible, so that they can be applied and be fruitful for further inquiry (Pintrich et al., 1993). Interestingly, it has been shown that working in an innovative and future-oriented school is positively related to teachers' engagement (Rosenholtz, 1989). Although this offers a hint that school communities which constantly evolve and look forward to the future may keep teachers engaged, we still do not clearly recognize the origins

of this phenomenon. Therefore, further research is required to dig deeper theoretically into looking at fruitful preconditions for teachers' epistemic growth to support the readiness for change implementation. The local and global evolution is ongoing, and the post-pandemic era will inevitably bring again new challenges, such as hybrid learning, for teacher profession.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: GDPR regulations were taken into account. All data that might enable the identification of an individual participant was deleted and replaced by a participant number. Requests concerning these datasets should be directed to the corresponding author. Requests to access these datasets should be directed to HL, heidi.lammassaari@helsinki.fi.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Review Board in the humanities and social and behavioral sciences of the University of Helsinki. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Bakkenes, I., Vermunt, J. D., and Wubbels, T. (2010). Teacher learning in the context of educational innovation: learning activities and learning outcomes of experienced teachers. *Learn. Instruct.* 20, 533–548. doi: 10.1016/j.learninstruc.2009.09.001
- Bakker, A. B. (2011). An evidence-based model of work engagement. *Curr. Direct. Psychol. Sci.* 20, 265–269. doi: 10.1177/0963721411414534
- Bakker, A. B., Albrecht, S. L., and Leiter, M. P. (2011). Key questions regarding work engagement. *Eur. J. Work Org. Psychol.* 20, 4–28. doi: 10.1080/1359432x.2010.485352
- Bakker, A. B., and Demerouti, E. (2007). The job demands-resources model: state of the art. *J. Manag. Psychol.* 2, 309–328. doi: 10.3390/ijerph17010069
- Bakker, A. B., and Demerouti, E. (2008). Towards a model of work engagement. *Career Dev. Int.* 13, 209–223. doi: 10.1108/13620430810870476
- Bakker, A. B., Schaufeli, W. B., Leiter, M. P., and Taris, T. W. (2008). Work engagement: an emerging concept in occupational health psychology. *Work Stress* 22, 187–200. doi: 10.1080/02678370802393649
- Bendixen, L. D. (2002). "A process model of epistemic belief change," in *Personal epistemology: The psychology of beliefs about knowledge and knowing*, eds B. K. Hofer and P. R. Pintrich (Hillsdale, NJ: Erlbaum), 191–208.
- Buehl, M. M., and Beck, J. S. (2015). "The relationship between teachers' beliefs and teachers' practices," in *Handbook of Research on Teachers' Beliefs*, eds H. Fives and M. Gregoire Gill (New York, NY: Routledge).
- Buehl, M. M., and Fives, H. (2016). "The role of epistemic cognition in teacher learning and praxis," in *Handbook of epistemic cognition*, eds J. A. Greene, W. Sandoval, and I. Braten (Milton Park: Routledge), 247–264.
- Cable, D. M., and Edwards, J. R. (2004). Complementary and supplementary fit: a theoretical and empirical integration. *J. Appl. Psychol.* 89, 822. doi: 10.1037/0021-9010.89.5.822
- Chinn, C. A., Buckland, L. A., and Ala, S. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educ. Psychol.* 46, 141–167. doi: 10.1080/00461520.2011.587722
- Chiu, T. K., Lin, T. J., and Lonka, K. (2021). Motivating online learning: the challenges of COVID-19 and beyond. *Asia-Pacific Educ. Res.* 30, 187–190. doi: 10.1007/s40299-021-00566-w

AUTHOR CONTRIBUTIONS

HL: leading writer. LH: researcher of the projects, leading data analysis, supervisor of HL, and a contributor to the writing process. KSA: Co-PI of the Study 1 project and a contributor to the writing process. KH: PI of the Study 2 project and a contributor to the writing process. KL: PI of the Study 1 project, supervisor for HL, and a contributor to the writing process. All authors contributed to the article and approved the submitted version.

FUNDING

We are grateful for funding by The Finnish Strategic Research Council (grants #312527, #336064, #327242, and #345264) and the Academy of Finland (grants #308351, #308352, #320371, and #336138).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.861437/full#supplementary-material>

- City of Helsinki Education Division (2015). *Ilmiömainen Helsinki. Tulevaisuuden koulun suuntaviivat 2015–2020*. Helsinki: Savion Kirjapaino Oy.
- City of Helsinki Education Division (2016). *Helsingin kaupungin opetuksen digitalisaatio-ohjelma vuosille 2016–2019: Koulutuksen ja oppimisen digistrategia*. Available online at: https://www.hel.fi/static/liitteet/opev/Opetusvirasto_Digitalisaatiohanke_Webjulkaisu_FINAL.pdf. (retrieved November 8, 2021)
- Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *J. Exp. Educ.* 60, 323–337. doi: 10.1080/00220973.1992.9943869
- Deng, F., Chai, C. S., Tsai, C.-C., and Lee, M. H. (2014). The relationships among Chinese practicing teachers' epistemic beliefs, pedagogical beliefs and their beliefs about the use of ICT. *J. Educ. Tech. Soc.* 17, 245–256.
- Edwards, J. R., and Cable, D. M. (2009). The value of value congruence. *J. Appl. Psychol.* 94:654. doi: 10.1037/a0014891
- Eldor, L. (2016). Work engagement: toward a general theoretical enriching model. *Hum. Res. Dev. Rev.* 15, 317–339. doi: 10.1177/1534484316655666
- Epskamp, S., and Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychol. Methods* 23:617. doi: 10.1037/met0000167
- Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). qgraph: network visualizations of relationships in psychometric data. *J. Stat. Soft.* 48, 1–18.
- Epskamp, S., Rhemtulla, M., and Borsboom, D. (2017). Generalized network psychometrics: combining network and latent variable models. *Psychometrika* 82, 904–927. doi: 10.1007/s11336-017-9557-x
- European Parliament (2015). *Innovative schools: Teaching and learning in the Digital Era. Workshop Documentation*. Brussels: European Parliament.
- Feucht, F. C. (2010). "Epistemic climate in elementary classrooms," in *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*, eds L. D. Bendixen and F. C. Feucht (Cambridge: Cambridge University Press), 55–93. doi: 10.1017/CBO9780511691904.003
- Finnish National Agency for Education (2014). *The National Core Curriculum for Basic Education*. Helsinki: Next Print Oy.
- Finnish National Agency for Education (2015). *National Core Curriculum for General Upper Secondary Schools*. Helsinki: Next Print Oy.
- Fives, H., and Buehl, M. M. (2012). "Spring cleaning for the 'messy' construct of teachers' beliefs: What are they? Which have been examined? What can they

- tell us?," in *Educational Psychology Handbook, Vol. 2. Individual differences and cultural and contextual factors*, eds K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, and M. Zeidner (Washington, D.C.: American Psychological Association), 471–499. doi: 10.1037/13274-019
- Fives, H., and Buehl, M. M. (2016). Teachers' beliefs, in the context of policy reform. *Policy Insights Behav. Brain Sci.* 3, 114–121. doi: 10.1177/2372732215623554
- Fives, H., Lacatena, N., and Gerard, L. (2015). Teachers' beliefs about teaching (and learning). *Internati. Handb. Res. Teach. Bel.* 62, 249–265.
- Fraser, B. J. (1998). Classroom environment instruments: development, validity and applications. *Learn. Env. Res.* 1, 7–34.
- Gillies, R. M., and Ashman, A. F. (2003). "An historical review of the use of groups to promote socialization and learning," in *Cooperative learning: The social and intellectual outcomes of learning in groups*, eds R. M. Gillies and A. F. Ashman (London: Routledge Falmer), 11–28.
- Greene, J., Sandoval, W., and Bråten, I. (2016). "An introduction to epistemic cognition," in *Handbook of epistemic cognition*, eds J. A. Greene, W. A. Sandoval, and I. Bråten (Milton Park: Routledge), 1–15. doi: 10.1093/acprof:oso/9780199591596.003.0001
- Guyon, H., Falissard, B., and Kop, J. L. (2017). "Modeling psychological attributes in psychology—an epistemological discussion: network analysis vs." latent variables. *Front. Psychol.* 2017:798. doi: 10.3389/fpsyg.2017.00798
- Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *J. Sch. Psychol.* 43, 495–513. doi: 10.1016/j.jsp.2005.11.001
- Hakanen, J. J., Peeters, M. C., and Perhoniemi, R. (2011). Enrichment processes and gain spirals at work and at home: a 3-year cross-lagged panel study. *J. Occupat. Org. Psychol.* 84, 8–30. doi: 10.1111/j.2044-8325.2010.02014.x
- Hakanen, J. J., Perhoniemi, R., and Toppinen-Tanner, S. (2008). Positive gain spirals at work: from job resources to work engagement, personal initiative and work-unit innovativeness. *J. Vocat. Behav.* 73, 78–91. doi: 10.1016/j.jvb.2008.01.003
- Hakkarainen, K. (2009). A knowledge-practice perspective on technology-mediated learning. *Internat. J. Comp. Supp. Collab. Learn.* 4, 213–231. doi: 10.1007/s11412-009-9064-x
- Hargreaves, A. (2003). *Teaching in the knowledge society: Education in the age of insecurity*. New York, NY: Teachers College Press.
- Heiskanen, H., and Lonka, K. (2012). Are epistemological beliefs and motivational strategies related to study engagement in higher education? *Soc. Behav. Sci.* 2012, 306–313. doi: 10.1016/j.sbspro.2012.11.414
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemp. Educ. Psychol.* 25, 378–405. doi: 10.1006/ceps.1999.1026
- Hofer, B. K. (2001). Personal epistemology research: implications for learning and teaching. *Educ. Psychol. Rev.* 13, 353–383. doi: 10.1023/A:1011965830686
- Hofer, B. K. (2004). Exploring the dimensions of personal epistemology in differing classroom contexts: student interpretations during the first year of college. *Contemp. Educ. Psychol.* 29, 129–163. doi: 10.1016/j.cedpsych.2004.01.002
- Hofer, B. K. (2016). "Epistemic cognition as a psychological construct: Advancements and challenges," in *Handbook of epistemic cognition*, eds J. A. Greene, W. A. Sandoval, and I. Bråten (Milton Park: Routledge), 31–50.
- Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research* 67, 88–140. doi: 10.1111/medu.13928
- Hu, L., and Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equat. Mod.* 1, 1–55. doi: 10.1080/10705519909540118
- Joram, E., and Gabriele, A. J. (1998). Preservice teachers' prior beliefs: transforming obstacles into opportunities. *Teach. Teach. Educ.* 14, 175–191. doi: 10.1016/S0742-051X(97)00035-8
- Kember, D., and Kwan, K. P. (2000). "Lecturers' approaches to teaching and their relationship to conceptions of good teaching. *Instruct. Sci.* 28, 469–490. doi: 10.1055/s-0031-1285895
- Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *Br. J. Educ. Psychol.* 77, 229–243. doi: 10.1348/000709905X90344
- Lammasaari, H., Hietajärvi, L., Lonka, K., Chen, S., and Tsai, C. C. (2021). Teachers' epistemic beliefs and reported practices in two cultural contexts. *Educ. Stud.* 2021:2000369.
- Leithwood, K. A., Menzies, T., Jantzi, D., and Leithwood, J. (1999). "Teacher burnout: A critical challenge for leaders," in *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (Cambridge, MA: Cambridge University Press), 85–114. doi: 10.1017/cbo9780511527784.006
- Lindqvist, P., and Nordänger, U. (2006). Who dares to disconnect in the age of uncertainty? Teachers' recesses and 'off-the-clock'work. *Teach. Teach.* 12, 623–637. doi: 10.1080/13540600601029637
- Lonka, K. (1997). *Explorations of constructive processes in student learning. [Doctoral dissertation, University of Helsinki]*. Helsinki: University Press.
- Lonka, K. (2018). *Phenomenal learning from Finland*. Helsinki: Edita Publishing Ltd.
- Lonka, K., and Lindblom-Ylänne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher Education* 31, 5–24. doi: 10.1007/bf00129105
- Lonka, K., Joram, E., and Bryson, M. (1996). Conceptions of learning and knowledge: does training make a difference?. *Contemp. Educa. Psychol.* 21, 240–260. doi: 10.1006/ceps.1996.0021
- Lonka, K., Ketonen, E., and Vermunt, J. D. (2021). University students' epistemic profiles, conceptions of learning, and academic performance. *Higher Educ.* 81, 775–793. doi: 10.1007/s10734-020-00575-6
- Lonka, K., Ketonen, E., Vekkila, J., and Lara, M. (2019). "Doctoral students' writing profiles and their relations to well-being and perceptions of the academic environment. *Higher Educ.* 77, 587–602. doi: 10.1007/s10734-018-0290-x
- Lonka, K., Sharafi, P., Karlgrén, K., Masiello, I., Nieminen, J., Birgegård, G., et al. (2008). "MED NORD—A tool for measuring medical students' well-being and study orientations. *Med. Teach.* 30, 72–79. doi: 10.1080/01421590701769555
- Luthans, F. (2002). The need for and meaning of positive organizational behavior. *Internat. J. Industr.* 2002, 695–706. doi: 10.1002/job.165
- Luthans, F., and Avolio, B. J. (2003). Authentic leadership development. *Posit. Org. Scholar.* 2003:258.
- Luthans, F., Norman, S. M., Avolio, B. J., and Avey, J. B. (2008). The mediating role of psychological capital in the supportive organizational climate—employee performance relationship. *J. Org. Behav.* 2008, 219–238. doi: 10.1002/job.507
- Macey, W. H., and Schneider, B. (2008). The meaning of employee engagement. *Industr. Org. Psychol.* 1, 3–30.
- Maggioni, L., and Parkinson, M. M. (2008). The role of teacher epistemic cognition, epistemic beliefs, and calibration in instruction. *Educational Psychology Review* 20, 445–461. doi: 10.1007/s10648-008-9081-8
- Maggioni, L., VanSledright, B., and Alexander, P. (2009). Walking on the borders: a measure of epistemic cognition in history. *J. Exp. Educ.* 77, 187–214. doi: 10.3200/jexe.77.3.187-214
- Maslach, C., and Leiter, M. P. (1999). "Teacher burnout: A research agenda," in *Understanding and preventing teacher burnout: A sourcebook of international research and practice*, eds R. Vandenberghe and A. M. Huberman (Cambridge, MA: Cambridge University Press), 295–303. doi: 10.1016/j.jsp.2020.01.001
- Maslach, C., and Leiter, M. P. (2005). Reversing burnout. *Standford Soc. Innov. Rev.* 3, 43–49.
- Maslach, C., Jackson, S. E., and Leiter, M. P. (1996). *MBI: Maslach burnout inventory*. Sunnyvale, CA: CPP, Incorporated.
- Maslach, C., Schaufeli, W. B., and Leiter, M. P. (2001). Job burnout. *Ann. Rev. Psychol.* 52, 397–422.
- Morrison, D., and Collins, A. (1995). Epistemic fluency and constructivist learning environments. *Educ. Technol.* 35, 39–45.
- Muis, K. R., and Duffy, M. C. (2013). Epistemic climate and epistemic change: instruction designed to change students' beliefs and learning strategies and improve achievement. *J. Educ. Psychol.* 105:213. doi: 10.1037/a0029690
- Muis, K. R., Trevors, G., and Chevrier, M. (2016). *Epistemic climate for epistemic change in Handbook of epistemic cognition*. Milton Park: Routledge, 343–371.
- Näätänen, P., Aro, A., Matthiesen, S., and Salmela-Aro, K. (2003). *Bergen Burnout Indicator-15*. Helsinki: Edita.
- National Advisory Board on Research Ethics (2019). *The Ethical Principles of Research With Human Participants and Ethical Review in the Human Sciences in Finland*. Available online at: <https://tenk.fi/en/advice-and-materials> (accessed January 2022).
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *J. Curr. Stud.* 19, 317–328.
- OECD (2018). *Future of Education and Skills 2030: OECD Learning Compass 2030*. <http://www.oecd.org/education/2030-project> (accessed November 8, 2021)
- OECD (2019). *Learning Compass 2030*. https://www.oecd.org/education/2030-project/teaching-and-learning/learning-learning-compass-2030/OECD_Learning_Compass_2030_concept_note.pdf. (accessed November 8, 2021)

- Olafson, L., and Schraw, G. (2010). "Beyond epistemology: Assessing teachers' epistemological and ontological worldviews," in *Personal epistemology in the classroom: Theory, research and implications for practice*, eds L. D. Bendixen and F. C. Feucht (New York, NY: Cambridge University Press), 516–551. doi: 10.1017/cbo9780511691904.016
- Packer, M. J., and Goicoechea, J. (2000). Sociocultural and constructivist theories of learning: ontology, not just epistemology. *Educ. Psychol.* 35, 227–241. doi: 10.1207/s15326985ep3504_02
- Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62, 307–332. doi: 10.3102/00346543062003307
- Pillay, H., Goddard, R., and Wilss, L. (2005). Well-being, burnout and competence: implications for teachers. *Austral. J. Teach. Educ.* 2005:30.
- Pintrich, P. R., Marx, R. W., and Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Rev. Educ. Res.* 63, 167–199.
- Postareff, L., and Lindblom-Ylänne, S. (2008). Variation in teachers' descriptions of teaching: broadening the understanding of teaching in higher education. *Learn. Instruct.* 18, 109–120.
- Pyhältö, K., Pietarinen, J., and Salmela-Aro, K. (2011). Teacher–working-environment fit as a framework for burnout experienced by Finnish teachers. *Teach. Teach. Educ.* 27, 1101–1110.
- Retelsdorf, J., Butler, R., Streblo, L., and Schiefele, U. (2010). Teachers' goal orientations for teaching: associations with instructional practices, interest in teaching, and burnout. *Learn. Instruct.* 20, 30–46.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. *Handbook Res. Teach. Educ.* 2, 273–290.
- Richardson, V., and Placier, P. (2001). "Teacher change," in *Handbook of research on teaching*, 4th Edn, ed. V. Richardson (Washington DC: American Educational Research Association), 905–947.
- Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organisational change*. San Francisco: Jossey-Bass.
- Rosenholtz, S. J. (1989). Workplace conditions that affect teacher quality and commitment: implications for teacher induction programs. *Elem. Sch. J.* 89, 421–439. doi: 10.1086/461584
- Rosenholtz, S. J., and Simpson, C. (1990). Workplace conditions and the rise and fall of teachers' commitment. *Soc. Educ.* 63, 241–257.
- Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 36. doi: 10.3389/fpsyg.2014.01521
- RStudio Team (2020). *RStudio: Integrated Development for R: RStudio*. Boston, MA: PBC.
- Russell, R. F., and Stone, A. G. (2002). A review of servant leadership attributes: developing a practical model. *Leadership Org. Dev. J.* 23:3.
- Salmela-Aro, K., Hietajärvi, L., and Lonka, K. (2019). Work burnout and engagement profiles among teachers. *Front. Psychol.* 2019:2254. doi: 10.3389/fpsyg.2019.02254
- Salmela-Aro, K., Rantanen, J., Hyvönen, K., Tilleman, K., and Feldt, T. (2011). Bergen Burnout Inventory: reliability and validity among Finnish and Estonian managers. *Internat. Archiv. Occupat. Env. Health* 84, 635–645. doi: 10.1007/s00420-010-0594-3
- Schaufeli, W. B., and Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *J. Org. Behav.* 2004, 293–315.
- Schaufeli, W. B., and Buunk, B. P. (2003). Burnout: an overview of 25 years of research and theorizing. *Handbook Work Health Psychol.* 2, 282–424.
- Schaufeli, W. B., Bakker, A. B., and Salanova, M. (2006). The measurement of work engagement with a short questionnaire: a cross-national study. *Educ. Psychol. Meas.* 66, 701–716. doi: 10.1177/0013164405282471
- Schaufeli, W. B., Salanova, M., González-Romá, V., and Bakker, A. B. (2002). The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. *J. Happ. Stud.* 3, 71–92.
- Schaufeli, W., and Enzmann, D. (1998). *The burnout companion to study and practice: A critical analysis*. Boca Raton: CRC Press.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psychol.* 82:498. doi: 10.1037/0022-0663.82.3.498
- Schommer, M. (1993). Epistemological development and academic performance among secondary students. *J. Educ. Psychol.* 85:406.
- Schommer-Aikins, M. (2011). *Spontaneous cognitive flexibility and an encompassing system of epistemological beliefs in Links between beliefs and cognitive flexibility*. Dordrecht: Springer, 61–77.
- Sharplin, E., O'Neill, M., and Chapman, A. (2011). Coping strategies for adaptation to new teacher appointments: intervention for retention. *Teach. Teach. Educ.* 27, 136–146.
- Sinatra, G. M. (2016). "Thoughts on knowledge about thinking about knowledge," in *Handbook of epistemic cognition* (Milton Park: Routledge), 479–491.
- Sjöblom, K., Salmela-Aro, K., and Hietajärvi, L. (2020). Measuring broad self-regulatory skills in multi-locational knowledge work. *EAWOP Pract. E-J.* 2020, 16–41.
- Skaalvik, E. M., and Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: a study of relations. *Teach. Teach. Educ.* 26, 1059–1069.
- Smylie, M. A. (1999). "Teacher stress in a time of reform," in *Understanding and preventing teacher burnout: A sourcebook of international research and practice*, eds R. Vandenberghe and A. M. Huberman (Cambridge, MA: Cambridge University Press), 59–84.
- Trigwell, K., and Prosser, M. (1996a). Changing approaches to teaching: a relational perspective. *Stud. High. Educ.* 1996, 275–284.
- Trigwell, K., and Prosser, M. (1996b). Congruence between intention and strategy in university science teachers' approaches to teaching. *High. Educ.* 1996, 77–87.
- Tsai, C. C. (2007). Teachers' scientific epistemological views: the coherence with instruction and students' views. *Sci. Educ.* 91, 222–243.
- Upadaya, K., and Salmela-Aro, K. (2020). Social demands and resources predict job burnout and engagement profiles among Finnish employees. *Anxiety* 2020, 403–415. doi: 10.1080/10615806.2020.1746285
- Upadaya, K., Vartiainen, M., and Salmela-Aro, K. (2016). "From job demands and resources to work engagement, burnout, life satisfaction, depressive symptoms, and occupational health. *Burnout Res.* 3, 101–108. doi: 10.1016/j.burn.2016.10.001
- Vaara, L., and Lonka, K. (2014). What kind of leadership fosters pedagogically innovative school culture? *Eur. J. Soc. Behav. Sci.* 1, 1626–1637.
- Van Dierendonck, D., and Nuijten, I. (2011). The servant leadership survey: development and validation of a multidimensional measure. *J. Bus. Psychol.* 26, 249–267. doi: 10.1007/s10869-010-9194-1
- Vedenpää, I., and Lonka, K. (2014). Teachers' and teacher students' conceptions of learning and creativity. *Creat. Educ.* 5:1821. doi: 10.1177/1529100611418056
- Verquer, M. L., Beehr, T. A., and Wagner, S. H. (2003). A meta-analysis of relations between person–organization fit and work attitudes. *J. Vocat. Behav.* 63, 473–489. doi: 10.1016/s0001-8791(02)00036-2
- Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: an analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Rev. Educ. Res.* 72, 131–175.
- Xanthopoulou, D., Bakker, A. B., Demerouti, E., and Schaufeli, W. B. (2009). Reciprocal relationships between job resources, personal resources, and work engagement. *J. Vocat. Behav.* 74, 235–244.
- Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*, Vol. 30. Los Angeles: University of California.

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 Lammasaari, Hietajärvi, Salmela-Aro, Hakkarainen and Lonka. 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.



Aiding Reflective Navigation in a Dynamic Information Landscape: A Challenge for Educational Psychology

Katarzyna Bobrowicz*, Areum Han, Jennifer Hausen and Samuel Greiff

Computer-Based Assessment Group, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Simone Belli,
Complutense University of Madrid,
Spain

*Correspondence:

Katarzyna Bobrowicz
katarzyna.bobrowicz@uni.lu

Specialty section:

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

Received: 22 February 2022

Accepted: 12 April 2022

Published: 02 May 2022

Citation:

Bobrowicz K, Han A, Hausen J
and Greiff S (2022) Aiding Reflective
Navigation in a Dynamic Information
Landscape: A Challenge
for Educational Psychology.
Front. Psychol. 13:881539.
doi: 10.3389/fpsyg.2022.881539

Open access to information is now a universal phenomenon thanks to rapid technological developments across the globe. This open and universal access to information is a key value of democratic societies because, in principle, it supports well-informed decision-making on individual, local, and global matters. In practice, however, without appropriate readiness for navigation in a dynamic information landscape, such access to information can become a threat to public health, safety, and economy, as the COVID-19 pandemic has shown. In the past, this readiness was often conceptualized in terms of adequate literacy levels, but the contemporarily observed highest-ever literacy levels have not immunized our societies against the risks of misinformation. Therefore, in this Perspective, we argue that democratization of access to information endows citizens with new responsibilities, and second, these responsibilities demand readiness that cannot be reduced to mere literacy levels. In fact, this readiness builds on individual adequate literacy skills, but also requires rational thinking and awareness of own information processing. We gather evidence from developmental, educational, and cognitive psychology to show how these aspects of readiness could be improved through education interventions, and how they may be related to healthy work-home balance and self-efficacy. All these components of education are critical to responsible global citizenship and will determine the future direction of our societies.

Keywords: open access, COVID-19, 21st century skills, health literacy, critical literacy, statistical literacy, metacognition, rational thinking

INTRODUCTION

We live in the age of nearly universal access to information. With decentralized news outlets, growing access to open science, and worldwide social media coverage, individuals can be more broadly and diversely informed than ever before. Open access to information is a key value of modern democratic societies, as only thoroughly informed citizens can participate in society and make informed decisions about the directions in which they wish their society to evolve.

It seems, however, that despite open and multisource access to information, individuals fail to make thoroughly informed choices at both societal and individual levels. In this Perspective, we aim to examine why such failures may happen and how they could be remedied in education.

The ongoing COVID-19 crisis exposed, and perhaps augmented, society-wide difficulties with critical processing of dynamically changing information. This may have impeded many important outcomes such as sufficiently high vaccination rates, which, in turn, impose health, social, and economic losses on entire societies. Misinformation and other environmental factors outside of an individual's control have certainly contributed to such losses (Okuhara et al., 2020; Engledow and Weiland, 2021; Montagni I. et al., 2021). However, in this Perspective, we focus on the lack of individual readiness for navigating the dynamic information landscape, which may exacerbate poor decision-making under the pressing conditions of the prevailing global crisis. As educational psychologists, we suggest that equipping citizens with such readiness is a pressing challenge for contemporary education that needs to be targeted at all stages of individual development across the lifespan. We are certainly not alone in this view, as prolific literacy research has shown for several decades (e.g., Cervetti et al., 2006; Leu et al., 2017). Literacies are essential competencies, skills and dispositions that support individual comprehension and use of "information in daily activities to achieve goals, develop knowledge and potential" (cf. Organization Economic Cooperation Development [OECD], 2000, p. x; Leu et al., 2017). Since improving literacies is only part of the solution, we sketch out a more complex and comprehensive picture of the challenge ahead by coupling literacy research with relevant research on rational thinking and cognition. We argue that fostering responsible citizenship demands, at the very least, literacies, rational thinking (i.e., reflective, effortful processing of information), and awareness of our own information processing (i.e., using psychological knowledge to take stock of own mental operations; **Figure 1**). Throughout the paper, we recount evidence-based solutions developed within each of these research fields to offer concrete suggestions for education stakeholders.

OPEN ACCESS, LITERACIES, AND RATIONAL THINKING

Open Access Endows the Citizen With New Responsibilities

Openness and Participation are critical to democratic citizenship (Knight Commission Report, 2009; Şimşek and Şimşek, 2013). Openness implies that everybody can reach the information that supports civic and personal decisions, which serves to facilitate Participation, that is, joining and using information systems to solve civic and personal issues (Knight Commission Report, 2009). As open and multisource access to information secures these values, it cannot be realistically controlled by democratic governments (Eysenbach, 2008). This emancipation from top-down control forces individual citizens to assume responsibility

for their own navigation in the dynamic information landscape (Weiland, 2017). Citizens need to learn how to identify and critically process relevant information (Eysenbach, 2008), and how to position themselves in relation to a myriad of views, values, and ideas (Weiland, 2017). The COVID-19 crisis is one example that clearly showed that citizens were not ready for such navigation.

Soon before the onset of the crisis, vaccine hesitancy, a delay in acceptance, or a refusal of vaccine despite its availability (MacDonald and SAGE Working Group on Vaccine Hesitancy, 2015; Montagni I. et al., 2021), was listed among the 10 critical global threats by the World Health Organization [WHO] (2019), Okuhara et al. (2020). Hesitant attitudes toward vaccines became increasingly studied during the COVID-19 crisis, showing that they were associated with poor health literacy (Montagni I. et al., 2021; Turhan et al., 2021; in particular, digital health literacy, Patil et al., 2021) and a mismatch between affect-laden anti-vaccination messages and statistics-based governmental communication (Okuhara et al., 2020). In the cognitively and emotionally challenging "infodemic," understood as an overabundance of information on virus-related matters (Duplaga et al., 2019; World Health Organization [WHO], 2020; Patil et al., 2021), the citizens were left to their own devices. Despite higher-than-ever literacy rates, the public was poorly prepared for universal access to information and failed to assume the responsibilities that come with such access. Paired with poor readiness for selecting, evaluating, and processing information, universal access to information became a threat, not an aid.

Literacies Are Key but Insufficient

Universal access to information is here to stay but using it to individual and societal advantage rather than disadvantage demands better readiness, and it is better to foster it late than never. The notion that the age of access demands specific skills, and that educational systems across the world should support individuals in honing these skills, is not new. For over two decades now, educational psychologists have investigated a myriad of literacy skills (i.e., literacies) that support individual navigation in the dynamically changing information landscape (Cervetti et al., 2006; Şimşek and Şimşek, 2013; Leu et al., 2017). Literacy became a popular buzzword in educational psychology and an umbrella term for multiple, multimodal and multifaceted skills (Cervetti et al., 2006). Literacy for twenty-first century citizenship is digital, complex, and dynamic, but it builds on traditional literacies, such as reading, writing and comprehension (Leu et al., 2017). These traditional literacies need to be applied in the digital world, in which information is no longer provided in a single modality (e.g., audio only), a single flow (e.g., without simultaneous advertising), or by a limited group of individuals sharing similar values, views, and ideas. The line between the addresser and the addressee is fuzzy, as each member of the online community may smoothly move on the continuum between these roles. Therefore, twenty-first century literacies demand not only technical readiness for the modern technologies, but, critically, control over one's own information seeking and processing, awareness of the complexity of the social world (Cervetti et al., 2006; Leu et al., 2017), and a balance

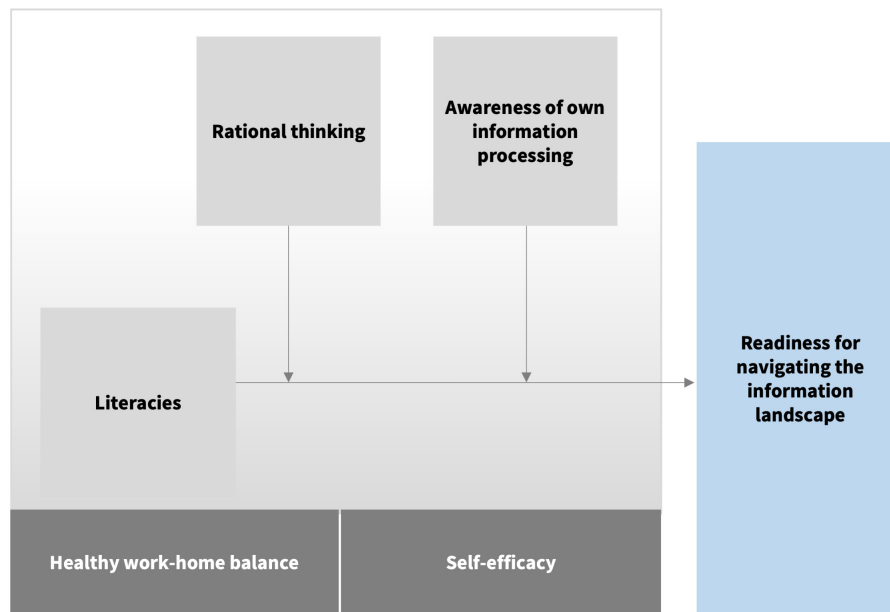


FIGURE 1 | The proposed model of the relationship between literacies and readiness for navigating the information landscape, moderated by rational thinking and awareness of own information processing. Here, literacies are defined as “competencies and dispositions that support individual comprehension and use of information in daily activities”; rational thinking as “reflective, effortful processing of information”; and awareness of own information processing as “using psychological knowledge to take stock of own mental operations”. Healthy work-home balance and self-efficacy are also components of the educational sphere that may contribute to readiness for navigating the information landscape.

between emotional engagement and emotional distance toward arguments presented by others (Gee, 2012; Olin-Scheller and Tengberg, 2017).

Technical readiness is typically a minor challenge for “digital natives” (Prensky, 2001; Valenza, 2006; Harris, 2008) or “insiders” (Lankshear and Knobel, 2006; Wimmer and Draper, 2019), who are children and adolescents raised amid rapidly evolving technologies from their birth (Neumann, 2016). However, evidence from educational and developmental psychology showed that digital natives, despite technical readiness, lack critical literacy and struggle with seeking, selecting, and evaluating information, and are not by default skilled or critical consumers of information (Valenza, 2006; Harris, 2008; Wineburg et al., 2016). Digital natives typically prioritize most accessible information sources and effortless processing (Shenton and Dixon, 2004; van Deursen et al., 2014; Loh and Kanai, 2015), and show age-specific difficulties in deploying attention to relevant information. Younger children (at 8–10 years) struggle with inhibiting irrelevant yet salient and engaging information (Eastin et al., 2006) and typically do not question its accuracy (Hirsh, 1999); older children lack sufficient knowledge base and analytical skills to contextualize and analyze the information, and often prioritize the form of the information over its content (Watson, 1998; Agosto, 2002; Sundar, 2008). This suggests that critical literacy learning should be adjusted to children’s cognitive developmental stage (Eastin, 2008) and fostered in education (Bernstein, 2000; Gee, 2012; Leu et al., 2017; Olin-Scheller and Tengberg, 2017) because it is central to the new responsibilities that follow from universal access to information.

Critical literacy, “the ability to argue from evidence, values and different perspectives (Skolverket., 2011) and to go beyond the individual and personal, and relate knowledge to more general and abstract notions (Bernstein, 2000; Gee, 2012),” as defined by Olin-Scheller and Tengberg (2017, p. 428), must be paired with age-appropriate statistical literacy and scientific literacy to prepare the prospective citizen for the influx of complex, heavily data-based information (Gal, 2002; Organization Economic Cooperation Development [OECD], 2004; Engel, 2017, 2019; Podkul et al., 2020; Engledowl and Weiland, 2021). To this end, in educational settings, complex information should not be simplified, but rather turned into manageable chunks and presented on multiple occasions (Spiro et al., 2003). Comparing poorly presented information, e.g., misleading data visualizations with well-executed visualizations provided by such tools as Engel (2019), Gapminder (2022) should be also a regular part of classroom activities (Engledowl and Weiland, 2021).

Critical literacy relies on cognitive skills, but it cannot be achieved without Students’ emotional investment. Educational psychologists showed that students need to learn how to self-regulate their own investment by striking a balance between emotional engagement and emotional distance toward their own and others’ arguments when discussing social issues (Gee, 2012; Olin-Scheller and Tengberg, 2017). Inhibiting own emotional attitudes is a prerequisite when evaluating own and others’ arguments and can be effectively trained in the classroom context, by, for instance, drawing Students’ attention to the affect-laden but irrelevant arguments that they present in response to a controversial issue (Olin-Scheller and Tengberg, 2017).

This requires metacognition, understood as active monitoring of own thoughts and emotions, and switching between own and others' perspectives (Flavell, 1976). Although this ability develops early (Marulis et al., 2016), it becomes an educational priority only in adolescence (Eysenbach, 2008; Olin-Scheller and Tengberg, 2017). In adolescence, citizens become increasingly autonomous, and drift away from reliance on traditional authority figures (i.e., intermediaries; Eysenbach, 2008) toward support from peers and less authoritative sources of information (i.e., apomedaries; Eysenbach, 2008). Individual autonomy continues to develop in adulthood, and is typically associated with gains in cognitive ability, literacies, and self-efficacy. Low capacity and/or insufficient belief in one's own capacity can push even adult citizens into reliance on traditional authority figures and uncritical, unreflective processing of information (Eysenbach, 2008). Therefore, critical literacy is only a bare minimum for individual readiness during global crises.

Literate Citizens Are Not Rational Thinkers

Navigating the information landscape demands critical literacy, paired with other relevant literacies, such as media literacy, health literacy, and statistical literacy. To date, various interventions have aimed to boost relevant literacies in individuals, perhaps stemming from a belief that a sufficiently literate individual will be able to scan, select, and evaluate the information landscape to make thoroughly informed choices. Empirical research findings on human rationality, however, suggest otherwise, showing that literacies are fundamental but not sufficient for individual readiness. Several models of human cognition showed that information processing roughly follows two separate routes: either a peripheral, fast route with mental shortcuts and simple rules (i.e., heuristics) or a central, cognitively effortful route with a reflective, metacognitive outlook on the available information (Sundar, 2008; Stanovich, 2009, 2018). Having access to these two processing routes (or systems, Stanovich, 2009) is highly adaptive, but overreliance on the heuristic route can lead to multiple cognitive biases (Sundar, 2008; Stanovich, 2009). Therefore, people who are sufficiently literate in terms of health and statistical knowledge may nevertheless suffer such cognitive biases (Stanovich, 2018) and make irrational choices as to whether, for example, to vaccinate themselves and/or their children against COVID-19 due to the anti-vaccination messages. The messages are typically communicated through affect-laden, salient imagery focused on vaccine toxicity and its side effects and consequently tend to trigger fast, heuristic processing (Okuhara et al., 2020). Even individuals that reflectively weigh the benefits and risks of the vaccine against the risks of COVID-19 infection may struggle with overriding the worry caused by the heuristic processing of such information (Okuhara et al., 2020).

As the above example showed, cognitive psychologists have repeatedly shown that individuals are prone to the same biases across social strata regardless of their literacy levels, but they may be able to override these biases with sufficient awareness and training (Stanovich, 2018). Instead of resorting to fast, automatic processing that prioritizes ease and speed

over rationality, individuals should be able to intentionally choose slow, reflective processing of the information landscape. Educating citizens on the multiple heuristics triggered by salient but irrelevant aspects of communication (Sundar, 2008) may be a good introduction to rational thinking. To some extent, training of rational thinking overlaps with literacy training, as they also involve technical components of statistical and scientific literacy (Stanovich, 2009). Such technical skills involve, among others, learning probability theory, switching between absolute and relative representations of data (e.g., 1,000 individuals vs. 25% of the tested sample vs. 1 in 4 tested people got infected) and overriding an automatic tendency to overestimate absolute numbers compared to percentages (Yamagishi, 1997; Stanovich, 2009). Thus far, however, it seems that citizens, lacking systematic training on this matter, have not had a chance to build an awareness of their own cognitive functioning. Without such chances at a society-wide level, many citizens will not be able to take control over their own cognitive functioning, and, thereby, will not achieve readiness for universal access to information. Their information processing will be governed externally, by those who manipulate vividness, accessibility, and salience of information (Stanovich, 2009), such as advertising companies in the modern market-based society. Thus far, the lack of awareness on own information processing had not threatened the economy, health, and survival, and had been overlooked in education. Now that it is generating considerable economic losses for the local and global communities, it will finally find its way to society-wide educational practices. We thus propose that developing a habit of taking stock of one's mental operations should become more prominent in the education of children, adolescents, and adults.

Better understanding of one's own information processing is a goal that demands concrete, structured action, with the concept of (ir)relevance of information at its core (Heine, 2006; Eysenbach, 2008; Hobbs and Jensen, 2009; Román et al., 2009; Olin-Scheller and Tengberg, 2017; Bajo et al., 2021). Identifying relevant information is key for critical and rational thinking, since changes of salient, irrelevant information can significantly impact human decision-making (i.e., framing effect; Stanovich, 2009). This skill has been championed by several scholars in recent years, pointing toward conditions in which students can efficiently take responsibility for handling access to information (Lankes, 2008; Weingarten, 2008; Rapp and McCrudden, 2018). Providing students with information of varying relevance and credibility, both in the classroom and in complex real-world contexts, gives them the best opportunity to determine on their own what matters regarding the information flow (Lankes, 2008; Weingarten, 2008; Rapp and McCrudden, 2018). Furthermore, delegating the responsibility for setting up own sub-goals, and selecting information without micromanaging instruction leads to more accurate relevance judgments (Heine, 2006). One way to achieve such tasks may be to construct self-regulated and metacognitive learning environments where students can figure out their own information processing with the help of technologies in learning analytics (e.g., Winne et al., 2019) and cooperation with interdisciplinary professionals. Completing such tasks may further build Students' confidence in their

own information processing abilities and steer them toward autonomous, critical thinking.

DISCUSSION

Over two centuries ago, a French mathematician argued during the French Revolution that citizens need *savoir liberateur*, knowledge built on information raising their awareness of the state of society (De Condorcet, 1994; Engel, 2019). The concept of a well-informed citizen is not new (e.g., Schutz, 1964; Strassheim, 2018), but it is, perhaps more than ever, pivotal for contemporary societies. The COVID-19 crisis showed that offering citizens universal access to information is insufficient, and, under some circumstances, detrimental to the common good. Global citizenship demands multiple literacies, rational thinking, and awareness of own information processing.¹ Although broad views of scientific literacy (Choi et al., 2011), critical statistical literacy (Engledowl and Weiland, 2021), and civic statistics (Engel, 2019) have incorporated metacognitive aspects of such literacies, neither rational thinking nor awareness of own information processing have been incorporated into twenty-first century skills (Dede, 2009). This demands better integration into curricula for school-aged children and adolescents, and informational interventions for adults.

Improving cognitive abilities or an overview of such abilities will not suffice when facing the challenges of the twenty-first century. Therefore, beyond the curricula and interventions at a broader level, we posit that readiness, relevance, and responsibility are important keywords for education in the 2020s, the decade of crucial decision-making that will shape the life of every individual in the global community. Given that readiness for responsible citizenship relies on rational thinking, and that overworked or emotionally burdened individuals may be more likely to engage in fast, heuristic processing rather than slow, effortful processing of continuously incoming information (cf. Gillard et al., 2009; Spiliopoulos et al., 2018), we suggest that children, adolescents, and adults are encouraged to replace unhealthy, workaholic habits with healthy work-home balance (e.g., Deloitte Consulting, 2010; Bannai and Tamakoshi, 2014; Anxo et al., 2017; Anxo and Karlsson, 2019). Therefore, unhealthy habits instilled in school-age children and adolescents through a tremendous amount of schoolwork should be limited to foster intentional deployment of cognitive resources from an early age. To our knowledge, this has not been investigated in the past, but future experimental studies could, for instance, longitudinally track adults on a spectrum of such habits and

their likelihood of using heuristic vs. reflective processing of information. Scholars from educational, developmental, and cognitive psychology need to assume this broad outlook on individual cognition, socioemotional skills and socioeconomic pressures to collaboratively and holistically foster responsible citizenship across all ages, from kindergartners to adults (see, e.g., Ellemers, 2021).

Literacy, rational thinking and work-life balance are vital parts of responsible citizenship, but it also requires sufficient self-efficacy that shields individuals against overreliance on authority figures (Eysenbach, 2008). Perceived lack of own impact on the society and non-meaningful, poorly paid jobs or unpaid internships cannot build such self-efficacy in the young citizens. Therefore, we urge adults to include young citizens in important decision-making, and to offer them meaningful, well-paid opportunities to contribute to society and to shape the world of tomorrow. Arguably, young citizens are leaving the educational system with better readiness for the contemporary digital world than older generations (Wimmer and Draper, 2019) and have high stakes in responsible, well-informed decision-making. However, without secure pay and a meaningful job, young citizens will not be able to achieve financial and intellectual autonomy critical to their participation in society.

Universal access to information is a double-edged sword. It gives all citizens a chance to shape their societies, but backfires without their readiness for responsible, well-reasoned choices. In this Perspective, based on evidence from educational, cognitive, and developmental psychology, we showed examples of concrete actions that can be taken by education stakeholders to foster citizen readiness. Since the twenty-first century globalization will certainly continue to scale up poor individual choices into international emergencies, citizen readiness, an essential remedy to such emergencies, should become a top priority for educational psychology.

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.

AUTHOR CONTRIBUTIONS

KB: conceptualization, investigation, writing—original draft, writing—review and editing, and artwork. AH: conceptualization, investigation, writing—review and editing, and artwork. JH and SG: conceptualization, investigation, and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

All authors were supported by the University of Luxembourg in the absence of external funding.

¹Global citizenship, in fact, demands more than education of individuals. In this Perspective, we have focused on such education, but we acknowledge that education of individuals must be coupled with broader environmental and societal changes. On the environmental level, for instance, citizens need external support in rational decision-making, e.g., by offering them prospectively, not temporarily, beneficial default choices (like pension funds or savings plans; Stanovich, 2009). On the societal level, for example, open and universal access to evidence-based information should be facilitated world-wide in order to alleviate post-colonial inequalities in such access (Santoro Lamelas and Belli, 2018).

REFERENCES

- Agosto, D. E. (2002). A model of young people's decision-making in using the web. *Libr. Inform. Sci. Res.* 24, 311–341.
- Anxo, D., and Karlsson, M. (2019). *Overtime Work A Review Of Literature And Initial Empirical Analysis. ILO Working Papers 995012793502676*.
- Anxo, D., Boulin, J.-Y., Cabrita, J., and Vermeylen, G. (2017). *Working Time Patterns for Sustainable Work*. Luxembourg: Publications Office of the European Union.
- Bajo, M. T., Gómez-Ariza, C. J., and Marful, A. (2021). Inhibitory control of information in memory across domains. *Curr. Dir. Psychol. Sci.* 30, 444–453. doi: 10.1177/09637214211039857
- Bannai, A., and Tamakoshi, A. (2014). The association between long working hours and health: a systematic review of epidemiological evidence. *Scand. J. Work Environ. Health.* 40, 5–18. doi: 10.5271/sjweh.3388
- Bernstein, B. (2000). *Pedagogy, Symbolic Control And Identity: Theory, Research, Critique*. Lanham, MD: Rowman & Littlefield Publishers.
- Cervetti, G., Damico, J., and Pearson, P. D. (2006). Multiple literacies, new literacies, and teacher education. *Theory Practice* 45, 378–386. doi: 10.1207/s15430421tip4504_12
- Choi, K., Lee, H., Shin, N., Kim, S., and Krajcik, J. S. (2011). Re-conceptualization of scientific literacy in south korea for the 21st century. *J. Res. Sci. Teach.* 48, 670–697. doi: 10.1002/TEA.20424
- De Condorcet, A. (1994). *Foundations of Social Choice and Political Theory*. Brookfield, VT: Edward Elgar.
- Dede, C. (2009). *Determining, Developing, and Assessing the Capabilities of "Future-Ready" Students*. Harvard University. Friday Institute White Paper Series, No 2.
- Deloitte Consulting (2010). "Annex 1 - study on health and safety aspects of working time," in *Study to support an impact assessment on further action at European level regarding directive 2003/88/EC and the evolution of working time organization, Report on behalf of the European Commission*. Available online at: https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_663072.pdf (accessed January 22, 2022).
- Duplaga, M., Sobiecka, K., and Wójcik, S. (2019). The reliability and validity of the telephone-based and online polish eHealth literacy scale based on two nationally representative samples. *Int. J. Environ. Res. Public Health* 16:3216. doi: 10.3390/ijerph16173216
- Eastin, M. S. (2008). "Toward a cognitive development approach to youth perceptions of credibility," in *Digital Media, Youth, And Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 29–48. doi: 10.1162/dmal.9780262562324.029
- Eastin, M. S., Yang, M. S., and Nathanson, A. (2006). Children of the net: an empirical exploration into the evaluation of internet content. *J. Broadcast. Electron. Media.* 50, 211–230.
- Ellemers, N. (2021). Science as collaborative knowledge generation. *Br. J. Soc. Psychol.* 60, 1–28. doi: 10.1111/bjso.12430
- Engel, J. (2017). Statistical literacy for active citizenship: a call for data science education. *Stat. Educ. Res. J.* 16, 44–49.
- Engel, J. (2019). "Statistical literacy and society," in *Actas Del Tercer Congreso Internacional Virtual De Educación Estadística*, eds J. M. Contreras, M. M. Gea, M. M. López-Martín, and E. Molina-Portillo.
- Engledowl, C., and Weiland, T. (2021). Data (Mis)representation and COVID-19: leveraging misleading data visualizations for developing statistical literacy across grades 6–16. *J. Stat. Data Sci. Educ.* 29, 160–164. doi: 10.1080/26939169.2021.1915215
- Eysenbach, G. (2008). "Credibility of health information and digital media: new perspectives and implications for youth," in *Digital Media, Youth, And Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 123–154. doi: 10.1162/dmal.9780262562324.123
- Flavell, J. H. (1976). "Metacognitive aspects of problem solving," in *The Nature Of Intelligence*, ed. J. B. Rensnick (Hillsdale, NJ: Lawrence Erlbaum Associates), 231–235.
- Gal, I. (2002). Adults' statistical literacy: meaning, components, responsibilities. *Intl. Stat. Rev.* 70, 1–25.
- Gapminder (2022). <https://www.gapminder.org> (accessed January 22, 2022).
- Gee, J. P. (2012). *Social Linguistics and Literacies. Ideology in Discourse*. London, NY: Routledge Falmer.
- Gillard, E., Van Dooren, W., Schaeken, W., and Verschaffel, L. (2009). Proportional reasoning as a heuristic-based process: time constraint and dual task considerations. *Exp. Psychol.* 56, 92–99. doi: 10.1027/1618-3169.56.2.92
- Harris, F. J. (2008). "Challenges to teaching credibility assessment in contemporary schooling," in *Digital Media, Youth, And Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 155–180. doi: 10.1162/dmal.9780262562324.155
- Heine, C. (2006). *Credibility in the Context of Searching. Unpublished White Paper*. Aurora, IL: Illinois Mathematics and Science Academy.
- Hirsh, S. (1999). Children's relevance criteria and information seeking on electronic resources. *J. Am. Soc. Inform. Sci.* 50, 1265–1283.
- Hobbs, R., and Jensen, A. (2009). The past, present, and future of media literacy education. *J. Media Lit. Educ.* 1, 1–11.
- Knight Commission Report (2009). *Knight Commission on the Information Needs of Communities in A Democracy: Sustaining Democracy in the Digital Age*. Washington, DC: The Aspen Institute.
- Lankes, R. D. (2008). "Trusting the internet: new approaches to credibility tools," in *Digital Media, Youth, and Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 101–122. doi: 10.1162/dmal.9780262562324.101
- Lankshear, C., and Knobel, M. (2006). *New Literacies: Everyday Practices And Classroom Learning*, 2nd Edn. New York: Open University Press.
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., and Henry, L. A. (2017). New literacies: a dual-level theory of the changing nature of literacy, instruction, and assessment. *J. Educ.* 197, 1–18. doi: 10.1177/002205741719700202
- Loh, K. K., and Kanai, R. (2015). How has the internet reshaped human cognition? *Neuroscientist* 22, 506–520. doi: 10.1177/1073858415595005
- MacDonald, N. E., and SAGE Working Group on Vaccine Hesitancy (2015). Vaccine hesitancy: definition, scope and determinants. *Vaccine* 33, 4161–4164. doi: 10.1016/j.vaccine.2015.04.036
- Marulis, L. M., Palincsar, A. S., Berhenke, A. L., and Whitebread, D. (2016). Assessing metacognitive knowledge in 3–5 year olds: the development of a metacognitive knowledge interview (McKI). *Metacogn. Learn.* 11, 339–368. doi: 10.1007/s11409-016-9157-7
- Montagni, I., Mebarki, A., Texier, N., Schück, S., Tzourio, S., and Confins group. (2021). Acceptance of a covid-19 vaccine is associated with ability to detect fake news and health literacy. *J. Public Health* 43, 695–702. doi: 10.1093/pubmed/fdab028
- Neumann, C. (2016). Teaching digital natives: promoting information literacy and addressing instructional challenges. *Reading Improvement* 53, 101–106.
- Okuhara, T., Ishikawa, H., Okada, H., Ueno, H., and Kiuchi, T. (2020). Dual-process theories to counter the anti-vaccination movement. *Prev. Med. Rep.* 20, 1–3. doi: 10.1016/j.pmedr.2020.101205
- Olin-Scheller, C., and Tengberg, M. (2017). Teaching and learning critical literacy at secondary school: the importance of metacognition. *Lang. Educ.* 31, 418–431. doi: 10.1080/09500782.2017.1305394
- Organization Economic Cooperation Development [OECD] (2000). *Literacy in the Information Age. Final Report of the International Adult Literacy Survey*. Available online at: <https://www.oecd.org/education/skills-beyond-school/41529765.pdf> (accessed January 22, 2022).
- Organization Economic Cooperation Development [OECD] (2004). "Scientific literacy," in *The Routledge Falmer Reader in Science Education*, ed. J. Gilbert (London and New York: Routledge Falmer, Taylor and Francis Group), 39–52.
- Patil, U., Kostareva, U., Hadley, M., Manganello, J. A., Okan, O., Dadaczynski, K., et al. (2021). Health literacy, digital health literacy, and COVID-19 pandemic attitudes and behaviors in U.S. college students: implications for interventions. *Int. J. Environ. Res. Public Health* 18:3301. doi: 10.3390/ijerph18063301
- Podkul, A., Vittet, L., Tranter, S., and Alduncin, A. (2020). The coronavirus exponential: a preliminary investigation into the public's understanding. *Harvard Data Sci. Rev. (Special issue)*. doi: 10.1162/99608f92.fec69745
- Prensky, M. (2001). Digital natives, digital immigrants. *Horizon* 9, 1–6.
- Rapp, D. N., and McCrudden, M. T. (2018). "Relevance before, during, and after discourse experiences," in *Relevance and Irrelevance: Theories, Factors, and Challenges*, eds J. Strassheim and H. Nasu (Berlin: De Gruyter), 141–160.
- Román, P., Soriano, M. F., Gómez-Ariza, C. J., and Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychol. Sci.* 20, 1053–1058.
- Santoro Lamelas, V., and Belli, S. (2018). Digital communication in scientific practices: limitations and possibilities for equality. *Prisma Soc.* 20, 316–332.

- Schutz, A. (1964). "The well-informed citizen: an essay on the social distribution of knowledge," in A. Schutz: *Collected Papers 2: Studies in Social Theory*, ed. A. Brodersen (Hague: Martinus Nijhoff), 120–134.
- Shenton, A. K., and Dixon, P. (2004). Issues arising from youngsters' information-seeking behavior. *Libr. Inf. Sci. Res.* 26, 177–200. doi: 10.29173/lirg172
- Şimşek, E., and Şimşek, A. (2013). New literacies for digital citizenship. *Contemp. Educ. Technol.* 4, 126–137. doi: 10.30935/CEDTECH/6097
- Skolverket. (2011). *Läroplan för Grundskolan, Förskolklassen Och Fritidshemmet 2011 (Curricula for Compulsory School, Pre-School And After-School Activities 2011)*. Stockholm, Sweden: Skolverket & Fritzes.
- Spiliopoulos, L., Ortmann, A., and Zhang, L. (2018). Complexity, attention, and choice in games under time constraints: a process analysis. *J. Exp. Psychol. Learn. Mem. Cogn.* 44, 1609–1640. doi: 10.1037/xlm0000535
- Spiro, R. J., Collins, B. P., and Thota, J. J. (2003). Cognitive flexibility theory: hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educ. Techn.* 43, 5–10.
- Stanovich, K. E. (2009). *What Intelligence Tests Miss: The Psychology Of Rational Thought*. New Haven, CT: Yale University Press.
- Stanovich, K. E. (2018). How to think rationally about world problems. *J. Intell.* 6:25. doi: 10.3390/jintelligence6020025
- Strassheim, J. (2018). "Relevance and irrelevance," in *Relevance and Irrelevance: Theories, Factors and Challenges*, eds J. Strassheim and H. Nasu (Berlin, Boston: De Gruyter), 1–18. doi: 10.1515/9783110472509-001
- Sundar, S. S. (2008). "The MAIN model: a heuristic approach to understanding technology effects on credibility," in *Digital Media, Youth, and Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 123–154. doi: 10.1162/dmal.9780262562324.073
- Turhan, Z., Dilcen, H. Y., and Dolu, I. (2021). The mediating role of health literacy on the relationship between health care system distrust and vaccine hesitancy during COVID-19 pandemic. *Curr. Psychol.* [Epub ahead of print]. doi: 10.1007/s12144-021-02105-8
- Valenza, J. K. (2006). They might be gurus. *Teacher Librarian* 34:18.
- van Deursen, A. J., Görzig, A., van Delzen, M., Perik, H. T. M., and Stegeman, A. G. (2014). Primary school children's internal skills: a report on performance tests of operational, formal, information, and strategic internal skills. *Int. J. Commun.* 8, 1343–1365.
- Watson, J. S. (1998). If you don't have it, you can't find it." a close look at students' perceptions of using technology. *J. Am. Soc. Inform. Sci.* 49, 1024–1036.
- Weiland, T. (2017). Problematising statistical literacy: an intersection of critical and statistical literacies. *Educ. Stud. Math.* 96, 33–47. doi: 10.1007/S10649-017-9764-5
- Weingarten, F. W. (2008). "Credibility, politics, and public policy," in *Digital Media, Youth, And Credibility*, eds M. J. Metzger and A. J. Flanagin (Cambridge, MA: The MIT Press), 181–202. doi: 10.1162/dmal.9780262562324.181
- Wimmer, J. J., and Draper, R. J. (2019). Insiders' views of new literacies, schooling, and the purpose of education: "we should be teaching them more important things". *Read. Psychol.* 40, 149–168. doi: 10.1080/02702711.2019.1607000
- Wineburg, S., McGrew, S., Breakstone, J., and Ortega, T. (2016). *Evaluating Information: The Cornerstone Of Civic Online Reasoning*. Available online at: <http://purl.stanford.edu/fv751yt5934> (accessed January 22, 2022).
- Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., et al. (2019). nStudy: software for learning analytics about learning processes and self-regulated learning. *J. Lear. Anal.* 6, 95–106. doi: 10.18608/jla.2019.6.27
- World Health Organization [WHO] (2019). *Ten threats to global health in 2019*. Available online at: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019> (accessed January 22, 2022).
- World Health Organization [WHO] (2020). *Managing The COVID-19 Infodemic: Promoting Healthy Behaviours and Mitigating The Harm From Misinformation And Disinformation*. A joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC. Geneva.
- Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: implications for risk communication. *Appl. Cogn. Psychol.* 11, 495–506. doi: 10.1002/(SICI)1099-0720(199712)11:6<495::AID-ACP481<3.0.CO;2-J

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 Bobrowicz, Han, Hausen and Greiff. 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.



Corrigendum: Aiding Reflective Navigation in a Dynamic Information Landscape: A Challenge for Educational Psychology

Katarzyna Bobrowicz*, Areum Han, Jennifer Hausen and Samuel Greiff

Computer-Based Assessment Group, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Keywords: open access, COVID-19, 21st century skills, health literacy, critical literacy, statistical literacy, metacognition, rational thinking

A Corrigendum on

Aiding Reflective Navigation in a Dynamic Information Landscape: A Challenge for Educational Psychology

by Bobrowicz, K., Han, A., Hausen, J., and Greiff, S. (2022). *Front. Psychol.* 13:881539. doi: 10.3389/fpsyg.2022.881539

OPEN ACCESS

Approved by:
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

***Correspondence:**
Katarzyna Bobrowicz
katarzyna.bobrowicz@uni.lu

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

Received: 08 May 2022
Accepted: 17 May 2022
Published: 09 June 2022

Citation:
Bobrowicz K, Han A, Hausen J and
Greiff S (2022) Corrigendum: Aiding
Reflective Navigation in a Dynamic
Information Landscape: A Challenge
for Educational Psychology.
Front. Psychol. 13:939190.
doi: 10.3389/fpsyg.2022.939190

In the original article, there was an error in Introduction, Paragraph 1. “We in the age of nearly universal access to information. With decentralized news live outlets” should read “We live in the age of nearly universal access to information. With decentralized news outlets”. A correction has been made to Introduction, Paragraph 1:

We live in the age of nearly universal access to information. With decentralized news outlets, growing access to open science, and worldwide social media coverage, individuals can be more broadly and diversely informed than ever before. Open access to information is a key value of modern democratic societies, as only thoroughly informed citizens can participate in society and make informed decisions about the directions in which they wish their society to evolve. It seems, however, that despite open and multisource access to information, individuals fail to make thoroughly informed choices at both societal and individual levels. In this Perspective, we aim to examine why such failures may happen and how they could be remedied in education.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

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 Bobrowicz, Han, Hausen and Greiff. 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 the Relationship Among English Language Learning Strategies, Language Achievement, and Attitude

Anita Habók^{1,2*}, Andrea Magyar² and Gyöngyvér Molnár^{1,2}

¹Institute of Education, University of Szeged, Szeged, Hungary, ²MTA–SZTE Digital Learning Technologies Research Group, Szeged, Hungary

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the Americas
– Nevis, United States

Reviewed by:

Rining (Tony) Wei,
Xi'an Jiaotong-Liverpool
University, China
Reza Kafipour,
Shiraz University of Medical Sciences,
Iran
Fidel Çakmak,
Alanya Alaaddin Keykubat University,
Turkey

*Correspondence:

Anita Habók
habok@edpsy.u-szeged.hu

Specialty section:

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

Received: 01 February 2022

Accepted: 15 April 2022

Published: 13 May 2022

Citation:

Habók A, Magyar A and
Molnár G (2022) Investigating the
Relationship Among English
Language Learning Strategies,
Language Achievement, and Attitude.
Front. Psychol. 13:867714.
doi: 10.3389/fpsyg.2022.867714

The main objective of the study was to ascertain whether English as a Foreign Language learners with various levels of English language achievement differ significantly in applying foreign language learning strategies. We also aimed to explore strategy use frequency in connection with attitude toward English language learning. Data were collected from 1,653 lower secondary students in Hungary through a revised version of the previously developed online Self-Regulated Foreign Language Learning Strategy Questionnaire (SRFLLSQ) based on Oxford's Strategic Self-Regulation (S²R) Model. The findings point to statistically significant differences in the frequency of English language strategy use among more and less proficient learners. Quantitative analyses also reported that, in spite of the students stated low or moderate levels of strategy use, it turned out as a statistically significant predictor of foreign language attitude and language achievement. The results draw attention to the relevance of strategy research in foreign language teaching as well as encourages strategy teaching within language instruction.

Keywords: self-regulated foreign language learning, language learning strategies, foreign language attitude, language achievement, lower secondary students

INTRODUCTION

Foreign language learning requires many underlying skills and techniques. Learners have to master a number of complex linguistic, personal, cultural and social skills, and competences and be aware of effective techniques and strategies to help them cope with various challenges during the learning process. The frequent use of learning strategies can help learners to become more competent and effective language users in the classroom and inspire them to achieve higher levels of mastery in the target foreign language (Wong and Nunan, 2011; Oxford, 2016). Since the mid-1970s, an immense amount of learning strategy research has attempted to establish the concept and identify strategies that help learners to become more effective language learners (Oxford, 1990; Cohen, 1998). It is a widely researched and highly debated area even today (Thomas and Rose, 2019; Thomas et al., 2021). The most well-known and widely used taxonomy of language learning strategies (LLS) was developed by Oxford (1990, 2011, 2016). In her recently reconsidered Strategic Self-Regulation (S²R) Model based on Vygotsky's (1978) sociocultural

theory of self-regulated learning (SRL) and Zimmerman's three-phase model (Zimmerman, 2000; Zimmerman and Schunk, 2011), she identified four main strategy categories: cognitive, affective, motivational, and social, each of them guided by the master category of "meta-strategies." These meta-strategies are metacognitive, meta-affective, metamotivational, and metasocial strategies, respectively (Oxford, 2016).

Oxford also developed a measurement tool (Strategy Inventory for Language Learning, SILL) for investigating LLS use, which is employed worldwide; however, it is based on her original conceptualization. Nevertheless, it is essential to relate the latest pedagogical theories to language teaching practice. Self-regulation theory, which was the basis for Oxford's new taxonomy, has been dominant since the beginning of this century. It is thus crucial to develop state-of-the-art measurement tools which can be used in the classroom by language teachers. In previous research, we have developed and validated a questionnaire based on Oxford's S²R Model (SRFLLSQ; Habók and Magyar, 2018b). To obtain a more comprehensive view of the role and possible effect of language learning strategies on certain other factors, such as attitude, motivation, and self-efficacy, it is essential to conduct further research. In this study, we aimed to examine LLS in relation to other crucial factors of language learning; we have investigated the relationships among the application of English language learning strategies, language achievement, and attitude toward English among lower secondary students in Hungary.

LITERATURE REVIEW

The Concept of Language Learning Strategies

Language learning strategies have been a research focus since the mid-1970s, as strategic language learning is a key to successfully acquiring a foreign language (Rubin, 1975). A number of definitions of LLS have emerged, with one of the most influential having proved to be that of Rebecca Oxford, who not only established a conceptualization, but also conducted empirical research. In her content-analytic study, Oxford involved 33 distinct definitions and interpretations of the term LLS and thus determine it as follows:

L2 learning strategies are complex, dynamic thoughts, and actions, selected and used by learners with some degree of consciousness in specific contexts in order to regulate multiple aspects of themselves (such as cognitive, emotional, and social) for the purpose of (a) accomplishing language tasks; (b) improving language performance or use; and/or (c) enhancing long-term proficiency. Strategies are mentally guided but may also have physical and therefore observable manifestations. Learners often use strategies flexibly and creatively; combine them in various ways, such as strategy clusters or strategy chains; and orchestrate them to meet learning needs. Strategies are teachable. Learners in their contexts decide which strategies to use. Appropriateness of

strategies depends on multiple personal and contextual factors (Oxford, 2016, p. 48).

Strategic language learners select their LLS according to their personal preferences, motivational intentions, and situational circumstances. Therefore, it is especially difficult to identify a system for these strategies. This is one of the reasons why an extremely large number of conceptualizations and debates have emerged (Thomas and Rose, 2019; Thomas et al., 2021). Thomas et al. (2021) have pointed out that with the emphasis on self-regulation, the field of strategy research has moved away from formal educational settings toward learner autonomy. They argue that this is a hazardous trend because definitions of LLS minimize teachers' role and classroom contexts that can also be an influencing factor in strategic behavior among students. Thomas and Rose (2019) have therefore suggested a separation of LLS from self-regulation and introduced the Regulated Language Learning Strategies Continuum to make it clear that self-regulation can be conceptually separated in defining LLS. By interpreting LLS as being both other- and self-regulated, Dörnyei and Skehan (2003) argue that strategy use cannot be regarded as either emotional or cognitive or even behavioral, thus opening up another debated area in the field.

The classification of LLS is another key area of argument. Oxford's original classification of six major fields (memory, cognitive, metacognitive, affective, compensation, and social strategies) was recently reconsidered and restructured on the basis of self-regulation theories. Her Strategic Self-Regulation (S²R) Model was developed based on Vygotsky's (1978) sociocultural theory of self-regulated learning (SRL). In her concept, she identified four main fields—cognitive, affective, motivational, and social strategies—each of them directed by a "master category of meta-strategies." These meta-strategies are metacognitive, meta-affective, metamotivational, and metasocial strategies (Oxford, 2016). Her taxonomy was again open to a number of debates as some theorists (Dörnyei, 2005; Hajar, 2019) argued that success in language learning cannot be assessed through the frequency of strategy use alone.

Research on Language Learning Strategies

Despite the debates, LLS researchers have been devoted to conducting quantitative research on LLS use and its connection with other individual factors, such as gender, learning style, motivation, attitude, and proficiency (e.g., Radwan, 2011; Alhaysony, 2017; Habók and Magyar, 2018a, 2019). The most widespread measurement tool for assessing L2 learners' strategy use is Oxford's Strategy Inventory for Language Learning (SILL; Oxford, 1990). This questionnaire has been translated into numerous languages and adapted for a vast number of cultures around the world. It is based on her original taxonomy and employs her original six strategy fields. Based on her renewed taxonomy, a number of reconsidered measurement tools have been developed since then, which have approached effective language learning from different perspectives (Wang et al.,

2013; Salehi and Jafari, 2015; Božinović and Sindik, 2017; Köksal and Dündar, 2017; Habók and Magyar, 2018b; An et al., 2021).

One major area for researchers has been to find out what set of strategies foreign language learners rely on the most (Platsidou and Sipitanou, 2015; Alhaysony, 2017; Charoento, 2017; Dawadi, 2017; Habók and Magyar, 2018a,b, 2019, 2020; Habók et al., 2021). Overall, results have concluded moderate use of LLS among participants. The most frequently used strategies are cognitive, metacognitive, and compensation strategies, while affective and memory strategies are the least preferred. Habók et al. (2021) have pointed out the different strategy preferences in different cultural contexts. Their results reinforced the preferred use of metacognitive strategies in both European and Asian contexts. However, there were statistically significant differences in the affective field with regard to the lower strategy use preference in the European sample.

A great deal of research has investigated strategy use in connection with other aspects (Platsidou and Kantaridou, 2014; Rao, 2016; Charoento, 2017; Habók and Magyar, 2018a, 2020). One of the most often used factors was language achievement, which has been specified and covered in a multitude of ways. Some studies have focused on level of language proficiency or achievement test results (Raofi et al., 2017; Taheri et al., 2019; An et al., 2021; Malpartida, 2021), others have relied on self-ratings (Charoento, 2017), and still others have involved language course marks (Habók and Magyar, 2018a; Sánchez, 2019; Bećirović et al., 2021). As a result, most research has pointed out that students with higher proficiency use LLS more frequently than those with less (Rao, 2016; Charoento, 2017; Raofi et al., 2017; Sánchez, 2019). Charoento (2017) highlighted that successful students mainly used metacognitive strategies and less proficient students preferred to use social strategies the most. Sánchez (2019) reported that the application of social, metacognitive, and cognitive strategies was the highest among high achievers. Some research failed to find any significant differences in LLS use between learners with lower and higher English proficiency levels (Rianto, 2020).

A relatively small number of studies have examined how LLS use predicts language proficiency. Some research has pointed out a positive correlation between strategy use and proficiency. Comprehensive work by Taheri et al. (2019) indicated a statistically significant correlation between LLS and second language achievement. Specifically, they confirmed a statistically significant relationship between cognitive, social, and compensation strategies and second language achievement. Platsidou and Kantaridou (2014) also found that language use is predicted by learning strategy use in a statistically significantly way and that it anticipates perceived language performance. Rao (2016) also reinforced that learners' English proficiency level determines their strategy use and highlighted that students with high proficiency use strategies more frequently than low-level learners. Habók and Magyar (2018a) stated that strategies have a statistically significant effect on proficiency through attitudes. In addition, these effects reflect general school achievement. Bećirović et al. (2021) observed that LLS can influence students' English as a foreign language (EFL) achievement. Specifically, cognitive strategies have a statistically significant positive effect on EFL achievement, while

other strategies showed negative or no significant impact. An et al. (2021) also reported the positive direct effect of SRL strategies on students' English learning outcomes. Lin et al. (2021) concluded the direct impact of learning strategies on learners' performance in literal and inferential comprehension.

Another research direction is the investigation of strategy use in relation to other underlying factors, like affective variables, such as motivation, attitude, self-efficacy, and self-concept. Educational research has pointed out that learners' attitude toward language learning is crucial since it can greatly impact learning results and language learning proficiency (Platsidou and Kantaridou, 2014). Studies have indicated that learners with a positive attitude toward language learning employ LLS more frequently and effectively. Platsidou and Kantaridou (2014) used confirmatory factor analysis to show that attitudes toward language learning predict the use of both direct and indirect learning strategies. Jabbari and Golkar (2014) reported a more frequent use of cognitive, metacognitive, compensation, and social strategies among students with a positive attitude toward language learning. Habók and Magyar (2018a) demonstrated the reverse effect: learners who apply LLS effectively were more likely to have improved learning experiences and positive attitudes toward foreign language learning.

It can be concluded that strategy research is often related to other vital areas of language learning, among which attitude plays an important role. However, only a limited number of researchers have developed measurement tools for investigating self-regulated foreign LLS and measured it in relation to attitude. In addition, most studies have focused on the strategy use of tertiary samples with relatively high levels of proficiency. This study aims to fill this gap and provides an insightful investigation into the connections among strategy use, attitude, and English language achievement among lower secondary students. Based on the relevant literature (Jabbari and Golkar, 2014; Platsidou and Kantaridou, 2014; Habók and Magyar, 2018a), we hypothesized the statistically significant effect of LLS on proficiency through attitude.

RESEARCH QUESTIONS

The research addresses the following three research questions:

1. Which EFL strategy was the most frequently used by 11-year-old lower secondary students?
2. Were there statistically significant differences among students in their language learning strategy use on the basis of their English language achievement?
3. Which language learning strategy type has a statistically significant impact on learners' English language achievement and attitude?

RESEARCH METHODS

Participants

In Hungary, students start primary school at the age of six. This lasts 4 year. Then, they continue their studies at the lower

secondary level. At the age of 14, they move on to upper secondary school. The participants of the present research were 11-year-old lower secondary students in Grade 5 ($N_{\text{total}} = 1,653$; $N_{\text{boys}} = 827$, $N_{\text{girls}} = 780$, $N_{\text{missing}} = 46$) from 64 schools in Hungary. Hungarian students take EFL in compulsory courses in school, and they usually start learning a foreign language at the age of nine. However, in some schools, they can start at the age of six. Typically, they can choose between English and German, but recently a preference for English has become more common. In upper secondary school, two foreign languages are compulsory, English and German or Italian or Spanish. The second language depends on curricular choice at the school level.

The English proficiency of the participating students was at beginner/elementary level (A1–A2). As for their engagement in learning, there were 17 students in the sample who spent 2 h or less per week on English. Around half of the learners ($N = 884$) devoted 3 h a week to this subject, and few participants dedicated four ($N = 303$) or five ($N = 357$) hours a week to the language. We also found 67 students who dealt with English six or more hours per week. In addition, we did not receive any answers to this question from 25 students.

Instrument

The revised and improved version of the Self-Regulated Foreign Language Learning Strategy Questionnaire (SRFLLSQ) was employed after our first measurement and validation (Habók and Magyar, 2018b). We reviewed the new findings on the theoretical background to foreign LLS research and continued revising the affective field. In addition, based on the relevant literature, we included the field of motivation in the questionnaire. We thus completed the measurement tool with motivational and metamotivational factors based on Oxford's Strategic S²R Model. Finally, the questionnaire covered four strategy areas: metacognitive (eight items), cognitive (six items), meta-affective (eight items), affective (eight items), metasocial (eight items), social (six items), motivational (four items), and metamotivational (four items; see **Appendix**). The fifth-grade students provided their responses on a five-point Likert scale, which ranged from 1 ("Never or almost never true of me") to 5 ("Always or almost always true of me"). The measurement tool was also complemented with a background questionnaire, which asked students about their foreign language school marks, which indicated students' English language achievement (1 = fail, lowest school mark; 5 = excellent, highest school mark). Students also self-reported their attitudes toward English learning on a five-point Likert scale, which again ranged from 1 to 5.

Procedure

First, the research was accepted by the IRB at the University of Szeged Doctoral School of Education. It was concluded that the research design follows IRB recommendations. The participating learners' parents were asked for written informed consent, which was handled by the participating schools. Second, an invitation was sent to schools to register for the measurement. In the call, schools were given information about the purpose of the measurement. Once the schools accepted the invitation,

they received further instructions on data collection and a link to log into the Online Diagnostic Assessment System (eDia), which is developed, supervised, and operated by the University of Szeged Centre for Research on Learning and Instruction (Csapó and Molnár, 2019). Students' participation was voluntary in the research. They logged into the system with an official student assessment code (developed by the Hungarian Educational Authorities), which provided complete anonymity for them. The researchers could not identify the respondents on this basis. The identification code was handled by the school administrators, but the students' results were not available to them. Thus, complete anonymity was guaranteed. The eDia system is familiar to students because they regularly use it for diagnostic purposes during official school hours. The students had already acquired the necessary ICT skills at primary level, further improved through remote learning. For the present questionnaire, the participants indicated their responses by clicking on radio buttons. The learners were given a school lesson in a personal classroom environment provided by the school. After logging in, the respondents filled in the questionnaire in Hungarian, which is their native language, because they do not yet have the foreign language skills to provide reliable answers in English. This took 20 min to complete the instrument. Teacher assistance was not required while the questionnaire was being completed, but it was available. The students had the option to ask for assistance on any technical problems.

Data Analysis

First, we employed classical test analysis and examined reliability, means, and standard deviations for the questionnaire fields with SPSS Statistics 23.0. In the case of frequency of strategy use, we aimed to find out how strategy use was perceived by our sample. We also compared the students' strategy use vis-à-vis their English language achievement and attitude using an independent sample *t*-test. To interpret effect size, we followed Wei et al.'s (2019) and Wei and Hu's (2019) benchmark: under 0.005 is small, 0.01 is typical or medium, 0.02 is large, and is 0.09 very large. We used R^2 unsquared; thus, the benchmark for the effect size index is 0.07, 0.10, 0.14, and 0.30, which, respectively, represents small, medium, large, and very large cut-off values. We applied path analysis to map the possible relationships and effects of our variables. We studied the goodness-of-fit indices by applying various cut-off values for many fit indices, including the Tucker–Lewis index (TLI), the normed fit index (NFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and Chi-square values (Kline, 2015). TLI, NFI, and CFI were regarded as eligible with a cut-off value of 0.95, and RMSEA values indicated an acceptable fit of 0.8 (Kline, 2015).

RESULTS

Descriptive Analysis

The questionnaire was reliable in all the fields (**Table 1**). As regards the whole sample, moderate strategy use was observed. The lowest strategy use was noted in the field of metasocial

strategies, and the most frequent strategy was found in the affective field. As regards the corresponding factors, the most frequent use was observed in the motivational field (see **Table 1**).

We also identified the frequency of strategy use among the more and less proficient learners. Students were divided into two categories based on their English language achievement (**Table 2**). Those with excellent and good school marks were placed in the more proficient learners' category, while learners with average, fair, or unsatisfactory school marks were grouped into the less proficient learner category. Students ($N=810$) who received excellent school marks met the requirements of the English language curriculum and performed at a high level. Learners ($N=500$) with good marks had minor gaps. Those ($N=229$) with an average school mark had major gaps in their knowledge, and those ($N=65$) with unsatisfactory school marks had difficulty following the curriculum and varying levels of difficulty in all areas of language learning. Finally, students ($N=9$) who received an unsatisfactory school mark are at a disadvantage which is difficult to overcome. No data were received from 40 students. Students' EFL achievement could be regarded as good with a mean of 4.2 ($SD=0.89$). As a result, the more proficient learners employed strategies with greater frequency in all of the fields, a statistically significant finding. The affective factor indicated above medium effect size, while the other factors reported small effect sizes.

TABLE 1 | Frequency of language learning strategy use in Grade 5.

Fields	Crb alpha	<i>M</i> (SD)	<i>M</i> (SD)
Metacognitive	0.79	3.47(0.74)	3.46(0.70)
Cognitive	0.72	3.43(0.78)	
Meta-affective	0.73	3.28(0.77)	3.55(0.72)
Affective	0.83	3.82(0.81)	
Metasocial	0.88	3.19(0.98)	3.28(0.92)
Social	0.85	3.41(0.94)	
Metamotivational	0.76	3.45(0.98)	3.60(0.84)
Motivational	0.66	3.75(0.90)	

TABLE 2 | Frequency of language learning strategy use among less and more proficient learners.

Fields	Less proficient learners <i>M</i> (SD)	More proficient learners <i>M</i> (SD)	<i>t</i>	<i>r</i> (effect size)
Metacognitive	3.02(0.71)	3.58(0.71)	-12.21*	0.085
Cognitive	3.07(0.74)	3.51(0.76)	-9.28*	0.050
Meta-affective	3.07(0.77)	3.33(0.76)	-5.24*	0.017
Affective	3.26(0.79)	3.96(0.76)	-13.85*	0.113
Metasocial	2.82(0.88)	3.27(0.98)	-7.30*	0.032
Social	2.98(0.86)	3.50(0.93)	-8.78*	0.046
Metamotivational	3.14(0.96)	3.51(0.98)	-5.91*	0.021
Motivational	3.32(0.97)	3.84(0.86)	-9.34*	0.052

*Differences are statistically significant at $p < 0.001$ level.

Multivariate Analyses

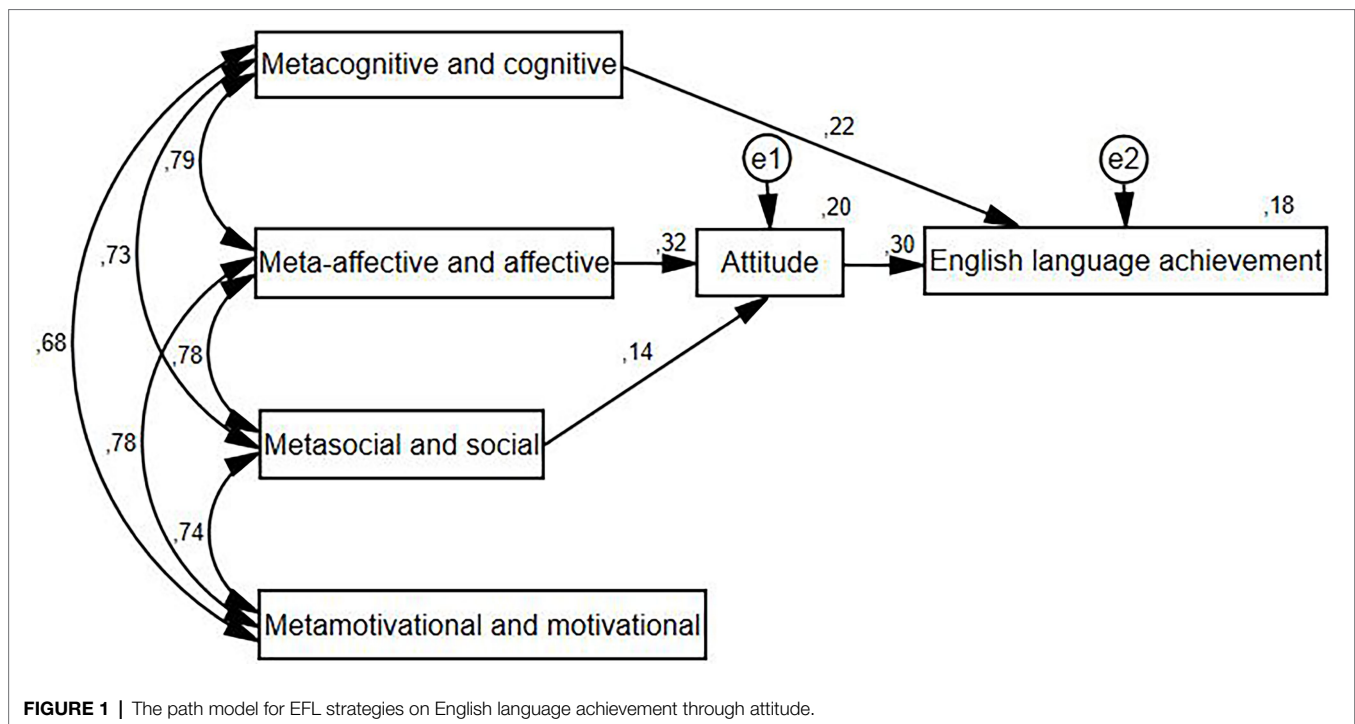
Finally, we investigated the effect of strategy use on English language achievement and attitudes. As Oxford's Strategic S²R Model considers strategies as being closely directed by their corresponding meta-strategies, we have regarded the strategies and their meta-strategy counterparts as common factors. The correlation coefficient was statistically significant between every strategy factor ($r=0.45-0.25$, $p < 0.001$). Our model showed acceptable fit indices (Chi-square = 35.574, $df=5$, $p=0.000$, CFI = 0.995, TLI = 0.977, NFI = 0.994, RMSEA = 0.061). We therefore concluded that English language achievement and attitude are statistically influenced by strategy use (**Figure 1**).

We found that the meta-affective and affective and metasocial and social categories directly influenced students' attitude toward English. A direct effect of attitude was observed on English language achievement. In addition, the metacognitive and cognitive categories had a direct effect on English language achievement, while English language achievement was indirectly affected by meta-affective and affective and metasocial and social factors. We could not confirm any significant effect of metamotivational and motivational factors on attitude or English language achievement.

DISCUSSION

We investigated the strategy use of 11-year-old lower secondary students in Grade 5 in connection with their language achievement and attitude toward the English language. Our first research question asked which LLS was the most frequently used by the sample. We found moderate strategy use with regard to a slightly modest application of the metasocial field, and the most frequent strategy use was observed in the affective field. These aspects of our findings partly correspond with most of the recent research with respect to moderate use of strategies; however, there are profound differences in the strategy preferences of the sample (Platsidou and Sipitanou, 2015; Alhaysony, 2017; Charoento, 2017; Dawadi, 2017; Habók and Magyar, 2018a,b, 2019, 2020; Habók et al., 2021). Raoofi et al. (2017) also pointed out the low level of social strategy use in their research. Another important statistically significant finding is that higher proficiency learners used learning strategies with greater frequency than their less proficient peers. This applies to every strategy field in agreement with Charoento's (2017) results.

Our second research question concerned differences in the use of LLS based on English language achievement. As concerns the sample, we regarded the EFL school mark as an indicator of English language achievement. The mean indicated that a considerable portion of the sample was grouped as more proficient. As a result, these students used LLS with greater frequency in all of the categories, which is a statistically significant finding. These results correspond with other research, which also reinforces this (Rao, 2016; Charoento, 2017; Raoofi et al., 2017; Sánchez, 2019). However, we also found that less proficient learners employed motivational strategies the most frequently, while their more proficient peers most often preferred the affective field, a result which is not reinforced by any previous findings. Apart from this, the strategy uses of both



subsamples followed the same order, with social and metasocial strategy use being the least preferred type for both. This may be due to the fact that our sample was mainly at the beginner/elementary level (A1–A2), so they cannot yet initiate conversations with others, even with native speakers. They also cannot understand many words and grammatical structures that are used by more proficient speakers, so social interaction is more difficult for them, even for the more advanced ones.

Our results on the role of LLS in English language achievement and attitude confirmed the statistically significant effect of LLS on background variables. English language achievement was directly influenced by the metacognitive and cognitive fields and attitudes and indirectly affected by the meta-affective and affective fields, as well as the metasocial and social fields. Our model could not confirm any direct or indirect effect of the metamotivational and motivational fields on attitude or English language achievement. This may be because motivational components form distinct factors and their role differs somewhat in predicting language achievement. These results are in line with previous findings (Platsidou and Kantaridou, 2014; Habók and Magyar, 2018a), which also concluded the outstanding role of attitudes, which is an important predictor of language achievement and reinforces the role of strategy use. In summary, strategy use influences English language achievement through attitude to language learning in a statistically significant way.

CONCLUSION

The main objective of the study was to find evidence for the role of strategy use in students' achievement at the

beginner/elementary level of English language learning. As a result, the strategy use preferences of the sample differed somewhat from the findings of previous research, as the affective and motivational fields were the ones the students preferred the most. This may be due to the fact that young children are more likely to use strategies that are rather emotional and related to their personality traits than strategies that require deeper understanding, specific learning techniques, and awareness, such as cognitive strategies. The use of social strategies was also quite low, probably owing to the low level of foreign language communication skills in the sample. As regards the different proficiency levels, more frequent strategy use was observed among the more proficient learners, a statistically significant finding. However, the patterns of strategy use were almost the same across the groups. The only difference was that the more proficient learners mostly preferred the affective field, while the less proficient ones mostly employed motivational strategies. This indicates that students at a higher level have more confidence to speak up and show how they feel about learning English. Learners with lower proficiency at this age often try to show that they are motivated, that is, that they are trying and want to achieve good results and present a good image of their own performance. The study also highlighted the importance of attitude; from the results, it can be concluded that, even at the beginner/elementary level, strategy use can affect language achievement and that a student's attitude is an important predictor and plays an important role as mediator between strategies and language achievement. This can have a positive impact on classroom performance and highlights the importance of teaching students about learning strategies.

LIMITATIONS

There are some limitations to consider in the study. First, the questionnaire was administered to fifth-grade students, who were at the beginner/elementary level of their English language learning. Thus, generalizability cannot be confirmed, and more research is needed across higher grades and higher proficiency learners. Second, we had difficulty identifying the affective domain in the first version of the questionnaire. For the fields in the present measurement tool, we have succeeded in identifying the affective and meta-affective domains of LLS. However, they still have to be optimized. Additional research is also called for with regard to the motivational components. Third, other underlying factors should be included in the investigation, such as self-efficacy, self-esteem, and self-concept.

PEDAGOGICAL IMPLICATIONS

The study points out that the role of learning strategies is substantial for the students in their language learning. Learning English is a complex process for Hungarian fifth graders. English pronunciation, vocabulary, and grammar are very different from those of Hungarian. For these learners, grammatical rules are often abstract phenomena, and it is difficult for them to associate meaning with the words they say and write. Furthermore, reading and listening comprehension are also influenced by many factors. The results draw attention to the paramount importance of teaching LLS, which can promote greater success among language learners. In addition, it is essential how consciously strategies are employed. Teachers are strongly urged to include strategy training in their courses. Strategy training can be conducted either in the form of an embedded sub-course in any of the subjects or in an independent form as an individual course. Strategy courses integrated into a school subject provide specific help for students learning that specific course material. For example, language learning strategies paid students in learning grammatical formulae or vocabulary in a foreign language, while general strategy courses help students to learn strategies that can be used in other school subjects, such as reading and writing strategies.

REFERENCES

- Alhaysony, M. (2017). Language learning strategies use by Saudi EFL students: the effect of duration of English language study and gender. *Theor. Prac. Lang. Stu.* 7, 18–28. doi: 10.17507/tpls.0701.03
- An, Z., Wang, C., Li, S., Gan, Z., and Li, H. (2021). Technology-assisted self-regulated English language learning: associations with English language self-efficacy, English enjoyment, and learning outcomes. *Front. Psychol.* 11:558466. doi: 10.3389/fpsyg.2020.558466
- Bećirović, S., Brdarević-Čeljo, A., and Polz, E. (2021). Exploring the relationship between language learning strategies, academic achievement, grade level, and gender. *J. Lang. Educ.* 7, 93–106. doi: 10.17323/jle.2021.10771
- Božinović, N., and Sindik, J. (2017). Construction of the questionnaire on foreign language learning strategies in specific Croatian context. *Coll. Antropol.* 41, 61–71.
- Charoento, M. (2017). Individual learner differences and language learning strategies. *Contem. Educ. Res. J.* 7, 57–72. doi: 10.18844/cej.v7i2.875
- Cohen, A. D. (1998). *Strategies in Learning and Using a Second Language*. Harlow: Longman.
- Csapó, B., and Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Front. Psychol.* 10:1522. doi: 10.3389/fpsyg.2019.01522
- Dawadi, S. (2017). Language learning strategies profiles of EFL learners in Nepal. *Eur. J. Educ. Soc. Sci.* 2, 42–55. <https://files.eric.ed.gov/fulltext/ED578188.pdf>
- Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z., and Skehan, P. (2003). "Individual differences in second language learning," in *The Handbook of Second Language Acquisition*. eds. C. J. Doughty and M. H. Long (Oxford: Blackwell), 589–630.
- Habók, A., Kong, Y., Ragchaa, J., and Magyar, A. (2021). Cross-cultural differences in foreign language learning strategy preferences among Hungarian. Chinese

Another implication of the study is that motivation and attitude also influence language achievement in a statistically significant way. Creating a learner-friendly and encouraging atmosphere is therefore essential. The findings from our research have provided important insights into these issues for classroom practice.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the datasets are confidential and cannot be shared with third parties. Requests to access the datasets should be directed to AH, habok@edpsy.u-szeged.hu.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRB at the Doctoral School of Education, University of Szeged. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AH and AM designed the study and implemented the data collection, as well as analyzing the data and participating in completing the manuscript. GM supervised the research and provided support. All the authors contributed to the editing and revision of the study and approved the final version of the manuscript.

FUNDING

This research was supported by the Research Programme for Public Education Development, Hungarian Academy of Sciences (grant KOZOKT2021-16) and a Hungarian National Research, Development and Innovation Fund grant (under the OTKA K135727 funding scheme).

- and Mongolian University Students. *Heliyon* 7, 1–7. doi: 10.1016/j.heliyon.2021.e06505
- Habók, A., and Magyar, A. (2018a). The effect of language learning strategies on proficiency, attitudes and school achievement. *Front. Psychol.* 8:2358. doi: 10.3389/fpsyg.2017.02358
- Habók, A., and Magyar, A. (2018b). Validation of a self-regulated foreign language learning strategy questionnaire through multidimensional modelling. *Front. Psychol.* 9:1388. doi: 10.3389/fpsyg.2018.01388
- Habók, A., and Magyar, A. (2019). The effects of EFL reading comprehension and certain learning-related factors on EFL learners' reading strategy use. *Cog. Educ.* 6, 1–19. doi: 10.1080/2331186X.2019.1616522
- Habók, A., and Magyar, A. (2020). The role of students' approaches in foreign language. *Cog. Educ.* 7, 1–21. doi: 10.1080/2331186X.2020.1770921
- Hajar, A. (2019). A critical review of research on language learning strategies used by Arab learners of English. *Stu. Self-Access Learn. J.* 10, 239–257. doi: 10.37237/100303
- Jabbari, M. J., and Golkar, N. (2014). The relationship between EFL learners' language learning attitudes and language learning strategies. *Int. J. Ling.* 6, 161–167. doi: 10.5296/ijl.v6i3.5837
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. 4th Edn. New York, NY: Guilford.
- Köksal, D., and Dünder, S. (2017). Factors affecting the use of self-regulated L2 learning strategies in Turkish FLE context. *J. Lang. Ling. Stu.* 13, 1–16. doi: 10.16986/HUJE.2017033805
- Lin, L., Lam, W.-I., and Tse, S. K. (2021). Motivational strategies, language learning strategies, and literal and inferential comprehension in second language Chinese reading: A structural equation modeling study. *Front. Psychol.* 12:707538. doi: 10.3389/fpsyg.2021.707538
- Malpartida, W. M. F. (2021). Language learning strategies, English proficiency and online English instruction perception during covid-19 in Peru. *Int. J. Instr.* 14, 155–172. doi: 10.29333/iji.2021.14410a
- Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. Boston: Heinle and Heinle.
- Oxford, R. L. (2011). *Teaching and Researching Language Learning Strategies*. Harlow: Longman/Pearson Education.
- Oxford, R. L. (2016). *Teaching and Researching Language Learning Strategies: Self-Regulation in Context*. 2nd Edn. New York: Routledge.
- Platsidou, M., and Kantaridou, Z. (2014). The role of attitudes and learning strategy use in predicting perceived competence in school-aged foreign language learners. *J. Lang. Lit.* 5, 253–260. doi: 10.7813/jll.2014/5-3/43
- Platsidou, M., and Sipitanou, A. (2015). Exploring relationships with grade level, gender and language proficiency in the foreign language learning strategy use of children and early adolescents. *Int. J. Res. Stu. Lang. Learn.* 4, 83–96. doi: 10.5861/ijrsl.2014.778
- Radwan, A. A. (2011). Effects of L2 proficiency and gender on choice of language learning strategies by university students majoring in English. *Asian EFL J.* 13, 114–162.
- Rao, Z. (2016). Language learning strategies and English proficiency: interpretations from information-processing theory. *Lang. Learn. J.* 44, 90–106. doi: 10.1080/09571736.2012.733886
- Raoofi, S., Binandeh, M., and Rahmani, S. (2017). An investigation into writing strategies and writing proficiency of university students. *J. Lang. Teach. Res.* 8, 191–198. doi: 10.17507/jltr.0801.24
- Rianto, A. (2020). A study of language learning strategy use among Indonesian EFL university students. *Reg. J.* 13, 231–256. doi: 10.18326/rgt.v13i2.231-256
- Rubin, J. (1975). What the 'good language learner' can teach us. *TESOL Q.* 9, 41–51. doi: 10.2307/3586011
- Salehi, M., and Jafari, H. (2015). Development and validation of an EFL self-regulated learning questionnaire. *Southern African Ling. App. Lang. Stu.* 33, 63–79. doi: 10.2989/16073614.2015.1023503
- Sánchez, N. S. (2019). Are low achievers low strategy users? *Lenguas Modernas* 53, 107–126.
- Taheri, H., Sadighi, F., Bagheri, M. S., and Bavali, M. (2019). EFL learners' L2 achievement and its relationship with cognitive intelligence, emotional intelligence, learning styles, and language learning strategies. *Cogent Educ.* 6:1655882. doi: 10.1080/2331186X.2019.1655882
- Thomas, N., and Rose, H. (2019). Do language learning strategies need to be self-directed? Disentangling strategies from self-regulated learning. *TESOL Q.* 53, 248–257. doi: 10.1002/tesq.473
- Thomas, N., Rose, H., and Pojanapunya, P. (2021). Conceptual issues in strategy research: examining the roles of teachers and students in formal education settings. *App. Ling. Rev.* 12, 353–369. doi: 10.1515/applirev-2019-0033
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Mental Processes*. Cambridge, MA: Harvard University Press.
- Wang, C., Schwab, G., Fenn, P., and Chang, M. (2013). Self-efficacy and self-regulated learning strategies for English language learners: comparison between Chinese and German college students. *J. Educ. Dev. Psychol.* 3, 173–191. doi: 10.5539/jedp.v3n1p173
- Wei, R., and Hu, Y. (2019). Exploring the relationship between multilingualism and tolerance of ambiguity: a survey study from an EFL context. *Bilingualism: Language and Cognition* 22, 1209–1219.
- Wei, R., Hu, Y., and Xiong, J. (2019). Effect size reporting practices in applied linguistics research: A study of one major journal. *SAGE Open* 9, 1–11. doi: 10.1177/2158244019850035
- Wong, L. L. C., and Nunan, D. (2011). The learning styles and strategies of effective language learners. *System* 39, 144–163. doi: 10.1016/j.system.2011.05.004
- Zimmerman, B. J. (2000). "Attaining self-regulation: A social cognitive perspective," in *Handbook of Self-Regulation*. eds. M. Boekaerts, P. R. Pintrich and M. Zeidner (San Diego, CA: Academic Press), 13–39.
- Zimmerman, B. J., and Schunk, D. H. (2011). "Self-regulated learning and performance: An introduction and an overview," in *Handbook of Self-Regulation of Learning and Performance*. eds. B. J. Zimmerman and D. H. Schunk (United States: Routledge), 1–12.

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 Habók, Magyar and Molnár. 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.

APPENDIX

Revised Version of the Self-Regulated Foreign Language Learning Strategy Questionnaire (SRFLLSQ).

When I learn English, ...

- | | |
|-----|---|
| | Metacognitive |
| 1. | I think of the relationships between what I already know and new things I learn in English. |
| 2. | I first skim an English passage, then go back and read carefully. |
| 3. | I look for opportunities to read as much as possible in English. |
| 4. | I write notes, messages, letters or reports in English. |
| 5. | I plan my schedule so I will have enough time to study English. |
| 6. | I pay attention when someone is speaking English. |
| 7. | I make summaries of information that I hear or read in English. |
| 8. | I try to find out how to be a better learner of English. |
| | Cognitive |
| 9. | I connect the sound of a new English word and an image or picture of the word to help me remember the word. |
| 10. | I use the English words I know in different ways. |
| 11. | I find the meaning of an English word by dividing it into parts that I understand. |
| 12. | I use new English words in a sentence so I can remember them. |
| 13. | I try to find patterns (grammar) in English. |
| 14. | I try not to translate word for word. |
| | Meta-affective |
| 15. | I notice if I am tense or nervous when I am studying or using English. |
| 16. | I encourage myself as I learn English so that I can learn what I would like. |
| 17. | I read in English as a leisure-time activity. |
| 18. | I organize my English language learning so that I always enjoy doing it. |
| 19. | I plan my English language learning so that I can perform better. |
| 20. | I have more success learning English when I feel like doing it. |
| 21. | I talk to others about how I feel when learning English. |
| 22. | I give myself a reward or treat when I do well in English. |
| | Affective |
| 23. | I try to relax whenever I feel afraid of using English. |
| 24. | I encourage myself to speak English even when I feel afraid of making a mistake. |
| 25. | I try to overcome my anxiety if learning English is difficult. |
| 26. | Making mistakes does not take away my desire to use English language. |
| 27. | It gives me a good feeling when I do well in English. |
| 28. | I am good at learning English. |
| 29. | I like learning English. |
| 30. | I am happy when I can use my knowledge of English in other school subjects. |
| | Metasocial |
| 31. | I try to learn about English-language cultures and/or other cultures through English. |
| 32. | I look for people I can talk to in English. |
| 33. | I look at English-language TV shows, movies or websites to get to know the cultures of English native speakers and/or other cultures through English. |
| 34. | I choose leisure activities where I encounter English-language cultures and/or other cultures through English as well. |
| 35. | I plan what I want to find out about the cultures of English speakers and/or other cultures through English. |
| 36. | I practise English with my peers. |
| 37. | I look for similarities and differences between my own culture and the cultures of English native speakers and/or other cultures through English. |
| 38. | Getting to know English-language cultures helps me to learn the language. |
| | Social |
| 39. | I start conversations in English. |
| 40. | I make up new words in English if I do not know the right ones. |
| 41. | When I speak with highly proficient speakers of English, I think it is important to get acquainted with their culture. |
| 42. | The more I learn about English-language culture(s) and/or other cultures through English, the more I love English. |
| 43. | I try to discover similarities and differences between my own culture and English-language cultures and/or other cultures I'm learning about through English. |
| 44. | I like talking to other students about English-language cultures and/or other cultures I'm learning about through English. |
| | Metamotivational |
| 45. | I pay attention to the types of English learning tasks that make me excited. |
| 46. | I plan ahead for what I'm going to learn in English in a week or two. |
| 47. | I add something motivating to my English learning environment, such as pleasant music. |
| 48. | I pay attention to making it interesting to learn English. |
| | Motivational |
| 49. | I give myself a reward for good progress or achievement. |
| 50. | I use positive self-talk about my reasons for achieving my aims. |
| 51. | I regard learning English as a game. |
| 52. | I believe that I'm fully capable of doing English learning tasks. |



Recent Development in University Student Learning Research in Blended Course Designs: Combining Theory-Driven and Data-Driven Approaches

Feifei Han^{1,2*}

¹ Department of Pedagogy and Psychology, Faculty of Education, The University of Hradec Králové, Hradec Králové, Czechia, ² Griffith Institute for Educational Research, Griffith University, Brisbane, QLD, Australia

Keywords: university student learning research, blended course designs, theory-driven approaches, data-driven approaches, a combined approach

UNIVERSITY STUDENT LEARNING IN BLENDED COURSE DESIGNS

OPEN ACCESS

Edited by:

Jesus de la Fuente,
University of Navarra, Spain

Reviewed by:

Chia-Lin Tsai,
University of Northern Colorado,
United States

*Correspondence:

Feifei Han
feifei.han@griffith.edu.au

Specialty section:

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

Received: 27 March 2022

Accepted: 28 April 2022

Published: 17 May 2022

Citation:

Han F (2022) Recent Development in
University Student Learning Research
in Blended Course Designs:
Combining Theory-Driven and
Data-Driven Approaches.
Front. Psychol. 13:905592.
doi: 10.3389/fpsyg.2022.905592

Advances in Internet and computer-based technologies have increased the growth of alternative learning spaces, creating entirely new ways of conceptualizing and assessing the learning experience of students (Wong, 2019). Research has repeatedly found positive impacts of various forms of technology-enhanced learning, such as learning embedded in social networking sites, web-conferencing, webinars, e-portfolio, digital games, mobile apps, and virtual and second worlds (Jarvoll, 2018; Winkelmann et al., 2020). The higher education sector has also shown a rapid development in the use of computer and web-based technologies, which has affected experiences of learning of a significant proportion of university students worldwide. This change has resulted in the widespread use of diverse online spaces developed for learning in the form of e-learning courses and/or a combination of face-to-face and online delivery systems, known as blended course designs (Shin et al., 2018). In a synergy of research foci and research methods of university student learning in blended contexts, Bliuc et al. (2007) defines blended course designs as “systematic combination of co-present (face-to-face) interactions and technologically-mediated interactions between students, teachers and learning resources” (p. 234).

More recently, the coronavirus pandemic (COVID-19) emergency has required higher education learning and teaching around the world to rapidly respond, in particular, redeploying even more learning and teaching activities to virtual learning spaces to promote physical distancing. As a result, more and more face-to-face courses have been delivered as blended courses (Mali and Lim, 2021). Under such circumstances, it becomes vital importance to understand student learning experience in blended course designs, its relation to various forms of learning outcomes, and the key factors which may impact such experience. However, there are many challenges to study the student experience in blended course designs as it is a complex phenomenon (Ma and Lee, 2021). Compared with learning in a single mode (i.e., fully face-to-face courses or fully online courses), in blended course designs, students are increasingly involved in decision-making in the learning process, such as with whom to work in a classroom tutorial and with whom to discuss in online forum, how many hours they learn online, whether to study in a physical library or log onto an online database, whether to collaborate in a face-to-face laboratory or whether to collaborate in virtual reality. These decisions require students to move back and forth between face-to-face and

online contexts, and across physical and virtual learning environments (Han and Ellis, 2020a). As a result, the complexity of student experience in blended learning involves an interplay of a wide range of factors, which include students' cognition (e.g., conceptions, approaches, and perceptions in learning) (Trigwell and Prosser, 2020); their choices of social interactions in learning (e.g., with whom to collaborate and the mode of the collaborations) (Hadwin et al., 2018); and the material elements across both physical and virtual learning spaces in face-to-face and online learning components (e.g., students' choices of learning spaces and their interactions with online learning activities) (Laurillard, 2013). As such, evaluation of university student learning experience in blended course designs requires research methods that move beyond approaches that do not routinely investigate the combined contributions of learners and the material things involved in learning to achieve academic performance, (Wu et al., 2010; López-Pérez et al., 2011).

THEORY-DRIVEN APPROACHES

Traditionally, research into student learning experience and academic performance in higher education has largely adopted theory-driven approaches, which test hypotheses derived from theories in educational psychology, learning sciences, and pedagogy and curriculum research (Trigwell and Prosser, 2020). Based on the accumulated empirical evidence which may support or refute the hypotheses, theories and models of learning have been constantly refined, modified, and updated. The theoretical frameworks departing from theory-driven perspectives predominantly assess various aspects in student learning experience using self-reported instruments and measurements, such as focus group, semi-structured interviews, and Likert-scale questionnaires (Han et al., 2022).

In a number of state-of-the-art articles, for instance, the instruments in the major frameworks on the student learning research in higher education have been reviewed and summarized, including Student Approaches to Learning Research, Self-regulated Learning Research, Information Processing Research, and Student Engagement Research (Lonka et al., 2004; Vermunt and Donche, 2017; Zusho, 2017). Uniformly, the primary means to collect data in these frameworks are self-reported questionnaires, such as Study Process Questionnaire (SPQ; Biggs et al., 2001), the Revised Approaches to Studying Inventory (RASI; Entwistle and McCune, 2004), and the Inventory Learning to Teach Process (ILTP; Endedijk et al., 2016) and Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, 2004); Inventory of Learning Patterns (ILP; Donche and Van Petegem, 2008), Course Experience Questionnaire (CEQ; Ramsden, 1991), and the Inventory of Perceived Learning Environments Extended (IPSEE; Könings et al., 2012).

However, the self-reported measures have been criticized for being subjective and have been questioned about their accuracy in describing students' use of learning approaches and strategies in real learning contexts (Zhou and Winne, 2012).

In addition, the self-reported measures and data also suffers from their limited capacities to represent the complex (e.g., using multiple indicators) and dynamic (e.g., changes over time) nature of student learning behaviors. To improve the insights of contemporary university student experiences of learning, suggestions have been put forward to expand the current self-reporting methods by including other types of measurements to study student learning (Vermunt and Donche, 2017). In this regard, learning analytics research is a promising avenue. For instance, Richardson (2017) suggested "The rapidly expanding field of learning analytics provides both researchers and practitioners with the opportunity to monitor students' strategic decisions in online environments in minute detail and in real time (p. 359)."

DATA-DRIVEN APPROACHES

The recent development of educational technology has produced prolific studies using learning analytics, which enables a capacity to collect rich and detailed digital traces of students' interactions with a variety of online learning resources and activities. The type of digital trace/log data, also known as the observational data, have the advantage of offering descriptions of student learning behaviors and strategies relatively more objectively and in a more granular details than using self-reported methods (Siemens, 2013; Baker and Siemens, 2014). Departing from data-driven approaches, learning analytics research has emerged as a growing area and has gradually gained popularity in student learning in higher education (Sclater et al., 2016). It employs advanced data mining techniques and algorithms to process the observational analytic data in relation to students' demographic information, which has been increasingly applied in various domains in higher education sector, such as advising students' career choice (Bettinger and Baker, 2013); detecting at risk students to improve retention (Krumm et al., 2014); providing personalized feedback (Gibson et al., 2017); identifying patterns of learning tactics and strategies (Chen et al., 2017); facilitating collaborative learning (Kaendler et al., 2015); monitoring students' affect in learning (Ocumpaugh et al., 2014); and predicting their academic learning outcomes (Romero et al., 2013). However, the data-driven approaches are often fragmented from educational theories and rely purely on empiricism, which limit the insights they can offer for directing pedagogical innovations and reforms, supporting learning design, fostering quality learning experience, and improving academic performance (Buckingham Shum and Crick, 2012).

Despite its wide applications, learning analytics research have received criticism that the data-driven approaches it relies on is fragmented from educational theory but overly focuses on quantitative number (Rodríguez-Triana et al., 2015). The patterns and models of student learning derived from such data-centric perspectives without proper guidance from educational theories are often result in erroneous interpretation, which have limited insights for generating actionable knowledge in order to locate learning barriers and to offer ideas for teaching practice and curriculum design (Wong and Li, 2020; Han and Ellis, 2021).

A COMBINED APPROACH

Recognizing the limitations of an overly data-centric approach used in learning analytics research, researchers have proposed to adopt a more holistic approach when designing research so that student learning behaviors and patterns can be captured in a comprehensive manner, and big data modeling and interpretation can be guided via sound theories (Lockyer et al., 2013; Rienties and Toetenel, 2016). This has resulted in an increasing amount of research using a combined approach, which employs both self-reported and observational instruments to measure student learning experience in a complementary manner (Gašević et al., 2015).

Reviewing the existing research using a combined approach suggests that these studies have two different aims. One aim is to examine how a combined approach may improve the explanatory power of predicting student learning outcomes by including both self-reported and observational measures of aspects in the processes of student learning (Reimann et al., 2014). Despite basing on different learning theories, the majority of studies in this line of inquiry have reported that combining the self-reported and observed measures of student learning have significantly improved the variance explained in the prediction of student academic achievement than using either self-reporting or observational method alone (Tempelaar et al., 2015; Han and Ellis, 2020b). For example, using multiple regression analyses, Pardo et al. (2017) found that students' reported anxiety in learning and use of self-regulated learning strategies could only explain 7% of variance in students' academic performance; whereas adding the observed frequency of students' online interactions into the regression model could explained 32% of variance in students' academic performance, significantly increasing the total variance explained by 25%. In another study which adopted Student Approaches to Learning research framework, Ellis et al. (2017) reported similar findings: while students' self-reported use of approaches to learning only predicted 9% of the variance in their academic learning outcomes, including students' observed online learning events in the multiple regression equation explained an extra of 25% in the students' learning outcomes.

Another aim of the studies which combine the self-reported and observational measures is to investigate the consistency between the two methods in terms of describing student experiences of learning (Reimann et al., 2014). Similarly, research with this aim has also departed from different learning perspectives and has diverse foci, such as using self-reported questionnaires to assess students' intrinsic motivation, test anxiety, self-efficacy, engagement, effort expenditure, achievement goal, learning orientations and motives on the one hand. On the other hand, a diversity of observed indicators of students' online learning, including frequency of clicks, completion of online learning tasks, duration of online learning events, as well as time-stamped sequences of online learning behaviors have been collected through digital traces to derive students' online learning tactics, strategies, and approaches

(Gašević et al., 2017; Pardo et al., 2017; Han et al., 2020; Sun and Xie, 2020; Ober et al., 2021). However, inconclusive results have been reported among studies with this aim. The non-alignment between the descriptions and categorisations of student learning experience by self-reports and observations have been found in a number of studies.

For instance, Gašević et al. (2017) found that there was no significant difference of self-reported using surface learning strategies between students who were categorized as deep and surface learners according to their observed online learning strategies. In contrast, consistency between students' self-reported learning orientations (as measured by their approaches to learning and perceptions of learning) and their online participation was observed in Han et al. (2020). Students who were categorized as having an "understanding" learning orientation (i.e., the learning was oriented toward gaining an in-depth understanding of the subject matter) were also observed to participate more online learning than those who were categorized as having a "reproducing" learning orientation (i.e., the learning was oriented toward reproducing facts and satisfying course requirements).

These inconsistent results are possibly caused by (1) different learning theories have been adopted by these studies (e.g., self-regulated learning, Student Approaches to Learning, student engagement); and (2) the diverse means of how students' online learning is measured (e.g., frequency, duration, temporal sequence, or a combination of multiple indicators). Until a firmer conclusion can be made, many more studies are required to examine the extent to which the self-reported and observational measures of student learning experience align with each other.

The strengths of a combined approach lie in multiple ways. First, it has an advantage of offering richer information in the way of predicting student learning outcomes over using a single approach, with each approach supplementing the other. While the observational measures are able to provide objective evidence as to what students actually do in their learning (Fincham et al., 2018), they do not, however, have capacity to reflect students' intents and motives behind the ways they learn as in the self-reported studies (Asikainen and Gijbels, 2017; Gerritsen-van Leeuwenkamp et al., 2019). Second, a combined approach can serve as a triangulation to check the validity of the results derived from either theory-driven or data-driven approaches. Third, the multiple data collection and analysis methods used in the combined approach also strengthen the analytical power of the analyses. All the merits of combining theory-driven and data-driven approaches point out its future applications to advance university student learning research and its potential to tackle the other complex issues of contemporary student experiences of learning.

AUTHOR CONTRIBUTIONS

FH contributed to the conception of the work, drafted the work and revised it critically for important intellectual content,

approved the final version of the paper to be published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

REFERENCES

- Asikainen, H., and Gijbels, D. (2017). Do students develop towards more deep approaches to learning during studies? A Systematic review on the development of students' deep and surface approaches to learning in higher education. *Educ. Psychol. Rev.* 29, 205–234. doi: 10.1007/s10648-017-9406-6
- Baker, R., and Siemens, G. (2014). "Educational data mining and learning analytics," in *The Cambridge Handbook of the Learning Sciences, 2nd ed*, eds Sawyer R (New York, NY: Cambridge University Press), 253–272.
- Bettinger, E., and Baker, R. (2013). The Effects of student coaching: an Evaluation of a randomized experiment in student advising. *Educ. Eval. Policy Anal.* 36, 3–19. doi: 10.3102/0162373713500523
- Biggs, J., Kember, D., and Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *Br. J. Educ. Psychol.* 71, 133–149. doi: 10.1348/000709901158433
- Blüch, A. M., Goodyear, P., and Ellis, R. A. (2007). Research focus and methodological choices in studies into students' experiences of blended learning in higher education. *Intern. High. Educ.* 10, 231–244. doi: 10.1016/j.iheduc.2007.08.001
- Buckingham Shum, S., and Crick, R. (2012). "Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics," In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (New York: ACM Press), 92–101.
- Chen, B., Resendes, M., Chai, C., and Hong, H. (2017). Two tales of time: Uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interact. Learn. Environ.* 25, 162–75. doi: 10.1080/10494820.2016.1276081
- Donche, V., and Van Petegem, P. (2008). "The validity and reliability of the short inventory of learning patterns," in *Style and cultural differences: How can organisations, regions and countries take advantage of style differences*, eds Broeck, Cools, Redmond, and Evans, (Gent: Vlerick Leuven Gent Management School), 49–59.
- Ellis, R. A., Han, F., and Pardo, A. (2017). Improving insights from learning analytics – The Value of combining observed and self-report data on university student learning. *Educ. Technol. Soc.* 20, 158–169.
- Endeldijk, M. D., Brekelmans, M., Sleegers, P., and Vermunt, J. D. (2016). Measuring students' self-regulated learning in professional education: bridging the gap between event and aptitude measurements. *Qual. Qty.* 50, 2141–2164. doi: 10.1007/s11135-015-0255-4
- Entwistle, N., and McCune, V. (2004). The conceptual bases of study strategy inventories. *Educ. Psychol. Rev.* 16, 325–345. doi: 10.1007/s10648-004-0003-0
- Fincham, E., Gašević, D., Jovanović, J., and Pardo, A. (2018). From study tactics to learning strategies: an analytical method for extracting interpretable representations. *IEEE Trans. Learn. Technol.* 12, 59–72. doi: 10.1109/TLT.2018.2823317
- Gašević, D., Dawson, S., and Siemens, G. (2015). Let's not forget: learning analytics are about learning. *TechTrends* 59, 64–75. doi: 10.1007/s11528-014-0822-x
- Gašević, D., Jovanovic, J., Pardo, A., and Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *J. Learn. Anal.* 4, 113–128. doi: 10.18608/jla.2017.42.10
- Gerritsen-van Leeuwenkamp, K., Joosten-ten Brinke, D., and Kester, L. (2019). Students' perceptions of assessment quality related to their learning approaches and learning outcomes. *Stud. Educ. Eval.* 63, 72–82. doi: 10.1016/j.stueduc.2019.07.005
- Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., and Knight, S. (2017). "Reflective writing analytics for actionable feedback," in *Proceedings of the Seventh International Learning Analytics and Knowledge Conference* (New York: ACM Press), 153–162. doi: 10.1145/3027385.3027436
- Hadwin, A., Järvelä, S., and Miller, M. (2018). "Self-regulation, co-regulation, and shared regulation in collaborative learning environments," in *Handbook of Self-Regulation of Learning and Performance*, eds Schunk, D. H., and Greene J. A., (London: Taylor and Francis Group), 83–106.
- Han, F., and Ellis, R. A. (2020a). Personalised learning networks in the university blended learning context. *Comunicar* 62, 19–30. doi: 10.3916/C62-2020-02
- Han, F., and Ellis, R. A. (2020b). Combining self-reported and observational measures to assess university student academic performance in blended course designs. *Australas. J. Educ. Technol.* 36, 1–14. doi: 10.14742/ajet.6369
- Han, F., and Ellis, R. A. (2021). Predicting students' academic performance by their online learning patterns in a blended course. *Educ. Technol. Soc.* 24, 191–204.
- Han, F., Ellis, R. A., and Pardo, A. (2022). The descriptive features and quantitative aspects of students' observed online learning: How are they related to self-reported perceptions and learning outcomes? *IEEE Trans. Learn. Technol.* doi: 10.1109/TLT.2022.3153001
- Han, F., Pardo, A., and Ellis, R. A. (2020). Students' self-report and observed learning orientations in blended university course design: How are they related to each other and to academic performance? *J. Comput. Assist. Learn.* 36, 969–980. doi: 10.1111/jcal.12453
- Jarvoll, A. B. (2018). "I'll have everything in diamonds!": Students' experiences with Minecraft at school. *Stud. Paedagog.* 23, 67–90. doi: 10.5817/SP2018-4-4
- Kaendler, C., Wiedmann, M., Rummel, N., and Spada, H. (2015). Teacher competencies for the implementation of collaborative learning in the classroom: a framework and research review. *Educ. Psychol. Rev.* 27, 505–536. doi: 10.1007/s10648-014-9288-9
- Könings, K. D., Brand-Gruwel, S., and Elen, J. (2012). Effects of a school reform on longitudinal stability of students' preferences with regard to education. *Br. J. Educ. Psychol.* 82, 512–532. doi: 10.1111/j.2044-8279.2011.02044.x
- Krumm, A., Waddington, R., Teasley, S., and Lonn, S. (2014). "A learning management system-based early warning system for academic advising in undergraduate engineering," in *Learning Analytics: From Research to Practice*, eds Larusson, J., and White, B., (New York, NY: Springer), 103–119.
- Laurillard, D. (2013). *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies*, London: Routledge.
- Lockyer, L., Heathcote, E., and Dawson, S. (2013). Informing pedagogical action: aligning learning analytics with learning design. *Am. Behav. Sci.* 57, 1439–1459. doi: 10.1177/0002764213479367
- Lonka, K., Olkinuora, E., and Mäkinen, J. (2004). Aspects and prospects of measuring studying and learning in higher education. *Educ. Psychol. Rev.* 16, 301–323. doi: 10.1007/s10648-004-0002-1
- López-Pérez, M., López-Pérez, M., and Rodríguez-Ariza, L. (2011). Blended learning in higher education: students' perceptions and their relation to outcomes. *Comput. Educ.* 56, 818–826. doi: 10.1016/j.compedu.2010.10.023
- Ma, L., and Lee, C. S. (2021). Evaluating the effectiveness of blended learning using the ARCS model. *J. Comput. Assist. Learn.* 37, 1397–1408. doi: 10.1111/jcal.12579
- Mali, D., and Lim, H. (2021). How do students perceive face-to-face/blended learning as a result of the Covid-19 pandemic? *Int. J. Manag. Educ.* 19, 100552. doi: 10.1016/j.ijme.2021.100552
- Ober, T. M., Hong, M. R., Rebouças-Ju, D. A., Carter, M. F., Liu, C., and Cheng, Y. (2021). Linking self-report and process data to performance as measured by different assessment types. *Comput. Educ.* 167, 104188. doi: 10.1016/j.compedu.2021.104188
- Ocuppaugh, J., Baker, R., Gowda, S., Heffernan, N., and Heffernan, C. (2014). Population validity for educational data mining models: a case study in affect detection. *Br. J. Educ. Technol.* 45, 487–501. doi: 10.1111/bjet.12156
- Pardo, A., Han, F., and Ellis, R. A. (2017). Combining university student self-regulated learning indicators and engagement with online learning events

FUNDING

This study was supported by the International mobilities for research activities of the University of Hradec Králové II, CZ.02.2.69/0.0/0.0/18_053/0017841.

- to predict academic performance. *IEEE Trans. Learn. Technol.* 10, 82–92. doi: 10.1109/TLT.2016.2639508
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ. Psychol. Rev.* 16, 385–407. doi: 10.1007/s10648-004-0006-x
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: the course experience questionnaire. *Stud. High. Educ.* 16, 129–150. doi: 10.1080/03075079112331382944
- Reimann, P., Markauskaite, L., and Bannert, M. (2014). e-R esearch and learning theory: what do sequence and process mining methods contribute? *Br. J. Educ. Technol.* 45, 528–540. doi: 10.1111/bjet.12146
- Richardson, J. T. (2017). Student learning in higher education: a commentary. *Educ. Psychol. Rev.* 29, 353–362. doi: 10.1007/s10648-017-9410-x
- Rienties, B., and Toetnel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: a cross-institutional comparison across 151 modules. *Comput. Hum. Behav.* 60, 333–341. doi: 10.1016/j.chb.2016.02.074
- Rodríguez-Triana, M., Martínez-Monés, A., Asensio-Pérez, J., and Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *Br. J. Educ. Technol.* 46, 330–343. doi: 10.1111/bjet.12198
- Romero, C., López, M. I., Luna, J. M., and Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Comput. Educ.* 68, 458–472. doi: 10.1016/j.compedu.2013.06.009
- Scater, N., Peasgood, A., and Mullan, J. (2016). *Learning Analytics in Higher Education: A Review of UK and International Practice*. London: Jisc.
- Shin, Y., Park, J., and Lee, S. G. (2018). Improving the integrated experience of in-class activities and fine-grained data collection for analysis in a blended learning class. *Interact. Learn. Environ.* 26, 597–612. doi: 10.1080/10494820.2017.1374980
- Siemens, G. (2013). Learning analytics: the emergence of a discipline. *Am. Behav. Sci.* 57, 1380–1400. doi: 10.1177/0002764213498851
- Sun, Z., and Xie, K. (2020). How do students prepare in the pre-class setting of a flipped undergraduate math course? A latent profile analysis of learning behavior and the impact of achievement goals. *Intern. High. Educ.* 46, 100731. doi: 10.1016/j.iheduc.2020.100731
- Tempelaar, D. T., Rienties, B., and Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Comput. Hum. Behav.* 47, 157–167. doi: 10.1016/j.chb.2014.05.038
- Trigwell, K., and Prosser, M. (2020). *Exploring University Teaching and Learning: Experience and Context*, London: Palgrave Macmillan.
- Vermunt, J. D., and Donche, V. (2017). A learning patterns perspective on student learning in higher education: state of the art and moving forward. *Educ. Psychol. Rev.* 29, 269–299. doi: 10.1007/s10648-017-9414-6
- Winkelmann, K., Keeney-Kennicutt, W., Fowler, D., Lazo Macik, M., Perez Guarda, P., and Joan Ahlborn, C. (2020). Learning gains and attitudes of students performing chemistry experiments in an immersive virtual world. *Interact. Learn. Environ.* 28, 620–634. doi: 10.1080/10494820.2019.1696844
- Wong, B. T. M., and Li, K. C. (2020). A review of learning analytics intervention in higher education (2011–2018). *J. Comput. Educ.* 7, 7–28. doi: 10.1007/s40692-019-00143-7
- Wong, R. (2019). Basis psychological needs of students in blended learning. *Interact. Learn. Environ.* 1[†]15. doi: 10.1080/10494820.2019.1703010
- Wu, J. H., Tennyson, R., and Hsia, T. L. (2010). A Study of student satisfaction in a blended e-learning system environment. *Computers and Education* 55, 155–164. doi: 10.1016/j.compedu.2009.12.012
- Zhou, M., and Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, 22, 413–419. doi: 10.1016/j.learninstruc.2012.03.004
- Zusho, A. (2017). Toward an integrated model of student learning in the college classroom. *Educ. Psychol. Rev.* 29, 301–324. doi: 10.1007/s10648-017-9408-4

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.

Copyright © 2022 Han. 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 Proactive-Reactive Resilience as a Mediational Variable Between the Character Strength and the Flourishing in Undergraduate Students

Jesús de la Fuente^{1,2*}, Begoña Urien¹, Elkin O. Luis¹, María Carmen González-Torres¹, Raquel Artuch-Garde³ and Alvaro Balaguer¹

¹ School of Education and Psychology, University of Navarra, Pamplona, Spain, ² School of Psychology, University of Almería, Almería, Spain, ³ School of Health and Psychology, Public University of Navarra, Pamplona, Spain

OPEN ACCESS

Edited by:

Jin Su Jeong,
University of Extremadura, Spain

Reviewed by:

Xiaoqin Zhu,
The Hong Kong Polytechnic
University, Hong Kong SAR, China
Ansar Abbas,
Airlangga University, Indonesia

*Correspondence:

Jesús de la Fuente
jdlfuente@unav.es

Specialty section:

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

Received: 17 January 2022

Accepted: 25 March 2022

Published: 18 May 2022

Citation:

de la Fuente J, Urien B, Luis EO, González-Torres MC, Artuch-Garde R and Balaguer A (2022) The Proactive-Reactive Resilience as a Mediational Variable Between the Character Strength and the Flourishing in Undergraduate Students. *Front. Psychol.* 13:856558. doi: 10.3389/fpsyg.2022.856558

The aim of this research was to delimit the predictive and mediational model of resilience between character strengths to predict flourishing, in a sample of undergraduate students. After signing their informed consent, 642 university students completed three validated scales (i.e., character strengths, resilience, and flourishing). Using an ex post facto design, regression, structural modeling, and mediation analyses were carried out, in order to construct a multi-causal predictive model. Results indicated a consistent predictive direct effect of character strengths on resilience and flourishing and of resilience on flourishing. As hypothesized, resilience also showed a mediating effect on the relationship between character strengths and flourishing. Additionally, results also revealed that the reactive and proactive factors of resilience were explained by different character strengths (e.g., emotional strength/cognitive, interpersonal strengths), reinforcing the idea that the two directions are complementary and necessary. Finally, several implications were established for the practice of positive psychology.

Keywords: strength character, proactive and reactive resilience, flourishing, university, mediational model

INTRODUCTION

Positive Psychology (PP) focuses on factors that promote human wellbeing, in contrast to the classic study of factors involved in disorders (Joseph and Linley, 2005), and it is defined as the scientific study of what makes life most worth living (Seligman and Csikszentmihalyi, 2000; Ghielen et al., 2018). Thus, positive psychology ultimately seeks to explain and predict psychological wellbeing (Dodge et al., 2012). In this regard, a wellbeing model based on actual positive psychology (PP 2.0) needs to enhance individuals' positive traits, attitudes, and behaviors and/or manage negative ones in order to promote wellbeing and decrease mental illness at the same time (Ivtzan et al., 2016). PP 2.0 recognizes that wellbeing involves a dialectical balance between the "light" and the "dark" aspects of life (Lomas, 2016).

The PP, has also influenced the emergence and interest by of new psychological constructs (Duckworth et al., 2005; Proyer et al., 2013; Ciarrochi et al., 2016) such as character strengths,

subjective wellbeing, psychological wellbeing, hedonic wellbeing, eudaimonic, flourishing, happiness, thriving or resilience which have become part of this new psychological dictionary (Dodge et al., 2012; Ruch and Hofmann, 2017; Schreiner, 2018).

The present paper adopts flourishing as an integrated approach to study wellbeing, which includes both hedonic and eudaimonic wellbeing traditions (Patnaik, 2021). Flourishing/Flowering defined as the favorable relationship between individuals, their environment, and culture, which in turn influences the experimentation of high levels of wellbeing and consequently of mental health have developed an important body of research (Keyes, 2002).

Although character strengths (i.e., positive stable personality traits; Wagner et al., 2020), resilience (i.e., the ability to resist negative events and/or to recover from them; Harms et al., 2018), and flourishing have emerged from different research domains, they share the common denominator of promoting psychological wellbeing, while they are related to each other in ways yet to be defined (Donaldson et al., 2015; Ciocanel et al., 2017). In addition to flourishing, character strengths and resilience are also complex constructs. Character strengths are organized under six broad virtues: (1) Wisdom and knowledge; (2) Courage; (3) Humanity; (4) Justice; (5) Temperance; and (6) Transcendence (Peterson and Seligman, 2004). On the other hand, resilience in recent studies suggests that it consists of two dimensions: reactive and proactive factors (de la Fuente et al., 2021a).

Thus, the present paper seeks to clarify the relationships between the mentioned variables by answering the following questions: Do character strengths predict resilience? To what degree do these two predict flourishing? To what extent do character strengths and resilience contribute to the flourishing of under-graduate students? From a research perspective, our work will contribute to establishing which character strengths relate to resilience, to confirming whether resilience comprises of two sub-dimensions, and to unfolding direct and indirect relationships among the variables under study. From an applied perspective, this paper could direct psychological interventions in order to improve flourishing in under-graduate students.

Wellbeing and Flourishing

Two different conceptual frameworks usually explain wellbeing (Chaves, 2021). The *hedonic paradigm* (subjective wellbeing as defined as by Diener's model) considers wellbeing as the result of the cognitive and affective evaluation that a person makes of his life, which leads him to experience high levels of positive affect, low levels of negative affect, high life satisfaction, and happiness (Diener, 2009). For its part, the *eudaimonic paradigm* (psychological wellbeing as defined by Ryff's model) understands wellbeing as human potential and highlight that wellbeing is facilitated by overcoming life challenges through applying human potential traits. He proposes six key components of wellbeing: purpose in life, mastery of the environment, positive relationships, self-acceptance, personal growth, and autonomy) (Ryff and Singer, 2006; Ryff, 2014, 2016). Since both perspectives can complement each other, tentative attempts have been proposed in order to unify these two traditions (Hedonistic and Eudaimonic), with the aim of comprehensively explaining

wellbeing. In fact, mental wellbeing is currently understood as the combination between the hedonic and eudaimonic perspectives (Ryan and Deci, 2001; Schotanus-Dijkstra et al., 2016; Chaves, 2021). Recent definitions of flourishing combine hedonic and eudaimonic elements to create a more comprehensive and holistic approach (Chaves, 2021).

From a temporal perspective, initially, the most influential perspective on human flourishing was the eudaimonic (VanderWeele, 2020). However, from a psychological perspective, flourishing also incorporates the hedonic experience (Schotanus-Dijkstra et al., 2016). In this respect, extant research (Keyes, 2007), as well as recent studies, also show the importance of improving flourishing so as to increase hedonic wellbeing (Allison et al., 2020; Lee et al., 2021).

Thus, these efforts lead to jointly considering both perspectives (hedonic and eudaimonic) of wellbeing. That is, to identify positive psychological states (subjective wellbeing) together with the value and impact of hedonic experiences that motivate and reinforce the search for basic needs satisfaction, and the update of human potential (significant vocation and a sense of purpose). Based on the above-mentioned, the present paper's approach to wellbeing follows this integrative perspective and accepts that flourishing comprises both wellbeing traditions.

These integrated eudaimonic and hedonic perspectives are based on influential theoretical approaches. Firstly, the model of human potential (Ryff and Singer, 2006), which encompasses emotional and physical health, evaluates autonomy, personal growth, self-acceptance, life purpose, mastery, and positive relationships. Secondly, self-determination theory (Ryan and Deci, 2000), considers that "self-realization" is a facilitator of wellbeing experience, and emphasizes the fundamental role that autonomy, competence, and relationships play in flourishing. Finally, Seligman's (2018) PERMA model, includes five fundamental aspects for human flourishing. These aspects are the experience of positive emotions, personal commitment to what is done, the establishment of positive interpersonal relationships, search for meaning, meaning, and purpose in life, and agency capacity. More recently, from a more interactive and contextual approach, the Self-vs. External-Regulation Theory, SR-ER model (de la Fuente et al., 2017; de la Fuente-Arias, 2017) has proposed that the achievement of psychological wellbeing is a consequence of the combination of the optimal levels of personal and contextual regulation.

Thus, Flourishing is a psychological construct referring to the experience that life is going well which combines a sense of feeling good and effective functioning. Flourishing is considered a personal indicator that corresponds to a high level of mental wellbeing (Huppert, 2009a,b). It is a global construct that integrates psychological variables from the hedonic and eudaimonic perspectives (e.g., perceived competence, emotional stability, engagement, meaning, optimism, positive emotions, positive relations, self-esteem, and vitality (Huppert and So, 2013).

Character Strengths and Flourishing

Character strengths spring from the search for personal values that make us strong and able to face life and adversity

(Park et al., 2004; Lavy, 2019; Karris-Bachik et al., 2020). Character strengths may be considered the psychological ingredients of virtues (Park et al., 2006). Based on this construct, the Values in Action Inventory of Strengths (VIA-IS; Park et al., 2004) was designed and it was also adapted to educational settings (Berkowitz, 2002; Vargas and González-Torres, 2009). Character strength meets the following criteria: 1. Ubiquity—is widely recognized across cultures; 2. Fulfilling—contributes to individual fulfillment, satisfaction, and happiness broadly constructed; 3. Morally valued—is valued in its own right and not for tangible outcomes it may produce; 4. Do not diminish others—elevate others who witness it, producing admiration, not jealousy; 5. Non-felicitous opposite—has obvious antonyms that are “negative”; 6. Traitlike—is an individual difference with demonstrable generality and stability; 7. Measurable—has been successfully measured by researchers as an individual difference; 8. Distinctiveness—is not redundant (conceptually or empirically) with other character strengths; 9. Paragons—is strikingly embodied in some individuals; 10. Prodigies—is precociously shown by some children or youth. 11. Selective absence—is missing altogether in some individuals; 12. Institutions—are the deliberate target of societal practices and rituals that try to cultivate them (Park et al., 2004).

Based on our characterization of flourishing which includes both hedonic and eudemonic wellbeing (Huppert and So, 2013), extant research broadly reports significant relationships between character strengths and the former (Brdar and Kashdan, 2010; Buschor et al., 2013; Hausler et al., 2017; Athota et al., 2020; Baumann et al., 2020; Villacís et al., 2021). Additionally, specific research which addresses wellbeing as flourishing confirms the positive relationship between both variables (Baumann et al., 2020; Niemiec, 2020). Further support comes from multicomponent interventions (Hendriks et al., 2019), which report flourishing amelioration via improving character strengths (VanderWeele, 2020). Along the same lines, extant research reports several experiences focus on improving both character strengths and flourishing (VanderWeele, 2017a,b, 2020; Schutte and Malouff, 2019). Similar results are found when studying the impact of character strengths on thriving as part of flourishing (Hausler et al., 2017).

Another line of studies analyses character strengths structural characteristics and their effects on flourishing (Heintz and Ruch, 2020). These studies report that the strongest cross-sectional associations of flourishing are with hope, curiosity, love, and gratitude (Emmons and Stern, 2013). Although extant longitudinal research reports that character strengths tend to be stable over time (Emmons and Stern, 2013; Wagner et al., 2020), other studies found that character strengths such as humor, spirituality, and prudence may be more susceptible to change (Wagner et al., 2020).

From the above-mentioned, a more in-depth analysis is necessary in order to specify which character strengths relate to flourishing as well as the strength of such a relationship.

Resilience and Flourishing

The role of resilience, whether in protecting against stress, or in contributing to flourishing, has been conceptualized from several

perspectives (e.g., rising above, adaptation, and adjustment; Aburn et al., 2016). Literature broadly supports the key role of resilience in coping with difficulties and in helping individuals to persist while pursuing their goals (Kim et al., 2016; Kachel et al., 2020). Additionally, the latest studies propose biased affective forecasting as a potential mechanism that promotes resilience and flourishing (Colombo et al., 2020).

Extant research confirms the positive relationship between resilience and flourishing (Yildirim and Belen, 2019). Additionally, recent studies analyze its value in personal recovery after health accidents (Rapport et al., 2020), in preventing psychopathological symptoms (Chmitorz et al., 2018), in mental health (Wu et al., 2020), or as a mediator between optimism and subjective wellbeing (He et al., 2013; Miranda and Cruz, 2020). However, a recent meta-analysis on resilience and flourishing dimensions shows that effect sizes among studies are heterogeneous, which points to a large variability within the reported results (Liu et al., 2020).

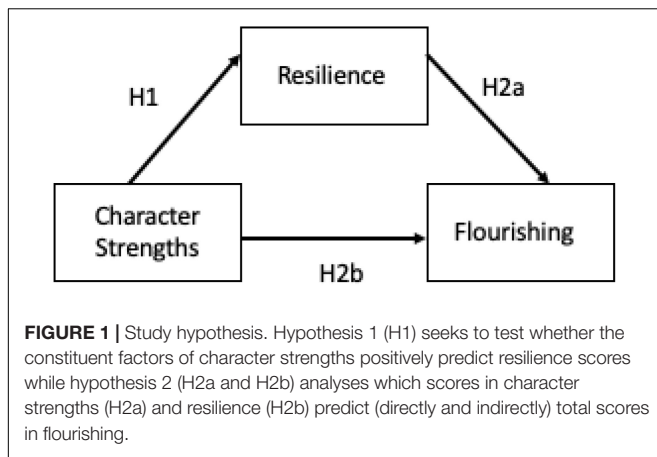
To further analyze resilience and based on the Connor-Davidson Resilience Scale (CD-RISC scale), recent research (de la Fuente et al., 2021a), suggests that resilience consists of two dimensions, that is, reactive and proactive. Reactive resilience comprises behavioral factors pertaining to endurance under adverse conditions (reactive factors). On the other hand, proactive resilience refers to the ability to bounce back and produce changes under unfavorable situations (proactive factors). Reactive resilience, stresses tolerance skills, and spirituality behaviors to predict emotion-focused coping strategies (stress, spirituality). Likewise, proactive resilience, i.e., the ability to adapt to change and perceived control (tenacity, control, change, spirituality), predict problem-focused strategies (de la Fuente et al., 2021a). Consequently, the two dimensions are complementary and necessary, although only the proactive factors could pertain to self-regulatory behavior (de la Fuente et al., 2021a). Additionally, the moderator role of this two-component model has also been tested between individual factors (big five) and the stress experience (de la Fuente et al., 2021b).

Thus, the present paper delves into the distinctive links between resilience dimensions (tenacity, control, change, stress, and spirituality) and flourishing in order to uncover the nuances of this relationship.

Character Strengths and Resilience

Extant studies have already found positive relationships between character strengths and resilience (Chung, 2008; Hausler et al., 2017). Specifically, character strengths show predictive power beyond other related factors (i.e., positive affect, self-efficacy, optimism, social support, self-esteem, life satisfaction) as well as sociodemographic variables (Martínez-Martí and Ruch, 2017).

More recent research with US university students reports a consistent relationship between VIA character strengths and happiness, wellbeing, resilience, academic success, and psychopathologies (Karris-Bachik et al., 2020). Specifically, three functions of character strengths are highlighted when facing adversity: buffering (i.e., the use of character strengths prevents problems); reappraisal (i.e., a person's character strengths explain or reinterpret problems); and resilience (i.e.,



character strengths support the bounce-back from life's setbacks (Colombo et al., 2020).

Despite the foregoing evidence, relations between character strengths and resilience are neither sufficiently established, nor do we understand precisely how these two constructs are related to flourishing. Nevertheless, it seems reasonable to assume that character strengths are foundational to resilient behaviors and to flourishing itself.

Aims and Hypothesis

Consequently, the objectives of the present paper are: (a) to establish predictive relationships between character strengths and resilience; and (b) to determine the joint contribution of these two variables to predict flourishing. Accordingly, *hypothesis 1* (H1) seeks to test whether the constituent factors of character strengths positively predict resilience scores while *hypothesis 2* (H2) analyses which scores in character strengths and resilience predict (directly and indirectly) total scores in flourishing (H2a). Specifically, resilience may be expected to play a mediating role on flourishing with a direct and indirect effect (H2b) (**Figure 1**).

MATERIALS AND METHODS

Participants

A convenience sample of 642 undergraduate students was formed from the two universities involved in this research project; these students subsequently completed questionnaires that assessed ten teaching-learning processes (i.e., ten academic subjects). The participants were enrolled in either Psychology or Primary Education degrees; 83.5% were female, and 16.5% were male. Mean age was 20.13 years ($\sigma_x = 5.8$), and age range was 19–45 years. Participation was anonymous and voluntary. Lecturers from various departments were invited to participate, and those who agreed then extended the invitation to their students. Participating lecturers and students were awarded a Certificate of Participation. Online questionnaires were applied to assess each specific teaching-learning process.

Instruments

The VIA Inventory of Strengths

The short form of the VIA Inventory of Strengths (VIA-IS-72) (Peterson and Seligman, 2004) includes 72 items, 24 factors and 6 dimensions, that allow subjects to self-assess on 24-character strengths. For each character strength, there are three Likert-style items with five possible responses, where 1 = “Very Much Unlike Me” and 5 = “Very Much Like Me.” For example, the statements include “I am a highly disciplined person” (self-regulation). According to measures of internal consistency reliability and validity, the VIA-72 is substantially equivalent to the original, long VIA-IS, as verified by its developers (VIA Institute on Character, 2020). VIA-72 psychometric report showed a mean Cronbach alpha (reliability) of 0.75 for the 24 scales, and Omega index of 0.71; coefficients ranged from 0.70 to 0.87 (VIA Institute on Character, 2020). The Spanish version of the VIA-72 was obtained for this study [CFI = 0.96, TLI = 0.94, RMSEA = 0.05, and SRMR = 0.06].

The Connor-Davidson Resilience Scale

The validated, Spanish version (Manzano-García and Ayala-Calvo, 2013) of the CD-RISC Scale (Connor and Davidson, 2003) was applied to measure resilience. Reliability and validity values are adequate in Spanish samples, and there is a five-factor structure: F1: Persistence/tenacity and strong sense of self-efficacy (*tenacity*); F2: Emotional and cognitive control under pressure (*stress*); F3: Adaptability/ability to bounce back (*change*); F4: Perceived Control (*control*), and F5, Spirituality (*spirituality*). Subjects are asked to rate statements such as “I give my best effort, no matter what the outcome may be” or “I believe I can achieve my goals, even if there are obstacles”. This Likert-type scale contains 25 items and five factors: (1) personal competence, high standards and *tenacity* ($\alpha = 0.80$), (2) self-confidence, tolerance of negative affect and strengthening effects of *stress* ($\alpha = 0.75$), (3) positive acceptance of *change*, and secure relationships ($\alpha = 0.77$), (4) *control* ($\alpha = 0.71$), and (5) *spiritual* influences (0.71) and an Omega index of 0.76. Adequate reliability and validity values had been obtained in Spanish samples, and a five-factor structure emerged [Chi-square = 1,619, 170; Degrees of freedom (350–850) = 265; $p < 0.001$; Ch/Df = 6,110; SRMR (Standardized Root Mean-Square) = 0.062; NFI (Normed Fit Index) = 0.957; RFI (Relative Fix Index) = 0.948; IFI (Incremental Fix Index) = 0.922; TLI (Tucker Lewis index) = 0.980; CFI (Comparative fit index) = 0.920; RMSEA (Root Mean Square Error) = 0.063; HOELTER = 240 ($p < 0.05$) and 254 ($p < 0.01$)].

The Flourishing Scale

A Spanish version validated in Spanish populations (Garzón-Umerenkova, 2016; Pozo-Muñoz et al., 2016; Garzón-Umerenkova et al., 2017) of the Flourishing Scale (Biswas-Diener et al., 2009) was applied. This scale seeks to measure flow or flow state (flourishing) and contains eight items on a five-point Likert scale. Responses range from 1 (strongly disagree) to 5 (strongly agree). Subjects are asked to rate statements such as “I have a useful and meaningful life” or “I am a good person and I live a good life”. Cronbach alpha for the Spanish sample is 0.85; and an Omega index of 0.81. The unidimensionality of

the scale and the metric invariance in the evaluated samples was confirmed [Chi-square = 79,392; Degrees of freedom (44–24) = 20; $p < 0.001$; Ch/Df = 3,970; SRMR (Standardized Root Mean-Square) = 0.052; NFI (Normed Fit Index) = 0.946; RFI (Relative Fit Index) = 0.953; IFI (Incremental Fit Index) = 0.959; TLI (Tucker Lewis index) = 0.955; CFI (Comparative fit index) = 0.958; RMSEA (Root Mean Square Error) = 0.039; HOELTER = 757 ($p < 0.05$) and 905 ($p < 0.01$)].

Procedure

University students voluntarily completed the validated questionnaires, after signing their informed consent via the online platform (de la Fuente et al., 2015). Five specific teaching-learning processes were assessed, corresponding to different academic subjects that were taken over a 2-year period. In September–October of 2018 and 2019, presage variable (characters strength) were assessed. Process variable (resilience) were measured in February–March of 2019 and 2020, and product variable (flourishing) in May–June of 2019 and 2020. Approval for this procedure was obtained from the Ethics Committee of University of Navarra, within a larger R&D Project (2018–2021; ref. 2018.170).

Data Analysis

An *ex post facto* prospective and transversal design (Lohr, 1999) was used. The *ex post facto* design come to solve the problem that occurs when the variable of interest has already occurred and/or it is not ethical to cause it. It is *prospective* because the independent variable precedes the dependent variable in the analysis. Additionally, it is cross-sectional because longitudinal follow-up is not possible and the data was collected within a short period of time. Three types of analyses were performed. The usual assumptions of regression analysis were tested beforehand.

(1) *Preliminary analysis*. First, the quality of the data was explored by testing for outliers and missing cases. *Univariate* outliers were tested by calculating the typical scores of each variable, considering cases with Z scores outside the ± 3 range to be potentially atypical cases (Tabachnick et al., 2007). On the other hand, the Mahalanobis distance (D2) was used to detect atypical combinations of variables (atypical multivariate cases), a statistical measure of an individual's multidimensional distance from the centroid or mean of the given observations (Lohr, 1999). This procedure detects significant distances from the typical combinations or centroids of a set of variables. Literature suggests removing univariate and multivariate outliers, or reassigning them to the nearest extreme score (Weston and Gore, 2006). The procedure was carried out using SPSS (v.26, IBM, Armonk, NY, United States), which provides a specific routine for missing values analysis that determines the magnitude of missing values and whether they are presented in a systematic or random manner.

Assumptions related to sample size, independence of errors, univariate, and multivariate normality, linearity, multicollinearity, recursion, and interval measurement level were also evaluated, showing acceptable reliability levels. Regarding the sample size, inclusion of 10–20 cases per parameter is recommended, and at least 200 observations (Kline, 2005).

Independence of error means that the error term of each endogenous variable must not be correlated with other variables. In order to test for univariate normality, the distribution of each observed variable was examined, and its indices of asymmetry and kurtosis. Asymmetry values greater than 3 and kurtosis greater than 10 suggest that the data should be transformed. Additionally, values less than 70 on the Mardia multivariate index indicate that the distance from the multivariate normality is not a critical deterrent to this analysis (Mardia, 1970). Although one of the assumptions is the level of interval measurement, in some cases, variables measured at a nominal or ordinal level were used, as long as the distribution of scores, particularly of the dependent variables, were not markedly asymmetric.

(2) *Predictive analysis*. For Hypothesis 1, multiple regression analysis was applied using SPSS (v.25)

(3) *Structural prediction and mediational models*. Hypotheses 2 were tested using a Structural Equation Model (SEM) and a mediational model for complex measurement (Ato and Vallejo, 2011). Model fit was assessed by first examining the ratio of chi-square to degrees of freedom, then the Comparative Fit Index (CFI), Normed Fit Index (NFI), Incremental Fit Index (IFI), Relative Fit Index (RFI) and Tucker-Lewis Index (TLI). All fit measures of the incremental model were above the suggested limit of 0.90 (Bentler, 1990). The value of the Comparative fit index (CFI) was equal to 0.928, which is also satisfactory. The results of the original scale were replicated. The value of the Root Mean Square Error of Approximation (RMSEA) was 0.08 less than the warning value of 0.09 (Ho, 2006). Ideally, these should be greater than 0.90. The Hoelter Index was also used to determine the adequacy of sample size. AMOS (v.22) was used for these analyses. Keith (2006) proposed the following beta coefficients as research benchmarks for *direct effects*: less than 0.05 is considered too small to be meaningful, above 0.05 is small but meaningful, above 0.10 is moderate, and above 0.25 is large. For *indirect effects*, we used Kenny's (2012) definition which considers them as the product of two effects; using Keith's benchmarks above, we propose a small indirect effect = 0.003, moderate = 0.01, and large = 0.06, values that are significant in the sphere of education.

RESULTS

Results from the analyses applied in order to test normality, a prerequisite for linear analysis, show an adequate distribution of sample variability (see Table 1).

Predictive Value of Character Strengths on Resilience, and Flourishing

Regression analyses showed, that D2 (*emotional strength*) was the strength dimension that established significant relationships with four of the five resilience dimensions (*adaption to change*), total resilience and flourishing. Specifically, it was the strongest predictor of *stress tolerance*, *flourishing* and *tenacity*. D1 (*cognitive strength*) explained *resilience* strongly and *adaption to change*. D3 (*interpersonal strength*) was only related to *control* and D6 (*strength from the meaning of life*) predicted *resilience*. As expected, D5 (*strength against excess*) is negatively related

TABLE 1 | Descriptive values of the variables under study ($n = 641$).

Variable	Min	Max	<i>M</i> (<i>SD</i>)	Mean std. error	Asymmetry	Standard asymmetry error	Kurtosis	Standard kurtosis error	Kolmogoroff-Sminoff
D1	1.93	5.00	3.655(0.595)	0.059	0.116	0.116	−0.303	0.343	0.065, $p < 0.200$
D2	1.80	5.00	3.823(0.551)	0.039	−0.105	0.175	0.044	0.348	0.056, $p < 0.200$
D3	2.74	5.00	4.000(0.530)	0.037	−0.205	0.174	−0.506	0.346	0.061, $p < 0.200$
D4	1.89	5.00	3.951(0.555)	0.040	−0.316	0.175	0.236	0.349	0.035, $p < 0.200$
D5	1.96	5.00	3.552(0.549)	0.038	0.233	0.170	0.070	0.377	0.064, $p < 0.200$
D6	2.10	5.00	3.578(0.603)	0.042	0.049	0.172	−0.281	0.343	0.043, $p < 0.200$
SCTOT	2.89	4.88	3.731(0.448)	0.038	0.355	0.206	−0.456	0.408	0.070, $p < 0.200$
Stress	1.43	5.00	3.627(0.539)	0.013	−0.162	0.061	−0.06	0.121	0.067, $p < 0.200$
Spirituality	1.00	5.00	3.257(1.00)	0.024	−0.071	0.060	−0.494	0.120	0.087, $p < 0.155$
Tenacity	1.00	5.00	3.912(0.591)	0.014	−0.450	0.060	0.286	0.121	0.074, $p < 0.174$
Change	1.63	5.00	3.942(0.613)	0.015	−0.489	0.060	0.462	0.120	0.058, $p < 0.188$
Control	1.20	5.00	3, 893(0.755)	0.018	−0.516	0.060	0.499	0.120	0.69, $p < 0.196$
RESTOT	1.52	4.86	3.725(0.476)	0.121	−0.492	0.062	0.558	0.125	0.049, $p < 0.200$
FLUORISHING	2.13	5.00	4.100(0.629)	0.031	−0.569	0.121	0.015	0.241	0.045, $p < 0.150$

D1, cognitive strength (Wisdom); D2, emotional strength (Courage); D3, interpersonal strength (Humanity); D4, Civic strength (Justice); D5, strengths that protect against excess (Temperance); D6, strength from the meaning of life (Transcendence). SCTOT, Total Strength Character; RESTOT, Total Resilience.

TABLE 2 | Multiple regression of the dimensions of character strengths, resilience and flourishing.

Strengths dimensions	Stress tolerance	Spiritual	Tenacity	Adaptation to change	Control	Total resilience	Flourishing
D1	0.193	0.110	0.028	0.295*	0.242	0.506***	0.028
D2	0.565***	0.344**	0.433*	0.215	0.340**	0.248**	0.443***
D3	0.003	0.137	0.087	0.117	0.276***	−0.108	0.087
D4	−0.034	−0.013	−0.110	−0.196	0.014	0.170	−0.110
D5	−0.161	−0.056	−0.116	−0.148	−0.266**	0.131	−0.116
D6	0.025	0.000	0.217	0.090	0.201	0.581***	0.035
$F(6, 124) =$	15,646***	7,185***	10,05***	15,646***	7,790***	4,501***	19,528***
R^2	0.421	0.249	0.337	0.208	0.262	0.170	0,327

D1, cognitive strength (Wisdom); D2, emotional strength (Courage); D3, interpersonal strength (Humanity); D4, Civic strength (Justice); D5, strengths that protect against excess (Temperance); D6, strength from the meaning of life (Transcendence). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

to control. Control and resilience were the criteria variables predicted by more character strengths dimensions. So, control was explained by emotional, interpersonal, and strength against excess while resilience was predicted by cognitive, emotional, and strength from the meaning of life. Overall, character strengths explained 42% of the variance of stress tolerance, 33.7% of tenacity and 32.7% of flourishing (see Table 2).

Predictive Value of the Factors of Character Strengths on Resilience Factors and Flourishing

In more detail, multiple linear regression analyses showed that the factors that most significantly and positively predicted total resilience were F9 (vitality and zest), F6 (bravery) and F14 (sense of justice), and F19 (self-regulation) and F17 (modesty, humility), negatively. The most significant and negative predictors of total flourishing were F20 (appreciation for beauty and excellence) and F9, positively. Referring to specific factors of resilience, F9 and F7 (perspective and diligence) significantly predicted tenacity. Stress tolerance was positively predicted by factors F12

(emotional intelligence) and F9, and negatively predicted by F13 (citizenship). Adaptation to change was positively predicted by factors F12 and F15 (leadership), and negatively by F17, F18 (prudence), and F24 (spirituality). The resilience factor with the highest number of predictive character strengths was perceived control, being predicted by F9, F14, F10 (love and be loved), F2 (love for knowledge) and F7 (positively) and by F1 (curiosity), F11 (kindness), F20 (appreciation for beauty), F22 (hope) and F17 (negatively). Factors F21 (gratitude) and F19 were positive predictors of spirituality (F24), whereas F24 did not predict them (see Table 3).

Structural and Mediational Prediction Model of Character Strengths, Resilience, and Flourishing Model Testing

Four structural predictive models were tested. The first referred to personal strengths predicting resilience. The second model tested resilience as a predictor of flourishing. The third model tested prediction between resilience and flourishing and the fourth,

TABLE 3 | Multiple regression between the dimensions and factors of strengths, resilience and flourishing.

PV	Stress	Spirituality	Tenacity	Change	Control	Total resilience	Flourishing
D1							
F1					−0.378**		
F2					0.242*		
F3							
F4							
F5							
D2							
F6						0.214*	
F7			0.277*		0.173*		
F8							
F9	0.291*		0.399**		0.478***	0.409**	0.567***
D3							
F10					0.291**		
F11					−0.316*		
F12	0.343**			0.293*			
D4							
F13	−0.275*						
F14					0.314*	0.282*	
F15				0.239*			
D5							
F16							
F17				−0.194*	−0.174*	−0.157*	
F18				−0.185*			
F19		0.179*					
D6							
F20					−0.205*	−0.239*	−0.219*
F21		0.342**					
F22					−0.205*		
F23							
F24				−0.173*			
<i>F</i>	<i>F</i> (24,111) = 5.563***	<i>F</i> (24,114) = 2.961***	<i>F</i> (24,111) = 4.474***	<i>F</i> (24,112) = 2.709***	<i>F</i> (24,112) = 4.214***	<i>F</i> (24,106) = 4.336***	<i>F</i> (24,137) = 4.494***
<i>R</i> ²	0.208	0.384	0.492	0.367	0.475	0.497	0.512

D1 = Cognitive strength (Wisdom): F1, Curiosity, interest in the world; F2, Love for knowledge and learning; F3, Judgment, critical thinking, open-mindedness; F4, Ingenuity, originality, practical intelligence; F5, Perspective; **D2 = Emotional strength (Courage):** F6, Bravery; F7, Perspective and diligence; F8, Integrity, honesty, authenticity; F9, Vitality and zest; **D3 = interpersonal strength (Humanity):** F10, Love, ability to love and be loved; F11, Kindness, friendliness, generosity; F12, Emotional, personal, and social intelligence; **D4 = Civic strength (Justice):** F13, Citizenship, social responsibility, loyalty, teamwork; F14, Sense of justice, fairness; F15, Leadership; **D5 = Strengths that protect against excess (Temperance):** F16, Ability to forgive, mercy; F17, Modesty, humility; F18, Prudence, discretion, caution; F19, Self-control, self-regulation; **D6 = Strength from the meaning of life (Transcendence):** F20, Appreciation for beauty and excellence, capacity for wonder; F21, Gratitude; F22, Hope, optimism, future-mindedness; F23, Sense of humor; F24, Spirituality, religious sense. **PV = Predictors Variables.** Bold values indicate predictor variables. **p* < 0.05; ****p* < 0.01; *****p* < 0.001.

tested strengths and resilience combined, as joint predictors of flourishing (Table 4). As shown in Table 4, the data fit best to the fourth model as hypothesis 2 stated.

Direct Effect

Direct statistic effects showed *character strengths* to be significant predictors of resilience (0.51) and flourishing (0.43), and resilience was also a predictor of flourishing (0.60) (see Table 5 and Figure 2).

Indirect Effect

Additionally, indirect effects were also observed between character strengths and the components of resilience.

Mediation Effects

These results demonstrate that character strengths have a significant overall effect on flourishing (0.60), both direct and

indirect. Thus, character strengths have a direct effect on flourishing (0.38) and an indirect effect -through resilience- on flourishing (0.22).

DISCUSSION

These results partially corroborate findings from previous research, which reveal relationships between character strengths and resilience (Chung, 2008) bringing to light the importance of those strengths.

Character Strengths and Resilience

Regarding *Hypothesis 1*, which studies the positive linear prediction of strengths on resilience, results uphold that character strengths are foundational to resilient behavior, by showing how character strengths, as a multidimensional construct,

TABLE 4 | Models of structural linear results.

Chi ²	FG	CH/df	SRMR	<i>p</i> <	NFI	RFI	IFI	TLI	CFI	RMSEA	HOELT.	<i>p</i> < 0.05; <i>p</i> < 0.01
(1) 120,687	34	3,55	0.09	0.000	0.898	0.835	0.924	0.875	0.923	0.098	0.108	0.125
(2) 75,147	9	8,35	0.10	0.000	0.910	0.789	0.920	0.810	0.918	0.062	0.421	0.522
(3) 106,774	9	11,864	0.12	0.000	0.958	0.901	0.961	0.909	0.961	0.075	0.303	0.338
(4) 256,454	52	4,932	0.05	0.000	0.951	0.954	0.956	0.955	0.966	0.047	0.500	0.526

Model 1: Strengths, Resilience; Model 2: Strengths, Flourishing; Model 3: Resilience, Flourishing; Model 4: Strengths, Resilience, Flourishing.

TABLE 5 | Total, direct, and indirect effects of the variables under study, and 95% bootstrap confidence intervals (CI).

Predictive variable	Criterion variable	Total effect	CI (95%)	Direct effect	CI (95%)	Indirect effect	CI (95%)	Results effects	CI (95%)
CS - >	Resilience	0.51	[0.39, 0.74]	0.51	[0.39, 0.74]	0.00	[-0.03, 0.02]	Direct only	[0.39, 0.74]
R - >	Flourishing	0.43	[0.37, 0.54]	0.43	[0.37, 0.54]	0.00	[-0.03, 0.02]	Direct only	[0.37, 0.54]
CS- >	Flourishing	0.60	[0.41, 0.76]	0.38	[.23, 0.48]	0.22	[0.16, 0.34]	Partial mediation	[0.16, 0.34]
CS- >	Stress	0.34	[0.32, 0.47]	0.00	[-0.04, 0.12]	0.34	[0.26, 0.48]	Full mediation	[0.26, 0.48]
CS- >	Spirituality	0.09	[0.02, 0.14]	0.00	[-0.03, 0.09]	0.09	[0.02, 0.14]	Full mediation	[0.02, 0.14]
CS- >	Tenacity	0.42	[0.32, 0.57]	0.00	[-0.04, -0.08]	0.42	[0.32, 0.57]	Full mediation	[0.32, 0.57]
CS- >	Change	0.41	[0.20, 0.38]	0.00	[-0.03, 0.04]	0.41	[0.20, 0.38]	Full mediation	[0.20, 0.38]
CS- >	Control	0.32	[0.43, 0.21]	0.00	[-0.07, 0.08]	0.30	[0.43, 0.21]	Full mediation	[0.43, 0.21]

CS, Character strengths; R, Resilience; CI, confidence interval. Bootstrapping sample size = 643.

significantly, and jointly predict resilience, as prior evidence had founded (Niemiec, 2020; Seale et al., 2020). Thus, the essential value of the present study lies in uncovering a more precise specification of this relationship, from a large sample of Spanish students. These results reveal how each factor of resilience was predicted by one or more character strengths. Specifically, emotional strength (D2) was the strongest predictor of resilient tenacity; cognitive strength (D1) enabled adaption to change; emotional strength (D2) and interpersonal strength (D3) positively predicted control, but the opposite occurred with the strength of temperance (D5); and finally, emotional strength (D2) predicted resilient spirituality. Although with less statistical power, these dimensions also predicted flourishing, confirming previous research (Park and Peterson, 2009).

The factors of character strengths also proved their predictive value for various components of resilience and flourishing. Referring to specific factors of resilience, F9 (vitality and zest) and F7 (perseverance) significantly predicted resilient tenacity. These results suggest a similar behavioral component between the two psychological constructs, but with factorial independence (Martínez-Martí and Ruch, 2017).

Factors F12 (emotional, personal and social intelligence) and F9 (bravery) acted as positive predictors of stress tolerance, whereas F13 (citizenship, social responsibility, loyalty, teamwork) was a negative predictor. This result concurs with abundant evidence that has established emotional intelligence as a protective factor against stress (Elshall et al., 2020; Trigueros et al., 2020). However, this potentially contradictory relationship between citizenship and resilience would need further analysis in future research. Factors F12 (emotional intelligence) and F15 (leadership) positively predicted adaptation to change, while F17 (modesty, humility), F18 (prudence, discretion, caution) and F24 (spirituality, faith, religious sense) were negative predictors. This result is of great interest because it shows that, there

is strong likelihood that personal strengths also depend on the circumstance in which one lives. Prudence can also be expected to protect against inappropriate decisions or changes (Reeves et al., 2014).

The resilience factor with the greatest number of predictive character strengths was perceived control, predicted by F9, F14, F10, F2, and F7 (positively) and by F1, F11, F20, F22, and F17 (negatively). These results seem to concur with evidence that personality or character factors predict resilience (Goodman et al., 2017). It is also of interest that factors F21 (gratitude) and F19 (self-control/self-regulation) were positive predictors of spirituality, concurring with previous studies on the relationship between self-regulation and spirituality (Desmond et al., 2013; Boyatzis et al., 2021). However, F24 (spirituality, religious sense) did not predict resilient spirituality, which would suggest that the psychological constructs that are being assessed are somewhat distinct.

One noteworthy result is the negative prediction found between temperance (character strength) and adaptation to change (resilience). This result might indicate that excessive regulation and control, in situations of change, would hinder good adjustment, because of perceived loss of control (Koltai and Stuckler, 2020), which could become deregulatory behavior (de la Fuente-Arias, 2017). Nonetheless, this question should be further studied in future research.

Strengths, Resilience, and Flourishing

Hypothesis 2 establishes the combination of character strengths and resilience in a significant, positive relationship with flourishing. The results suggest that the behavioral components inherent in the construct of character strengths predict resilience, and jointly, these two constructs predict flourishing. In short, they predict wellbeing, with similar elements present in both psychological constructs (Chung,

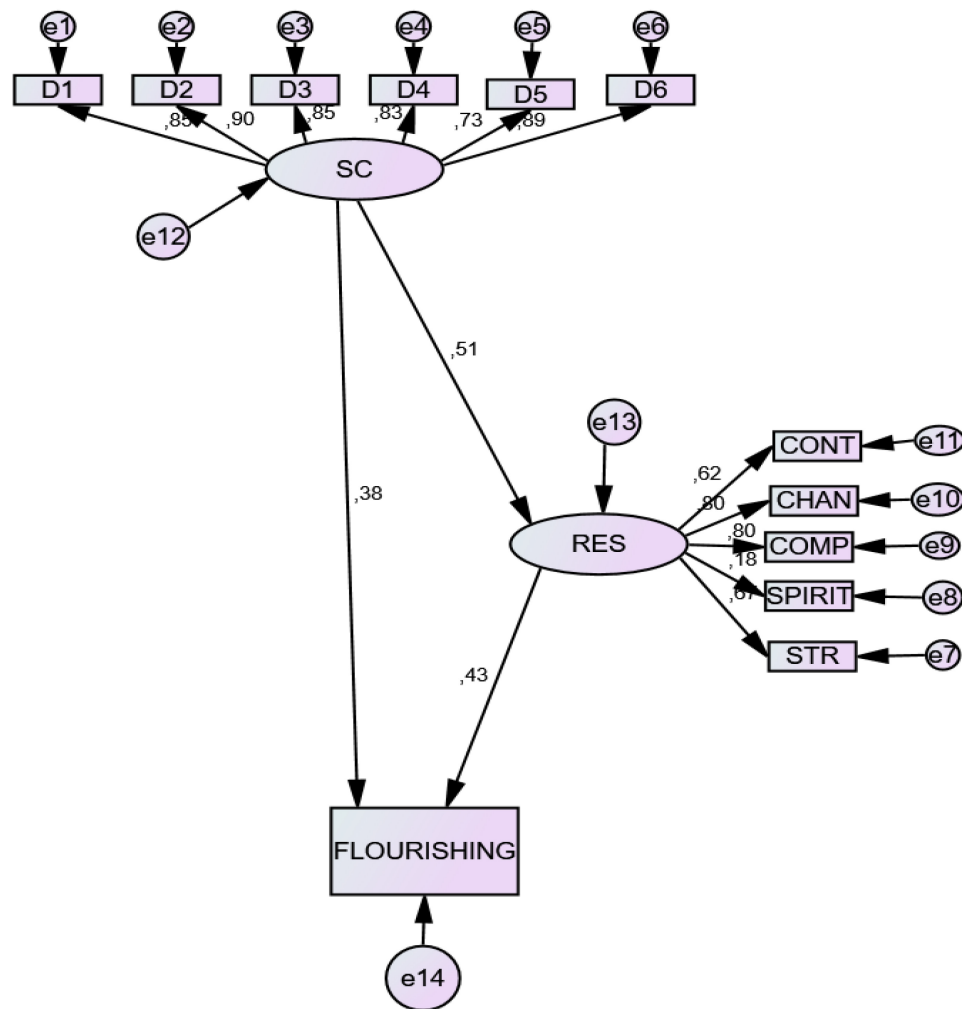


FIGURE 2 | Structural predictive model of Character Strengths, Resilience, and Flourishing. D1, cognitive strength (Wisdom); D2, emotional strength (Courage); D3, interpersonal strength (Humanity); D4, Civic strength (Justice); D5, strengths that protect against excess (Temperance); D6, strength from the meaning of life (Transcendence). STR, Gestion of Stress; SPIRIT, Spirituality; COMPT, Tenacity; CHAN, Adaptation to change; CONT, Self-Control.

2008; Martínez-Martí and Ruch, 2017; Karris-Bachik et al., 2020). Nonetheless, independence among the three constructs is also clearly shown. Previous research had produced evidence for some of these results. For example, a predictive relationship between strengths and wellbeing had been demonstrated, as well as positivity (Botha, 2020; Wagner et al., 2020). Resilience has consistently appeared as a predictor of wellbeing (Harms et al., 2018). In fact, the resilience variable has appeared as a mediating variable in the relationship between strengths and flourishing; this mediating effect is consistent with other previous work (de la Fuente et al., 2021b).

Limitations and Future Research

Despite the consistent results, this investigation has several limitations. On the one hand, it should be replicated in broader samples, to compensate for the gender imbalance which is

typical of the university samples used here. On the other hand, factor invariance in clinical samples should also be verified. Only in that case can these results be applicable to such samples. Lastly, it remains to be established whether these relationships are generalizable to other populations or areas such as leadership styles (Abbas et al., 2020b), the implementation of coping strategies in the digital age (Abbas et al., 2020a) or coping strategies of rural students at urban universities (Ali et al., 2021).

Implications for Psycho-Educational Intervention

According to Seligman and Csikszentmihalyi (2000) the field of positive psychology, at the subject level, is about valued subjective experiences: wellbeing, contentment, and satisfaction (in the past); hope and optimism (for the future); and flow and happiness (in the present) (2). At the individual level,

it is about positive individual traits: the capacity for love and vocation, courage, interpersonal skill, aesthetic sensibility, perseverance, forgiveness, originality, future mindedness, spirituality, high talent, and wisdom. At the group level, it is about the civic virtues and the institutions that move individuals toward better citizenship: responsibility, nurturance, altruism, civility, moderation, tolerance, and work ethic (Seligman and Csikszentmihalyi, 2000, p. 5).

The present study at the individual level supports the importance of, and a connection between the three constructs analyzed here. For this reason, the university academic curriculum would benefit from inclusion of these three variables, in order to ensure that students receive whole-person training. The importance of academic intervention in improving personal strengths has already been documented (Vargas and González-Torres, 2009; Lavy, 2019). The formation of university students in the role of character strengths, as an essential psychological tool, for the achievement of psychological wellbeing and flourishing, through the cultivation of resilience, is very important in our university system. Even more so, at a time when hedonic wellbeing proliferates in our university classrooms, as the only way to achieve personal wellbeing.

CONCLUSION

These findings have made it possible to corroborate in great detail how the distinct character strength dimensions allow the resilience dimensions in a sample of Spanish students to be predicted. In this sense, the present study showed that: (a) personal strengths can also depend on contextual circumstances; (b) prudence protects against risky decisions; (c) personality or character factors predict resilience; (d) gratitude and self-regulation predict positively spirituality; and (e) the perception of loss of control in situations of change together with excessive regulation and external control hinder the adequate adaptation to new circumstances. Finally, it was established that strengths and resilience present a significant positive predictive effect on flourishing, although clearly denoting the interdependence between the three constructs.

REFERENCES

- Abbas, A., Eliyana, A., Ekowati, D., Saud, M., Raza, A., and Wardani, R. (2020b). Data set on coping strategies in the digital age: the role of psychological well-being and social capital among university students in Java Timor, Surabaya, Indonesia. *Data Brief* 30:105583. doi: 10.1016/J.DIB.2020.105583
- Abbas, A., Saud, M., Suhariadi, F., Usman, I., and Ekowati, D. (2020a). Positive leadership psychology: authentic and servant leadership in higher education in Pakistan. *Curr. Psychol.* 1–13. doi: 10.2307/j.ctt1v2xw3s.4
- Aburn, G., Gott, M., and Hoare, K. (2016). What is resilience? An integrative review of the empirical literature. *J. Adv. Nurs.* 72, 980–1000. doi: 10.1111/jan.12888
- Ali, S., Sarker, M. F. H., Islam, M. S., Islam, M. K., and Al Mahmud, R. (2021). Pursuing higher education: adaptation challenges and coping strategies of rural students at urban universities in Bangladesh. *Tertiary Educ. Manag.* 27, 91–106. doi: 10.1007/S11233-021-09067-3

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 Investigación, University of Navarra. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JF, BU, EL, and MG-T contributed to the conceptualization, data curation, formal analysis, investigation, methodology, and writing—original draft. JF contributed to the design, supervision, data curation, formal analysis, methodology, and was responsible for the overall content as the guarantor. BU, EL, MG-T, RA-G, and AB contributed to investigation, methodology, and resources. All authors contributed to reviewing and editing the final manuscript and to the article and approved the submitted version.

FUNDING

This research was funded by the R&D Project PGC2018-094672-B-I00, University of Navarra (Ministry of Science and Education, Spain), the R&D Project UAL18-SEJ-DO31-A-FEDER (University of Almería, Spain), and the European Social Fund.

ACKNOWLEDGMENTS

The researchers would like to acknowledge the support of participants for their time and willingness to support research.

- Allison, L., Waters, L., and Kern, M. L. (2020). Flourishing classrooms: applying a systems-informed approach to positive education. *Contemp. Sch. Psychol.* 25, 395–405. doi: 10.1007/s40688-019-00267-8
- Athota, V. S., Budhwar, P., and Malik, A. (2020). Influence of personality traits and moral values on employee well-being, resilience and performance: a cross-national study. *Appl. Psychol.* 69, 653–685. doi: 10.1111/apps.12198
- Ato, M., and Vallejo, G. (2011). Los efectos de terceras variables en la investigación psicológica. *Anal. Psicol. Ann. Psychol.* 27, 550–561.
- Baumann, D., Ruch, W., Margelisch, K., Gander, F., and Wagner, L. (2020). Character strengths and life satisfaction in later life: an analysis of different living conditions. *Appl. Res. Qual. Life* 15, 329–347. doi: 10.1007/s11482-018-9689-x
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull.* 107:238. doi: 10.1037/0033-2909.107.2.238
- Berkowitz, M. W. (2002). “The science of character education,” in *Bringing in a New Era in Character Education*, ed. W. Damon (Palo Alto, CA: Hoover Institution Press), 43–63.

- Biswas-Diener, R., Kashdan, T. B., and King, L. A. (2009). Two traditions of happiness research, not two distinct types of happiness. *J. Posit. Psychol.* 4, 208–211. doi: 10.1080/17439760902844400
- Botha, T. (2020). *Flourishing Beyond Borders: Character Strengths, Resilience and Self-Perceived Well-Being of the Accompanying Expatriate Partner During International Relocation*. Doctoral dissertation. Potchefstroom: North-West University, South Africa.
- Boyatzis, R. E., Goleman, D., Dhar, U., and Osiri, J. K. (2021). Thrive and survive: assessing personal sustainability. *Consult. Psychol. J. Pract. Res.* 73, 27–50. doi: 10.1037/cpb0000193
- Brdar, I., and Kashdan, T. B. (2010). Character strengths and well-being in Croatia: an empirical investigation of structure and correlates. *J. Res. Pers.* 44, 151–154. doi: 10.1016/j.jrp.2009.12.001
- Buschor, C., Poyer, R. T., and Ruch, W. (2013). Self- and peer-rated character strengths: how do they relate to satisfaction with life and orientations to happiness? *J. Posit. Psychol.* 8, 116–127. doi: 10.1080/17439760.2012.758305
- Chaves, C. (2021). “Wellbeing and Flourishing” in *The Palgrave Handbook of Positive Education*, eds M. L. Kern and M. L. Wehmeyer (Cham: Palgrave Macmillan), 3–11. doi: 10.1007/978-3-030-64537-3_11
- Chmitorz, A., Kunzler, A., Helmreich, I., Tüscher, O., Kalisch, R., Kubiak, T., et al. (2018). Intervention studies to foster resilience—A systematic review and proposal for a resilience framework in future intervention studies. *Clin. Psychol. Rev.* 59, 78–100. doi: 10.1016/j.cpr.2017.11.002
- Chung, H. F. (2008). *Resiliency and Character Strengths Among College Students*. Doctoral dissertation. Tucson, AZ: University of Arizona.
- Ciarrochi, J., Atkins, P. W. B., Hayes, L. L., Sahdra, B. K., and Parker, P. (2016). Contextual positive psychology: policy recommendations for implementing positive psychology into schools. *Front. Psychol.* 7:1561. doi: 10.3389/fpsyg.2016.01561
- Ciocanel, O., Power, K., Eriksen, A., and Gillings, K. (2017). Effectiveness of positive youth development interventions: a meta-analysis of randomized controlled trials. *J. Youth Adolesc.* 46, 483–504. doi: 10.1007/s10964-016-0555-6
- Colombo, D., Fernández-Álvarez, J., Suso-Ribera, C., Cipresso, P., García-Palacios, A., Riva, G., et al. (2020). Biased affective forecasting: a potential mechanism that enhances resilience and well-being. *Front. Psychol.* 11:1333. doi: 10.3389/fpsyg.2020.01333
- 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
- de la Fuente-Arias, J. (2017). Theory of self- vs. externally- regulated learning TM: fundamentals, evidence, and applicability. *Front. Psychol.* 8:1675. doi: 10.3389/fpsyg.2017.01675
- de la Fuente, J., González-Torres, M. C., Artuch-Garde, R., Vera-Martínez, M. M., Martínez-Vicente, J. M., and Peralta-Sánchez, F. J. (2021b). Resilience as a buffering variable between the big five components and factors and symptoms of academic stress at University. *Front. Psychiatry* 12:600240. doi: 10.3389/fpsyg.2021.600240
- de la Fuente, J., López, M., Zapata, L., Sollinas, G., and Fadda, S. (2015). Improving mental Health trough and online self-assessment and self-help e-Utility in university Students. *Prog. Educ.* 33, 63–74.
- de la Fuente, J., Sander, P., Martínez-Vicente, J. M., Vera, M., Garzón-Umerenkova, A., and Fadda, S. (2017). Combined effect of levels in personal self-regulation and regulatory teaching on meta-cognitive, on meta-motivational, and on academic achievement variables in undergraduate students. *Front. Psychol.* 8:232. doi: 10.3389/fpsyg.2017.00232
- de la Fuente, J., Santos, F. H., Garzón-Umerenkova, A., Fadda, S., Solinas, G., and Pignata, S. (2021a). Cross-sectional study of resilience, positivity and coping strategies as predictors of engagement-burnout in undergraduate students: implications for prevention and treatment in mental well-being. *Front. Psychiatry* 12:596453. doi: 10.3389/fpsyg.2021.596453
- Desmond, S. A., Ulmer, J. T., and Bader, C. D. (2013). Religion, self-control, and substance use. *Deviant Behav.* 34, 384–406. doi: 10.1080/01639625.2012.726170
- Diener, E. (2009). “Subjective well-being,” in *The Science of Well-Being*, ed. E. Diener (Dordrecht: Springer), 11–58. doi: 10.1007/978-90-481-2350-6
- Dodge, R., Daly, A. P., Huyton, J., and Sanders, L. D. (2012). The challenge of defining wellbeing. *Int. J. Wellbeing* 2, 222–235. doi: 10.5502/ijw.v2i3.4
- Donaldson, S. I., Dollwet, M., and Rao, M. A. (2015). Happiness, excellence and optimal human functioning revisited: examining the peer-reviewed literature linked to positive psychology. *J. Posit. Psychol.* 10, 185–195. doi: 10.1080/17439760.2014.943801
- Duckworth, A. L., Steen, T. A., and Seligman, M. E. P. (2005). Positive psychology in clinical practice. *Annu. Rev. Clin. Psychol.* 1, 629–651. doi: 10.1146/annurev.clinpsy.1.102803.144154
- Elshall, S. E., El-Kholy, S. M., and Dorgham, L. S. (2020). The relationship among emotional intelligence, stress and coping strategies for nursing students. *J. Educ. Res. Rev.* 8, 45–56. doi: 10.33495/jerr_v8i4.20.135
- Emmons, R. A., and Stern, R. (2013). Gratitude as a psychotherapeutic intervention. *J. Clin. Psychol.* 69, 846–855. doi: 10.1002/jclp.22020
- Garzón-Umerenkova, A. (2016). *Flourishing Scale Inventory, FSI; Spanish Version*. Bogotá: Fundación Universitaria Konrad Lorenz.
- Garzón-Umerenkova, A. G., de la Fuente, J., Martínez-Vicente, J. M., Sevillano, L. Z., Pichardo, M. C., and García-Berbén, A. B. (2017). Validation of the Spanish Short Self-Regulation Questionnaire (SSSRQ) through Rasch analysis. *Frontiers in Psychology*, 8(MAR), 276. doi: 10.3389/fpsyg.2017.00276/BIBTEX
- Ghielen, S. T. S., van Woerkom, M., and Meyers, M. C. (2018). Promoting positive outcomes through strengths interventions: a literature review. *J. Posit. Psychol.* 13, 573–585. doi: 10.1080/17439760.2017.1365164
- Goodman, F. R., Disabato, D. J., Kashdan, T. B., and Machell, K. A. (2017). Personality strengths as resilience: a one-year multiwave study. *J. Pers.* 85, 423–434. doi: 10.1111/jopy.12250
- Harms, P. D., Brady, L., Wood, D., and Silard, A. (2018). “Resilience and well-being,” in *Handbook of Well-Being*, eds E. Diener, S. Oishi, and L. Tay (Salt Lake City, UT: DEF Publishers), 1–12.
- Hausler, M., Strecker, C., Huber, A., Brenner, M., Höge, T., and Höfer, S. (2017). Distinguishing relational aspects of character strengths with subjective and psychological well-being. *Front. Psychol.* 8:1159. doi: 10.3389/fpsyg.2017.01159
- He, F., Cao, R., Feng, Z., Guan, H., and Peng, J. (2013). The impacts of dispositional optimism and psychological resilience on the subjective well-being of burn patients: a structural equation modelling analysis. *PLoS One* 8:e82939. doi: 10.1371/journal.pone.0082939
- Heintz, S., and Ruch, W. (2020). Character strengths and job satisfaction: differential relationships across occupational groups and adulthood. *Appl. Res. Qual. Life* 15, 503–527. doi: 10.1007/s11482-018-9691-3
- Hendriks, T., Schotanus-Dijkstra, M., Hassankhan, A., de Jong, J., and Bohlmeijer, E. (2019). The efficacy of multi-component positive psychology interventions: a systematic review and meta-analysis of randomized controlled trials. *J. Happiness Stud.* 21, 357–390. doi: 10.1007/s10902-019-00082-1
- Ho, R. (2006). *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*. Boca Raton, FL: CRC Press.
- Huppert, F. A. (2009a). A new approach to reducing disorder and improving well-being. *Perspect. Psychol. Sci.* 4, 108–111. doi: 10.1111/j.1745-6924.2009.01100.x
- Huppert, F. A. (2009b). Psychological well-being: evidence regarding its causes and consequences. *Appl. Psychol. Health Well Being* 1, 137–164. doi: 10.1111/j.1758-0854.2009.01008.x
- Huppert, F. A., and So, T. T. (2013). Flourishing across Europe: application of a new conceptual framework for defining well-being. *Soc. Indic. Res.* 110, 837–861. doi: 10.1007/s11205-011-9966-7
- Ivtzan, I., Lomas, T., Hefferon, K., and Worth, P. (2016). *Second Wave Positive Psychology: Embracing the Dark Side of Life*. London: Routledge.
- Joseph, S., and Linley, P. A. (2005). Positive adjustment to threatening events: an organismic valuing theory of growth through adversity. *Rev. Gen. Psychol.* 9, 262–280. doi: 10.1037/1089-2680.9.3.262
- Kachel, T., Huber, A., Strecker, C., Höge, T., and Höfer, S. (2020). Development of cynicism in medical students: exploring the role of signature character strengths and well-being. *Front. Psychol.* 11:328. doi: 10.3389/fpsyg.2020.00328
- Karris-Bachik, M. A., Carey, G., and Craighead, W. E. (2020). VIA character strengths among US college students and their associations with happiness, well-being, resiliency, academic success and psychopathology. *J. Posit. Psychol.* 16, 1–14. doi: 10.1080/17439760.2020.1752785
- Keith, T. Z. (2006). *Multiple Regression and Beyond*. Boston, MA: Pearson Education, Inc.

- Kenny, D. (2012). *Effect Size of the Indirect Effect*. Available online at: <http://davidakenny.net/cm/mediate.htm> (accessed March 18, 2022).
- Keyes, C. L. M. (2002). The mental health continuum: from languishing to flourishing in life. *J. Health Soc. Behav.* 43, 207–222. doi: 10.2307/3090197
- Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing: a complementary strategy for improving National Mental Health. *Am. Psychol.* 62, 95–108. doi: 10.1037/0003-066X.62.2.95
- Kim, J., Han, A., and Park, S.-H. (2016). Stress-related growth experience: listening to Korean adolescents who have visual impairments. *J. Constr. Psychol.* 30, 427–438. doi: 10.1080/10720537.2016.1246991
- Kline, R. (2005). *Principles and Practices of Structural Equation Modeling*, 2nd Edn. New York, NY: Guilford Press.
- Koltai, J., and Stuckler, D. (2020). Recession hardships, personal control, and the amplification of psychological distress: differential responses to cumulative stress exposure during the US Great Recession. *SSM Popul. Health* 10:100521. doi: 10.1016/j.ssmph.2019.100521
- Lavy, S. (2019). A review of character strengths interventions in twenty first century schools: their importance and how they can be fostered. *Appl. Res. Qual. Life* 15, 573–596. doi: 10.1007/s11482-018-9700-6
- Lee, M. T., Bialowolski, P., Weziak-Bialowolska, D., Mooney, K. D., Lerner, P. J., McNeely, E., et al. (2021). Self-assessed importance of domains of flourishing: demographics and correlations with well-being. *J. Posit. Psychol.* 16, 1–8. doi: 10.1080/17439760.2020.1716050
- Liu, J. J. W., Ein, N., Gervasio, J., Battatton, M., Reed, M., and Vickers, K. (2020). Comprehensive meta-analysis of resilience interventions. *Clin. Psychol. Rev.* 82:101919. doi: 10.1016/j.cpr.2020.101919
- Lohr, S. L. (1999). *Sampling: Design and Analysis*. Pacific Grove, CA: Brooks/Cole Publishing.
- Lomas, T. (2016). Flourishing as a dialectical balance: emerging insights from second wave positive psychology. *Palgrave Commun.* 2:16018.
- Manzano-García, G., and Ayala-Calvo, J. C. (2013). New Perspectives: towards an Integration of the concept "burnout" and its explanatory models. *Anal. Psicol.* 29, 800–809. doi: 10.6018/analesps.29.3.145241
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530. doi: 10.1093/biomet/57.3.519
- Martínez-Martí, M. L., and Ruch, W. (2017). Character strengths predict resilience over and above positive affect, self-efficacy, optimism, social support, self-esteem, and life satisfaction. *J. Posit. Psychol.* 12, 110–119. doi: 10.1080/17439760.2016.1163403
- Miranda, J. O., and Cruz, R. N. C. (2020). Resilience mediates the relationship between optimism and well-being among Filipino university students. *Curr. Psychol.* 39, 1–10. doi: 10.1007/s12144-020-00806-0
- Niemiec, R. M. (2020). Six functions of character strengths for thriving at times of adversity and opportunity: a theoretical perspective. *Appl. Res. Qual. Life* 15, 551–572. doi: 10.1007/s11482-018-9692-2
- Park, N., and Peterson, C. (2009). Character strengths: research and practice. *J. Coll. Character* 10, 1–9. doi: 10.2202/1940-1639.1042
- Park, N., Peterson, C., and Seligman, M. E. P. (2004). Strengths of character and well-being. *J. Soc. Clin. Psychol.* 23, 603–619. doi: 10.1521/jscp.23.5.603.50748
- Park, N., Peterson, C., and Seligman, M. E. (2006). Character strengths in fifty-four nations and the fifty US states. *J. Posit. Psychol.* 1, 118–129. doi: 10.1080/17439760600619567
- Patnaik, G. (2021). *Positive Psychology for Improving Mental Health & Well-Being*. Chennai: Notion Press.
- Peterson, C., and Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Classification and Handbook*. New York, NY: Oxford University Press.
- Pozo-Muñoz, C., Garzón-Umerenkova, A., Bretónes-Nieto, B., and Charry-Poveda, C. (2016). Psychometric properties and dimensionality of the 'Flourishing Scale' in Spanish-speaking population. *Electron. J. Res. Educ. Psychol.* 14, 175–192. doi: 10.14204/ejrep.38.15044
- Proyer, R. T., Gander, F., Wellenzohn, S., and Ruch, W. (2013). What good are character strengths beyond subjective well-being? The contribution of the good character on self-reported health-oriented behavior, physical fitness, and the subjective health status. *J. Posit. Psychol.* 8, 222–232. doi: 10.1080/17439760.2013.777767
- Rapport, L. J., Wong, C. G., and Hanks, R. A. (2020). Resilience and well-being after traumatic brain injury. *Disabil. Rehabil.* 42, 2049–2055. doi: 10.1080/09638288.2018.1552327
- Reeves, R. V., Venator, J., and Howard, K. (2014). *The Character Factor: Measures and Impact of Drive and Prudence*. Center on Children & Families. Available online at: https://www.brookings.edu/wp-content/uploads/2016/07/22_character_factor_opportunity_reeves.pdf (accessed March 18, 2022).
- Ruch, W., and Hofmann, J. (2017). "Fostering humour" in *Positive Psychology Interventions in Practice*, ed. C. Proctor (Cham: Springer International Publishing), 65–80. doi: 10.1007/978-3-319-51787-2_5
- Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68
- Ryan, R. M., and Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and Eudaimonic well-being. *Annu. Rev. Psychol.* 52, 141–166. doi: 10.1146/annurev.psych.52.1.141
- Ryff, C. D. (2014). Psychological well-being revisited: advances in science and practice. *Psychother. Psychosom.* 83:10. doi: 10.1159/000353263
- Ryff, C. D. (2016). "Eudaimonic well-being and education: probing the connections," in *Well-Being and Higher Education. A Strategy of Chance and the Realization of Education's Grater Purposes*, ed. D. W. Harward (Washington, DC: Bringing Theory to Practice), 37–48.
- Ryff, C. D., and Singer, B. H. (2006). Know thyself and become what you are: a eudaimonic approach to psychological well-being. *J. Happiness Stud.* 9, 13–39. doi: 10.1007/S10902-006-9019-0
- Schotanus-Dijkstra, M., Pieterse, M. E., Drossaert, C. H., Westerhof, G. J., de Graaf, R., Ten Have, M., et al. (2016). What factors are associated with flourishing? Results from a large representative national sample. *J. Happiness Stud.* 17, 1351–1370. doi: 10.1007/s10902-020-00325-6
- Schreiner, L. A. (2018). "Applying principles of positive psychology to student success and well-being: enabling all students to thrive," in *Positive Psychology, established and emerging Issues*, ed. D. S. Dunn (London: Routledge), 197–219. doi: 10.4324/9781315106304-12
- Schutte, N. S., and Malouff, J. M. (2019). The impact of signature character strengths interventions: a meta-analysis. *J. Happiness Stud.* 20, 1179–1196. doi: 10.1007/S10902-018-9990-2
- Seale, J. P., Hadfield, K., Seale, D. M., Pande, Y., Lewis, T. M., and Harris, S. K. (2020). The effects of a spiritually-based character strengths training curriculum on alcohol use among Zambian youth: a randomized controlled trial. *Alcohol. Clin. Exp. Res.* 44:144.
- Seligman, M. (2018). PERMA and the building blocks of well-being. *J. Posit. Psychol.* 13, 333–335. doi: 10.1080/17439760.2018.1437466
- 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
- Tabachnick, B. G., Fidell, L. S., and Ullman, J. B. (2007). *Using Multivariate Statistics*, Vol. 5. Boston, MA: Pearson, 481–498.
- 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
- VanderWeele, T. J. (2017a). Religious communities and human flourishing. *Curr. Dir. Psychol. Sci.* 26, 476–481. doi: 10.1177/0963721417721526
- VanderWeele, T. J. (2017b). "Religion and health: a synthesis," in *Spirituality and Religion within the Culture of Medicine: From Evidence to Practice*, eds M. J. Balboni and J. R. Petzet (New York, NY: Oxford University Press), 357–401.
- VanderWeele, T. J. (2020). Activities for flourishing: an evidence-based guide. *J. Posit. Sch. Psychol.* 4, 79–91. doi: 10.24095/hpcdp.40.10.01
- Vargas, L., and González-Torres, M. C. (2009). Revitalization of character education in the current educational psychology arena: contributions from the sciences of Prevention and Positive Psychology. *Electron. J. Res. Educ. Psychol. (EJREP)* 7, 1379–1418. doi: 10.25115/ejrep.v7i19.1339
- VIA Institute on Character (2020). *The VIA Character Survey*. Available online at: <https://www.viacharacter.org/survey/account/Register> (accessed March 18, 2022).
- Villacis, J. L., de la Fuente, J., and Naval, C. (2021). Good character at college: the combined role of second-order character strength factors and Phronesis motivation in undergraduate academic outcomes. *Int. J. Environ. Res. Public Health* 18:8263. doi: 10.3390/ijerph18188263
- Wagner, L., Gander, F., Proyer, R. T., and Ruch, W. (2020). Character strengths and PERMA: investigating the relationships of character strengths with a

- multidimensional framework of well-being. *Appl. Res. Qual. Life* 15, 307–328. doi: 10.1007/s11482-018-9695-z
- Weston, R., and Gore, P. A. A. (2006). Brief guide to structural equation modeling. *Couns. Psychol.* 34, 719–751. doi: 10.1177/0011000006286345
- Wu, Y., Sang, Z., Zhang, X.-C., and Margraf, J. (2020). The relationship between resilience and mental health in Chinese college students: a longitudinal cross-lagged analysis. *Front. Psychol.* 11:108. doi: 10.3389/fpsyg.2020.00108
- Yildirim, M., and Belen, H. (2019). The role of resilience in the relationships between externality of happiness and subjective well-being and flourishing: a structural equation model approach. *J. Posit. Psychol. Wellbeing* 3, 62–76.

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 de la Fuente, Urien, Luis, González-Torres, Artuch-Garde and Balaguer. 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.



Cyberbullying and Cyberhate as Two Interlinked Instances of Cyber-Aggression in Adolescence: A Systematic Review

Giovanni Fulantelli¹, Davide Taibi^{1*}, Lidia Scifo¹, Veronica Schwarze² and Sabrina C. Eimler²

¹ Institute for Educational Technology, National Research Council of Italy, Palermo, Italy, ² Institute of Computer Science, Institute of Positive Computing, University of Applied Sciences Ruhr West, Bottrop, Germany

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Caterina Fiorilli,
Libera Università Maria SS. Assunta
University, Italy
Giorgio Falgares,
University of Palermo, Italy

*Correspondence:

Davide Taibi
davide.taibi@itd.cnr.it

Specialty section:

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

Received: 31 March 2022

Accepted: 29 April 2022

Published: 27 May 2022

Citation:

Fulantelli G, Taibi D, Scifo L,
Schwarze V and Eimler SC (2022)
Cyberbullying and Cyberhate as Two
Interlinked Instances of
Cyber-Aggression in Adolescence: A
Systematic Review.
Front. Psychol. 13:909299.
doi: 10.3389/fpsyg.2022.909299

In this paper we present the results of a systematic review aimed at investigating what the literature reports on cyberbullying and cyberhate, whether and to what extent the connection between the two phenomena is made explicit, and whether it is possible to identify overlapping factors in the description of the phenomena. Specifically, for each of the 24 selected papers, we have identified the predictors of cyberbullying behaviors and the consequences of cyberbullying acts on the victims; the same analysis has been carried out with reference to cyberhate. Then, by comparing what emerged from the literature on cyberbullying with what emerged from the literature on cyberhate, we verify to what extent the two phenomena overlap in terms of predictors and consequences. Results show that the cyberhate issue related to adolescents is less investigated than cyberbullying, and most of the papers focusing on one of them do not refer to the other. Nevertheless, by comparing the predictors and outcomes of cyberbullying and cyberhate as reported in the literature, an overlap between the two concepts emerges, with reference to: the parent-child relationship to reduce the risk of cyber-aggression; the link between sexuality and cyber-attacks; the protective role of the families and of good quality friendship relationships; the impact of cyberbullying and cyberhate on adolescents' individuals' well-being and emotions; meaningful analogies between the coping strategies put in practice by victims of cyberbullying and cyberhate. We argue that the results of this review can stimulate a holistic approach for future studies on cyberbullying and cyberhate where the two phenomena are analyzed as two interlinked instances of cyber-aggression. Similarly, prevention and intervention programs on a responsible and safe use of social media should refer to both cyberbullying and cyberhate issues, as they share many predictors as well as consequences on adolescents' wellbeing, thus making it diminishing to afford them separately.

Systematic Review Registration: <http://www.crd.york.ac.uk/PROSPERO>, identifier: CRD42021239461.

Keywords: cyberbullying, cyberhate, cyber-aggression, adolescents, social media, social network sites

INTRODUCTION

Social media have become online environments in which face-to-face activities of everyday life are transferred in the network mediated world with a wider audience (potentially millions of users) and no time constraints (they are open 24 h a day). According to the Digital 2020 July Global Statshot report¹, July 2020 can be considered a milestone in the history of the internet, since for the first time more than half of the world's total population was using social media, with a total number of 3.96 billion active social media users (Kemp, 2020). Due to the coronavirus pandemic lockdowns, this number increased up to 4.33 billion active social media users in April 2021 (55.1% of world population, with an annual increase close to 14%) (Kemp, 2021a)².

The share of adolescents contributing to these numbers is impressive: 90% of US teens aged 13–17 years use social media (AACAP, 2018)³; a similar percentage concerns Europe, where 87% of people aged 16 to 24 years use social networks, ranging from 79% in Italy up to 97% in Denmark (Eurostat, 2020⁴; research from GWI shows that 99.6% of South-East Asian internet users aged 16–24 years use social media (Kemp, 2021b).

According to these statistics, adolescents are among the most frequent users of social media. Following Shapiro and Margolin (2014), the main motivations for adolescents using social media are “to stay in touch with friends, make plans, get to know people better, and present oneself to others”.

Despite these perceived or real benefits, social media can become a place in which antisocial behaviors such as bullying, harassment and hate speech, proliferate and evolve leveraging the peculiarity of the online world (ElSherief et al., 2018). As a consequence, participation in social media exposes adolescents more to the risks associated with their use. In particular, several studies have pointed out the relationship between social media use and cyber-violence, broadly defined as violent acts perpetrated through the social media (Peterson and Densley, 2017; Backe et al., 2018; Nagle, 2018), and how adolescents can become victims or perpetrators of aggressive behaviors (Chisholm, 2006; O’Keeffe et al., 2011; Peterson and Densley, 2017). Furthermore, some authors have highlighted the difficulties in classifying and defining the spectrum and diversity of online violent behaviors, their specificity compared to similar offline behaviors, and the limits that result from a lack of clear definitions of online violent behaviors (Grigg, 2010; Pyzalski, 2012; Peterson and Densley, 2017).

It is precisely from the difficulties in defining the concept of cyberbullying that Grigg (2010) comes to the conclusion that it is necessary to move to a concept at a higher level of abstraction that includes all the online behaviors characterized by a high level of aggression, thus introducing the concept of cyber-aggression: “The study examined current definitions and concepts of

cyberbullying and how these differ in its findings; and considered different ways to foster positive online behavior for the context of practitioners. The concept of cyber-aggression is used to describe a wide range of behaviors other than cyberbullying. The findings indicate that there is a need to include a broader definition in line with the current trend of a range of behaviors that are common with internet and mobile phone usage” (p. 143). Following Grigg’s idea of cyber-aggression, Corcoran et al. (2015) argue that, in order to overcome the problems related to the variations across definitions of cyberbullying, it is necessary to consider the broader issue of cyber-aggression.

In addition to cyberbullying, cyberhate is another important example of online aggressive behavior which is more and more involving young victims and perpetrators. This is related to two main factors: firstly, the amount of online hate material (such as hateful messages and, more in general, content that harm the reputation of or instigate violence against groups or individual as member of groups) is rapidly increasing, and the risk for adolescents to be exposed to hateful online material is increasing accordingly (Hawdon et al., 2019; Harriman et al., 2020); secondly, adolescents are becoming one of the preferred target for online recruitment by organized hate groups and individuals (Smith, 2009; Costello et al., 2020). Similar to cyberbullying, several authors consider cyberhate as a subset of cyber-aggression (Mardianto et al., 2019; Tennakoon, 2021; Bedrosova et al., 2022).

As mentioned before, adolescents are particularly vulnerable to cyber-aggression, not only because of the time spent online: US statistics in 2018 showed that nearly all teens aged 13–17 years (95%) have access to a smartphone and 45% of them reported that they were online “almost constantly” (Anderson and Jiang, 2018); adolescents are at high risk also because most of them have fewer psychological tools than the majority of adults to defend themselves against cyber-aggression, such as resilience, competence, literacy, critical thinking, experiences.

Cyberbullying and cyberhate play a dramatic role in the relationships between adolescents’ well-being and use of social media. In fact, cyberbullying is the most frequent form of cyber-aggression involving adolescents, while cyberhate is the form of cyber-aggression that is spreading most rapidly among young people. Furthermore, the two phenomena are not totally distinct, rather there are some overlaps between them (Goerzig et al., 2019; Bedrosova et al., 2022). However, to the best of our knowledge, there is no approach addressing cyberbullying and cyberhate as two distinct but interlinked phenomena which makes it hard to evaluate them in empirical research settings. As a consequence, analyzing the two phenomena not in a separate manner, but on the contrary assuming that there may be important links between them, can help to develop models for the identification of predictor variables, to broaden the assessments of the possible impacts they have on the lives of adolescents, and to put in practice more effective and efficient prevention strategies. This approach can offer a concrete model which would be of interest to academia to explain how theoretical models can help to derive practical interventions to limit the spread of these phenomena. In the long term, the review can support practitioners in the school contexts in developing intervention measures aimed at avoiding toxic dynamics on the internet.

¹<https://datareportal.com/reports/digital-2020-july-global-statshot>

²<https://datareportal.com/reports/digital-2021-april-global-statshot>

³https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Social-Media-and-Teens-100.aspx

⁴<https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210630-1>

Accordingly, in this paper we present the results of a systematic review aimed at investigating what the literature reports on cyberbullying and cyberhate, whether and to what extent the connection between the two phenomena is made explicit, and whether it is possible to identify what exactly the overlapping factors are. Specifically, for each of the analyzed papers, we have identified the predictors of cyberbullying behaviors and the consequences of cyberbullying acts on the victims; the same analysis has been carried out with reference to cyberhate. Then, by comparing what emerged from the literature on cyberbullying with what emerged from the literature on cyberhate, we verify to what extent the two phenomena—as reported in the literature—overlap in terms of predictors and consequences.

Cyberbullying and Cyberhate as Two Interconnected Phenomena

In order to reflect on the relationship between cyberbullying and cyberhate in terms of differences and similarities, we focus on the distinguishing characteristics of each of them, being aware that the definition of the two concepts is not an objective of this review. In both cases, we focus on the comparisons between these two forms of cyber-aggression and their equivalent forms in face-to-face contexts: bullying and hate speech.

Cyberbullying does not simply refer to the transition from 'traditional' bullying in face-to-face contexts to bullying in online contexts, where, according to the Centers for Disease Control and Prevention, bullying is defined as “*any unwanted aggressive behavior(s) by another youth or group of youths who are not siblings or current dating partners that involves an observed or perceived power imbalance and is repeated multiple times or is highly likely to be repeated.*”⁵ In fact, the characteristics of social media lead to a re-interpretation of the concepts of aggression, repetition and the imbalance of power (Whittaker and Kowalski, 2015). Firstly, the characteristics of physical aggression related to vocal tone and facial expression assume a different value in the virtual world, in which other shapes of aggression come into the scene through hate speech or online harassment, humiliation or exclusion. Concerning repetition, in an online environment a harassing statement can be potentially viewed, “liked” and shared by other users multiple times, therefore the repetition of the act over time is not more crucial, only one shared content humiliating a victim could have a destructive effect on the victim's self-esteem. Finally, power imbalance is difficult to detect in a virtual environment in which power can be expressed in a multitude of ways. For example, users with a high level of digital knowledge can conduct cyberattacks by using sophisticated tools, so that power imbalance might also reflect differences in technological expertise (Whittaker and Kowalski, 2015).

Other characteristics of the virtual environment affect the proliferation of cyberbullying and differentiate cyberbullying from traditional bullying. In the virtual environment cyberbullies can use anonymous accounts to attack their victims. There is another relevant difference between cyberbullying and bullying

as stated by Englander (2017) who state that cyberbullying is connected to the widespread use of digital devices, thus leading cyberbullying to happen mainly outside of school, whereas traditional bullying most often happens in school.

Moving to the cyberhate concept, there are not universally accepted definitions of hate speech and cyberhate (MacAvaney et al., 2019). Specifically, to hate speech, differences amongst the definitions concern several aspects. Firstly, the authors and the spreaders of the hate messages can be individuals, organized groups or a combination of individuals and organized groups (Blaya and Audrin, 2019). Then, the target or victim is one of the most variable concepts in the definitions of hate speech, and it strongly reflects the differences in the contexts of use and in the historical period of definition. Some of the definitions consider only specific groups, such as the one proposed by the Council of Europe in the Additional Protocol to the Convention on Cybercriminality (Council of Europe, 2003) that states that individuals or communities become target of attacks because of their “race, color, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors”; the constraints posed to the potential target groups of hate speech is due to the origin of this document, conceived as a normative instrument to contrast racism and xenophobia. On the other end of the spectrum of definitions, the Center for Equal Opportunities and Opposition to Racism in Brussels lists sex, sexual orientation or political or religious beliefs, in addition to skin color, supposed race, ethnic origin, as reasons to unleash the haters.

A further distinction concerns the purpose of haters. The Council of Europe states that hate speech aims at advocating, promoting or inciting hatred, discrimination or violence (Council of Europe, 2003). Blaya and Audrin (2019) clarify that the purpose of haters is to attract new members toward their ideals, thus building and strengthening group identity to counter and reject others' collective identity. The mechanisms of propaganda, insulting and discrimination are therefore central to the hate speech definitions (Council of Europe, 2003; Anti-Defamation League, 2010; Blaya and Audrin, 2019). One of the definition that better resumes and mediates the different perspectives to hate speech is the one proposed in 1997 by the Council of Europe, which refer to hate speech as “all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility toward minorities, migrants and people of immigrant origin” (Council of Europe, 1997).

Moving to the differences between hate speech and online hate speech or cyberhate, the most evident difference is the media used for disseminating hate content and messages. The Council of Europe (2003) provides a comprehensive view of hate speech material, which includes “any written material, any image or any other representation of ideas or theories”. Drawing on this definition, Blaya and Audrin (2019) adopt the term “cyberhate” to refer to all hateful online forms of expression (text, images, videos, pictures, graphic representations) to generate hatred against people and communities. Cyberhate is based on the spreading of hateful material through electronic communication

⁵<https://www.cdc.gov/violenceprevention/youthviolence/bullyingresearch/fastfact.html>

technologies (e.g., websites, social media, blogs, online games, instant messaging services, e-mail). As for cyberbullying, the use of Internet-based communication media amplifies the effect of cyberhate, and exacerbates the negative consequences of hate speech.

Differences between cyberbullying and cyberhate are a direct consequence of the definitions provided so far. Firstly, as already mentioned, the final purpose of cyberhate is to promote or incite hatred, discrimination or violence against a community or group in order to disaggregate social cohesion and mine democracy; instead, the final aim of a cyberbullying is to harm an individual.

Bullying is known with its repetitive act to the same individual, unlike hate speech which is more general and not necessarily intended to hurt a specific individual (Al-Hassan and Al-Dossari, 2019). As Chetty and Alathur (2018) claim, hate speech may harm the victims directly or indirectly. Direct hate speech is similar to bullying, since the victims are injured immediately by hate speech content. However, in an indirect hate speech, the harm perpetrated by the original hater is only a part of the final goal, since the hater incites other people or organized groups to attack the victims, and a delayed harm is perpetrated by the latter, not by an original actor. In a typical racist hate speech scenario, hateful content on racism in (real or online) public settings might motivate other people to initiate harassment, intimidation, violence against ethnical minorities (Seglow, 2016). This introduces a second difference: one of the aims of the cyberhaters is to involve as many people as possible as active agents in the attack; this is not a priority for the perpetrator of cyberbullying, who is normally an individual or a small group of peers. Consequently, the online services used by the two categories of haters are different, even if some overlaps exist.

Strictly related to the previous one, another difference is that cyberhate targets communities more than individuals, while cyberbullying victims are individuals, usually young individuals in the setting of a particular community, like a school (López and López, 2017). Following Blaya and Audrin (2019), cyberhate can also harm individuals and affect them emotionally, but the main negative consequences are on whole communities. Another difference is that the perpetrator of cyberbullying usually personally knows his/her victim, which is often unknown to most of the cyberhaters.

Finally, the idea of victimization changes accordingly: specifically to cyberhate, Machackova et al. (2020) distinguish between cyberhate exposure and cyberhate victimization, and they define the former as “the experience of encountering hateful content online but not necessarily feeling victimized by it”. This distinction is not necessary in cyberbullying, since it is always possible to identify a victim.

Despite the several differences between the concepts of cyberbullying and cyberhate, they share important characteristics that lead to a partial overlap between the two concepts and promote the study of similar solutions. As mentioned before, in both cases the use of Internet-based technologies amplifies the consequences of the attacks, both in terms of geographical space and persistence over time. Secondly, the lack of face-to-face contact between the perpetrator and the victim, both in cyberbullying and in cyberhate, makes online forms of expression

(text, images, videos, pictures, graphic representations) the common language used by the perpetrators. Then, school-aged children, adolescents and young people are particularly exposed to cyberbullying and cyberhate, both as victims and perpetrators: Li et al. (2021) have illustrated that cyberbullying perpetration among school-aged children is a transnational phenomenon; likewise, organized hate groups specifically target adolescents as new recruits (Lee and Leets, 2002; Gerstenfeld et al., 2003), so that they are particularly vulnerable to online media activist groups due to their presence in social media, and the particular stage of their development who makes them sensitive to feelings of personal commitment, of social utility and of belonging they can provide (Atran and Ginges, 2015); finally, the concept of victimization is central to both, cyberbullying and cyberhate, in particular by taking into account that a large number of victims and perpetrators are adolescents.

In addition to these general factors that characterize the nature of the two phenomena, the aim of this review is, as anticipated in the introduction, to analyze in more detail the individual factors that—based on the results of the empirical studies presented in the selected papers—allow us to understand the level of overlap between cyberbullying and cyberhate. This will make it possible to better target interventions aimed at preventing the phenomena of cyber-aggression committed by or directed against adolescents.

MATERIALS AND METHODS

The PRISMA method was followed for the review methodology and data extraction (Liberati et al., 2009; Moher et al., 2009; Page et al., 2021). A protocol for this review was registered on PROSPERO in April 2021 (PROSPERO registration number CRD42021239461. The registration number is available at <http://www.crd.york.ac.uk/PROSPERO>).

Participants and Procedure

To identify the literature, the following databases were searched: PsycInfo, Scopus, PubMed, APA PsycArticles, EBSCO. The search for electronic literature databases was dated from January 2021 to February 2021. In each database, the following terms were searched: social media, adolescents, boys, girls, young adults, teenagers, cyberbullying, cybermobbing, online hate speech, cyberhate, cyber victimization. The search strategy was carried out combining these keywords with boolean operators as AND, OR, NOT. Of the selected empirical studies, the chronological age of research participants between 10 and 24 years were considered. Moreover, the broad age range of research participants was a forced choice because several studies included both adolescents and young adults. The selection of the studies was a process of evaluation of synonyms and related keywords, as the scientific literature concerning the topic of the review is vast and articulated.

In order to ensure that no relevant studies were missed, additional studies were identified by hand-searching the reference lists of reviews and research papers. Missing papers were requested from study authors by email.

The papers/records included in the review and analyzed are 24.

Inclusion and Exclusion Criteria

All the records were independently screened by four review authors to identify studies that potentially met the inclusion and exclusion criteria as outlined below.

The following inclusion criteria were adopted:

- studies described different types of cyberbullying and cyberhate in social media
- quantitative study and empirical papers published between 2000 and 2021
- studies were published in the English language
- empirical studies, experimental and quasi-experimental design
- peer-reviewed studies
- people aged between the ages of 10 and 24.

In the exclusion criteria, duplicates and irrelevant records have been eliminated.

In particular, the following exclusion criteria were adopted:

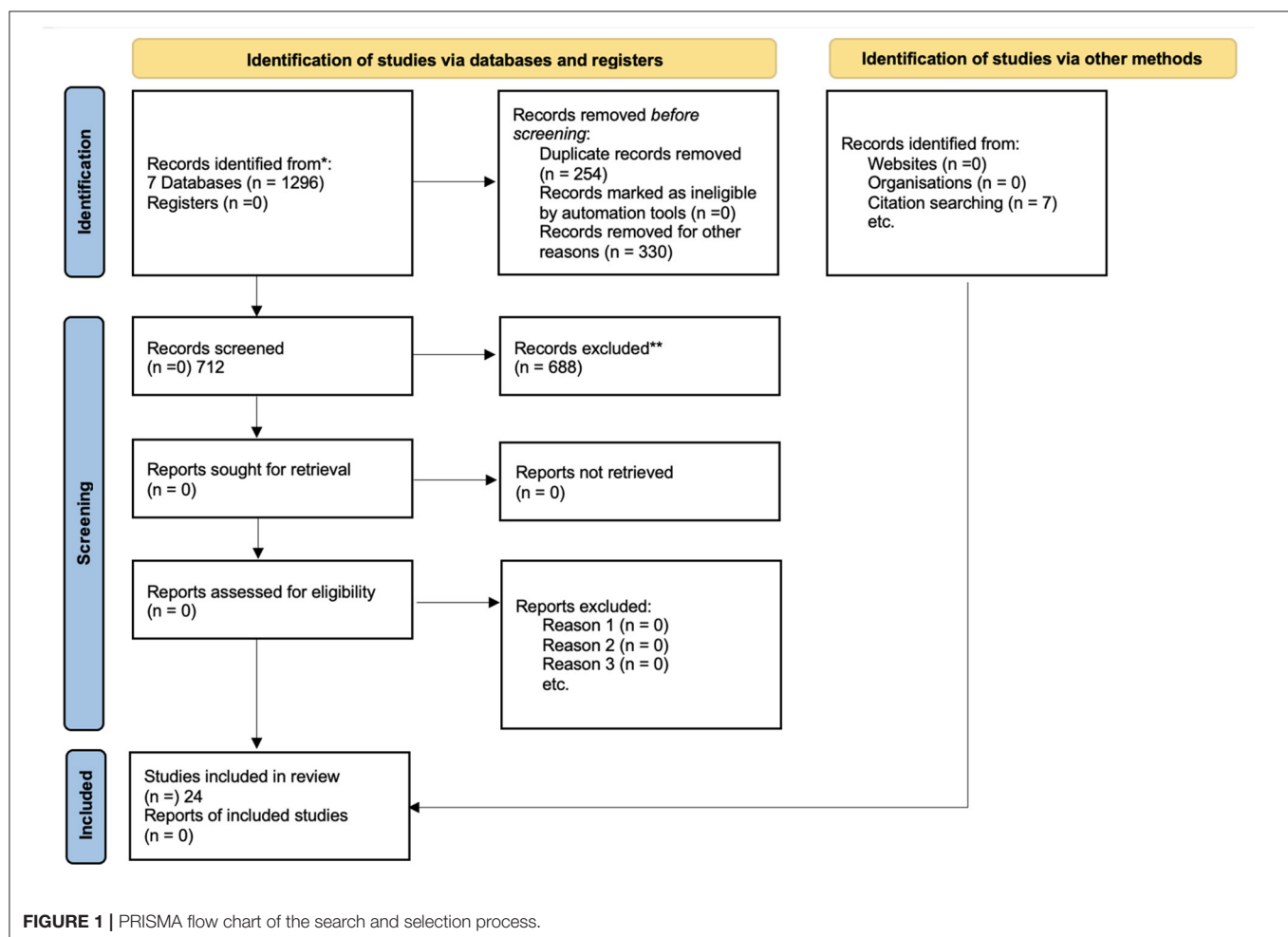
- qualitative studies

- studies were published before 2000
- studies were not published in the English language
- cross-sectional design, single case
- studies were not peer-reviewed
- gray literature, e.g., dissertations, conference abstracts, research reports, chapter(s) from a book, Ph.D. theses, reports on ID guidelines.

Data Analysis

The initial database search identified a total of 1296 records, after a careful selection, according to the PRISMA checklist and the inclusion and exclusion criteria, 24 papers were analyzed for the data extraction. The flow diagram following the models of Page et al. (2021) is included as **Figure 1**.

Papers were firstly reviewed through the title and abstract to determine if they could be included or excluded. The final papers were organized for the data extraction according to the following variables: digital object ID, article title, abstract, journal title, journal year, authors, social media as the subject of investigation, which kind of cyberbullying/cyberhate, content of the studies,



method and procedure participants, age participants, sex, stage of life, country, aim of the study, result.

The synthesis of the main variables considered essential to the topic of the systematic review process is reported in **Table 1**.

To systematically report and compare the heterogeneous findings of the articles included, a coding scheme was created, with categories derived from the reported results. Since some papers contained a large number of different worth reporting findings, the results were not exclusively assigned to one category, but are presented in the context of different dimensions. Given that the focus of this review was on cyberbullying and cyberhate as two interlinked instances of cyber-aggression, both phenomena contributed to the main distinction of the results, with 19 papers referring to cyberbullying and 5 papers referring to cyberhate. Next, the results were categorized according to whether they related to the predictors or the impact of cyberbullying or cyberhate. Thus, a 2×2 matrix was created, which served as a grid for the analysis of the results. The grid was further divided into subcategories as follows: as predictors, we considered (1) socio-demographic variables (e.g., age/school grade level, gender, race/ethnic background), (2) individual and contextual factors (e.g., empathy, sexuality, appearance/overweight, school performance, relationships with friends and family) and (3) the overlap between traditional and cyberbullying as well as the overlap of cyberhate with other forms of aggressive behavior. Similarly, the effects of both phenomena were contrasted and considered categories such as effects on health and well-being (e.g., psychological distress, depressive symptoms, somatic symptoms, suicidal) and coping strategies.

RESULTS

Predictors of Cyberbullying Socio-Demographic Variables

Numerous studies investigated the potential impact of various socio-demographic variables on cyberbullying behaviors. Age seems to play a substantial role in the prevalence of these behaviors. For example, Ybarra et al. (2011) found evidence for increasing age being predictive for exposure to and experience of online harassment, surveying 10 to 15 year old over three years. In this sense, Ševčíková and Šmahel (2009) found that of the older adolescents (16–19 years) 14.1% were in the role of the target, while of the younger adolescents (12–15 years) 7.6% experienced the target perspective. However, the situation was reversed for aggressors, as the higher proportion of aggressors was among 12–15 year old's (1.8%) rather than 16-to-19 years old's (1.2%). Mishna et al. (2010) reported an increased likelihood for older girls (grades 10 and 11) being exposed to cyberbullying compared to older boys, although this difference was not seen between girls and boys in lower grades (grades 6 and 7). Contrarily, while Ortega-Ruiz et al. (2009) indicate a significant peak in victimization at the age of 14, they also indicate that victimization decreased significantly between the ages of 12 and 17. Regarding cyberbullying, Schneider et al. (2012) report similar findings, with this phenomenon decreasing slightly from 9th grade to 12th grade (from 17.2 to 13.4%). Likewise, Waasdorp et al. (2018) observed that compared to middle school students, high school

students were less likely to be affected by the various forms of victimization with the exception of cyberbullying.

Age (school grade level) also seemed to moderate the association between cyberbullying victimization and students' engagement. While a positive association between cyberbullying victimization and emotional engagement was stronger for high school students than for middle school students, a negative association between cyberbullying victimization and cognitive-behavioral engagement was stronger for middle school students than for high school students (Yang et al., 2020).

Gender was also hypothesized to play a role in frequency and type of cyberbullying behaviors students encountered. Grading et al. (2009) observed that 8% of male students reported having sent mean text messages, e-mails, videos or photographs, while only 3% of female students reported doing so. Similarly, Baldry et al. (2019) observed that boys tended to be only bullies and bullies/victims, while girls tended to be uninvolved and only victims. These results are in line with those found by Ortega-Ruiz et al. (2009) who found that more females reported being victims of cyberbullying, both *via* mobile phone (6.3% females vs. 2.4% males) and *via* the Internet (9.1% females vs. 6% males) and Schneider et al. (2012) who found that the victims of cyberbullying are more often girls than boys (11.1% vs. 7.6%). Contrarily, Low and Espelage (2013) found that female middle school students had higher levels of cyberbullying, with the extent decreasing over time. That girls are not only victims of cyberbullying, but also take on the role of perpetrators, was also shown by Mishna et al. (2010). Here, however, the form of cyberbullying was a relevant factor. According to this, boys were more likely to be victims or perpetrators of direct bullying (e.g., threatening), while girls were more likely to be victims or perpetrators of indirect bullying (e.g., spreading rumors). In this study 3% of participants believed they were bullied because of their gender and 1% indicated bullying others because of their target's gender. In addition, Mehari and Farrell (2018) observed that aggressive behaviors did not show different patterns of relations according to gender and Longobardi et al. (2020) found that gender had no significant effect on any of the variables in their study.

The ethnic background also seems to play a role in online victimization and harassment. At the first of three measurement time points, for example, Low and Espelage (2013) showed that African American youth reported higher levels of endorsement of cyberbullying compared to White participants. These findings are contrasted by other researchers, who found no significant differences in overall reporting of cyberbullying by race or ethnicity (Schneider et al., 2012). These results are in line with those found by Mishna et al. (2010), in which Canadian students who did not speak English at home were not at a higher risk of being bullied. In addition, no differences in the prevalence of cyberbullying were found when the language spoken at home was considered. However, it was found that students who spoke English at home were more likely to spread rumors online than students who did not speak English at home. In this study 6% of the participants believed they were bullied online because of their race, while 3% of participants reported online bullying because of the target's race. Concerning the role of race, Ybarra et al.

TABLE 1 | Main characteristics of the papers included in the review.

Author/s year	Characteristics of the sample	Subject of investigation	Hate speech, (Cyber)bullying	Social media	Main findings	Paper key word
Mitchell et al. (2011)	2,051 adolescents age range 10–17 years	To examine rates of victimization, and the association between online and offline victimization. In particular, the symptoms of trauma and delinquency among adolescents were analyzed	Sexual victimizations, and psychological and emotional abuse	Not specified	Results show that many victims are at risk because they have very complex previous emotional experiences. Offline victimization experiences are associated with online victimization	Internet; Victimization; Adolescents; Trauma; Delinquency; Life adversity
Bossler et al. (2012)	434 adolescents students in a Kentucky middle and high school	To explore online harassment experiences. In particular, it was examined whether routine computer activities including use of computers increase the risk of victimization	Online harassment, bullying	MySpace, Facebook	Results show that having a social networking site and sharing personal information on these online platforms seemed to make harassment more likely than using tools such as email or instant messaging	Online harassment, routine activities theory, bullying
Ybarra et al. (2011)	1,588 adolescents age range 10–15 years	To examine technology-mediated exposures (e.g., hate sites, death sites) and experiences (e.g., bullying) and how they are associated with psychosocial challenges (e.g., violent behavior, depressive symptomatology)	Cyberbullying, Internet harassment unwanted sexual solicitation, unwanted sexual experiences	Text-Messaging, Violent Web sites, Hate sites, Design-based, War, death, and “terrorism” sites Design-based, Cartoons sites Design-based, Violent x-rated (“adult”) sites	Results show that while youth move online, victimization rates increase. Almost all violent experiences and exposures online are caused by the use intensity and frequency of the Internet and text messaging	Youth, violent media, Internet harassment, unwanted sexual solicitation, unwanted sexual experiences, cyberbullying
Oksanen et al. (2014)	723 Adolescents age range 15–18 years	To analyze data collected from a sample of Finnish Facebook and YouTube adolescent users. This research investigated the extent of exposure to and victimization by online hate material among young social media users	Hate material, victimization	Facebook and YouTube	Results show that exposure to hate material was associated with high online activity, poor attachment to family and physical offline victimization. Online hate material primarily focused on sexual orientation, physical appearance, and ethnicity	Victimization, internet, adolescents, youth, hate material
Pauwels and Schiils (2016)	6,020 Adolescents age range 16–24 years	To apply Social Learning theory to the explanation of political violence, focusing on exposure to extremist content through new social media and Facebook	Youth delinquency, youth delinquency, differential association exposure	Facebook, NSM	Results show that the most violent effects of new social media and Facebook are found for those measures where individuals actively seek out extremist content on the Internet. It is necessary to check background	Nonstate actors, Radicalization, Terrorism / counterterrorism, Violent extremism

(Continued)

TABLE 1 | Continued

Author/s year	Characteristics of the sample	Subject of investigation	Hate speech, (Cyber)bullying	Social media	Main findings	Paper key word
Räsänen et al. (2016)	723 adolescents age range 15–18 years	To examine whether the risk of online hate victimization is more likely when youth visited online sites containing potentially harmful content	Victimization, online hate social media	Facebook	Variables such as personality characteristics, moral values and peer influences Results show that being involved in the production of hate material and in researching such content one puts young people in danger	Social media; routine activity theory; victimization; online hate
Baldry et al. (2019)	5,058 adolescents age range 11–18 years	To investigate post-traumatic stress symptoms affecting the involvement in school bullying and cyberbullying of boys and girls according to the different bullying roles	Cyber bullying School bullying Cybervictimization	Not specified	Results show that school and cyberbullying are risk factors for development of post-traumatic stress symptoms differently affecting adolescents according to their role	School children, post-traumatic stress disorder, schools, symptoms, abuse, adolescents, aggressive behavior, boys, girls, human diseases, risk factors, risk groups, students, children
Longobardi et al. (2020)	345 adolescents age range 11–16 years	To analyze the association between Instagram popularity and subjective happiness and evaluate the relationship between roles of cyber victimization and social media addiction	Cyber victimization, Social media addiction	Instagram	Results show that social media use and cyber victimization were positively correlated, and both showed a negative correlation with perceived subjective happiness	Cyber victimization, Instagram, Peer exclusion, social media addiction, Happiness, Well-being, Adolescents
Blaya et al. (2020)	1,900 students age range 12–20 years	To examine the association between school bullying and cyberhate victimization and perpetration	Cyberhate, bullying, victimization and perpetration	Not specified	Results show that bullying and cyberhate are a common experience for many young people. In particular, the overlap between bullying and cyberhate and between traditional bullying and cyberbullying is evident	Cyberhate, Young people, Victimization, Involvement, School bullying, Overlap
Wachs et al. (2020)	1,480 adolescents age range 12–1ears	To investigate adolescents' coping strategies for cyberhate, while considering differences in adolescents' sex, age, socioeconomic status (SES), and victim status	Cyberhate, coping strategies, adolescents, cyberbullying	Not specified	Results show that different coping strategies are used by adolescents, with differences depending on sex, age, socioeconomic status, and victim status	Cyberhate Coping strategies Cybervictimization Hate speech Cyber discrimination

(Continued)

TABLE 1 | Continued

Author/s year	Characteristics of the sample	Subject of investigation	Hate speech, (Cyber)bullying	Social media	Main findings	Paper key word
Ang and Goh (2010)	396 adolescents age range 12–18 years	To examine the association between affective empathy, cognitive empathy, and gender on cyberbullying among adolescents	Cyberbullying among adolescent (for an example, hurting someone by sending them rude text messages)	Not specified	Results show that both boys and girls who had low cognitive empathy had higher scores on cyberbullying than those who had high cognitive empathy	Cyberbullying, Affective empathy, Cognitive empathy, Gender
Barlett et al. (2019)	3,079 youth participants average age 13.12 years (Wave 1) 1,957 youth participants (Wave 2) 909 youth participants (Wave 3)	To testing (a) the longitudinal stability in positive cyberbullying attitudes (CA), (b) whether any change in positive CA over time predict subsequent cyberbullying perpetration, and (c) the cross-lagged relations between positive attitudes toward CA and behavior over time	Positive cyberbullying attitudes, cyberbullying perpetration. (Relation between attitude and behavior)	Not specified	Results show a modest stability in positive CA and perpetration over time. Latent class analysis classified participants into either stable high attitudes, stable low attitudes, increasing attitudes, or decreasing attitudes	Cyberbully, cyberbullying attitudes
Gradingier et al. (2009)	761 adolescents age range 14–19 years	To analyzed whether students in the world both traditional and cyber belonging to groups of bullies or victims and bully and victims differed regarding adjustment	Traditional bullying, cyberbullying, traditional victimization, and cybervictimization	Mobile phones and computers (Not further specified)	Results show that the highest risks for poor adjustment were observed in students who were identified as combined bully-victims (traditional and cyber). In addition, gender differences were examined	Cyberbullying, cybervictimization, adjustment, aggression, configural frequency analysis
Schneider et al. (2012)	20,406 students 9th through 12th grade	To examine the prevalence of cyberbullying and school bullying victimization and their associations with psychological distress	Bullying victimization and psychological distress, including depressive symptoms, self-injury, and suicidality	Not specified	Results show that victimization was higher among no heterosexually identified youths. Victims report lower school performance and school attachment. Distress was highest among victims of both cyberbullying and school bullying	Adolescent, Bullying, Psychological epidemiology, Stress, Psychological etiology
Low and Espelage (2013)	1,023 early adolescents age range 10–15 years	To understanding the role of maladaptive family social dynamics on cyber-bullying and nonphysical bullying (i.e., verbal and relational) involvement through individual risk and protective factors	Cyber-bullying and nonphysical bullying (i.e., verbal and relational)	Not specified	Findings validate the importance of familial socialization. Cyber-bullying shows a significant overlap with nonphysical bullying, in particular, nonphysical bullying levels were associated with both higher family	Bullying, cyberbullying, longitudinal predictors, race, gender

(Continued)

TABLE 1 | Continued

Author/s year	Characteristics of the sample	Subject of investigation	Hate speech, (Cyber)bullying	Social media	Main findings	Paper key word
Mehari and Farrell (2018)	677 adolescents age range 11–15 years	To assess whether the dimensional model that cyberbullying that fits into a framework of adolescent aggression considered both form (overt or relational) and media (in-person or electronic) best fit the data	Form (overt or relational) and media (in-person or electronic) of aggression	Not specified	Violence and lower parental monitoring Results show that cyberbullying is a new form of aggression, a counterpart to overt and relational aggression, and this conceptualization fits the data quite well	Aggression, cyberbullying, adolescence, measurement of aggression, electronic media
Mishna et al. (2010)	2,186 students grades 6, 7, 10, and 11	To examine technology use, cyberbullying behaviors, and the psychosocial impact of bullying and being bullied, among a large sample of middle and high school students in a large urban center	Cyberbullying and technology use	Not specified	Results show that bullying was perpetrated by and toward friends and that bullies do not often disclose that they have been bullied. After the assault, the online bully reported feeling angry, sad, and depressed	Adolescents, Canada, cyber bullying, victimization, gender, ethnicity
Ortega-Ruiz et al. (2009)	1,671 adolescents age range 12–17 years	To examine the emotional impact caused to victims of traditional bullying or cyberbullying through technologies such as cell phones and the Internet	Direct bullying, indirect bullying, bullying inflicted via mobile phone, bullying inflicted via internet	Mobile phones and internet (no further specification)	Results show that traditional bullying affected more youth than cyberbullying. Cyberbullying produces emotional profiles like traditional bullying. The most common emotional response is anger and other negative emotions	Bullying, cyberbullying, emotions, victimization, adolescents
Patchin and Hinduja (2010)	1,963 students age range 10–16 years	Examines the relationship between cyberbullying and their level of self-esteem in the experience of middle school students	Cyberbullying	Email, MySpace or other (not further specified) web page	Results show that students who were both victims and perpetrators in cyberbullying had significantly lower self-esteem than those who had little or no experience with cyberbullying	Aggressive behavior, behavior, human behavior, programmes, psychological factors, school buildings, school kids, schoolchildren, teenagers, United States of America
Schultze-Krumbholz and Scheithauer (2009)	71 students grades 7, 8, and 10 average age 14.05 years	To identify characteristics of cyberbullies and cybervictims to be considered as potential risk and/or protective factors in a future study with a larger sample of students. Specifically, the	Cyberbullying and cybervictimization	Email, mobile phones and internet in general	Results show that higher frequency of cyberbullying compared with traditional bullying, and an overlap between cyberbullying and cybervictimization. Also, cyberbullies and cybervictims	Cyberbullying, social behavioral correlates, cybervictimization, empathy, frequency analysis

(Continued)

TABLE 1 | Continued

Author/s year	Characteristics of the sample	Subject of investigation	Hate speech, (Cyber)bullying	Social media	Main findings	Paper key word
		Research aims to assess the quality of several measurement instruments			Showed less empathy and higher relational aggression	
Ševčíková and Šmahel (2009)	Different age groups, including 223, respectively, 224 younger adolescents (age range 12–15 years) and 248, respectively, 249 older adolescents (age range 16–19 years)	To explore the frequency of online aggressive acts (as victim and aggressor)	Cyberbullying and aggressive behavior	Not specified	Adolescents (12–19 years old) were more often the target of aggressive behavior than older respondents	Harassment cyberbullying Internet Czech Republic
Waasdorp et al. (2018)	26,494 high school youth and 16,749 middle school youth	To analyze if weight status exacerbates the association between victimization and internalizing symptoms in bullied obese youth	Association between different forms of victimization, weight status, and adjustment	Not specified	Results highlight an increased risk of psychosocial adjustment problems among obese and overweight youth who are frequent victims of bullying. The odds of experiencing cyber victimization were higher than the odds of experiencing other forms of victimization	Internalization, Obesity, Overweight, Victimization, Bullying, At Risk Populations, Peers, Symptoms, Test Construction, Adolescent Characteristics, Internalizing Symptoms, Adolescence (13-17 yrs), Male, Female
Yang et al. (2020)	16,237 adolescents 6th–12th grade	To explore the relationship between cyberbullying victimization (CBV), student emotional engagement, and cognitive-behavioral engagement at both the student and school level	Traditional bullying victimization, cyberbullying victimization	Not specified	The most relevant findings suggest that CBV had a small but significant positive association with emotional engagement and a small but significant negative association with cognitive-behavioral engagement	Bullying victimization, cyberbullying victimization, school climate, student engagement
Zaborskis et al. (2018)	3,814 adolescents mean age 15.67 years	To analyze the prevalence of bullying and cyberbullying and their association with suicidal behavior among school-aged children in Israel, Lithuania, and Luxembourg	Cyberbullying and suicidality	Not specified	Results show that cyberbullying is a strong predictor of adolescent suicidality	Adolescents, bullying, cyberbullying, suicidality, associations

(2011) assessed the relative likelihood of reporting experiences of violence, such as bullying victimization, where the minority race (Black/African American) was found to be protective of all victimization experiences.

Individual and Contextual Factors

Other predictors for cyberbullying included personal traits and attitudes. For instance, Ang and Goh (2010) documented a three-way interaction in which high affective empathy moderated the effects of low cognitive empathy on cyberbullying for girls compared to boys. With regard to potential risk and/or protective factors, Schultze-Krumbholz and Scheithauer (2009) also considered empathy in their study. They found that cyberbullies and cybervictims showed less empathy and higher levels of relational aggression compared to students who did not engage in cyberbullying.

Barlett et al. (2019) examined the relationship between positive cyberbullying attitudes and subsequent cyberbullying perpetration in a longitudinal study of 3,000 Singaporean adolescents over a three-year period. They found that children with stable high or increasingly positive attitudes toward cyberbullying behaviors were also more likely to engage in such acts.

Using Data from a school-based census of about 20,400 youth, Schneider et al. (2012) showed that cyberbullying is far more frequent among nonheterosexual youth (33.1%), compared to heterosexual youth (14.5%). In their study, Mishna et al. (2010) asked the more than 2,100 participants whether they thought sexuality led to their bullying (2%) or was the reason they bullied others (2%).

In their study, including data from more than 43,200 adolescents from middle and high schools, Waasdorp et al. (2018) showed that overweight youth were more likely to report being a victim of cyberbullying (obese youth had even a 66% higher risk of being victims of cyberbullying). The findings of their study are supported by the work of Mishna et al. (2010), in which more than one in ten (11%) felt they have been victims of cyberbullying because of their appearance. According to this study, other characteristics that participants felt led to their bullying were disability (2%), family (2%) and school performance (5%).

School performance was also part of the investigation in other studies, with particular respect to the relationship between lower school performance and online victimization. While Bossler et al. (2012) reported that lower school performance is a predictor of online victimization, in the study of Schneider et al. (2012) no causality relationship between these two factors was analyzed, thus considering cyberbullying victimization as a potential predictor for school performance and vice versa. Referring to school as an important place for adolescents, a lower school attachment was also found to increase the likelihood of victimization online (Schneider et al., 2012) and, contrary to expectations, the negative influences of cyberbullying victimization on cognitive-behavior engagement were actually enhanced when students perceived a positive school climate. At the same time, the positive relationship between cyberbullying victimization and emotional engagement was mitigated by perceptions of a positive school climate (Yang et al., 2020).

Considering cyberbullying from the perpetrator's point of view, again appearance seems to be an important factor. Thus, further in the study of Mishna et al. (2010) 6% of the perpetrators stated that the appearance of the victim was the reason for their attacks. According to this study, other characteristics that participants stated as a reason for bullying others were disability (1%), school performance (3%) and family (2%). That cyberbullying occurs between parties who are familiar with each other or would even consider each other friends could also be demonstrated by Mishna et al. (2010). Here, friends (52%) were the most frequent targets of cyberbullying behavior. The influence of friends was also highlighted by Bossler et al. (2012), who positively associated a higher percentage of friends misbehaving on the computer with victimization. Besides the situation with friends, factors that an adolescent faces at home, and especially the quality of the caregiver-child relationship, seem to influence the likelihood of cyberbullying (Ang and Goh, 2010). In this context, parental monitoring and also the use of protective software seem to be associated with higher levels of cyberbullying (Bossler et al., 2012; Low and Espelage, 2013). At the same time, however, general use of technology also appears to be an indicator of an increased likelihood of being exposed to and experiencing violent media (Ybarra et al., 2011).

Overlap Between Traditional Bullying and Cyberbullying

The overlap between traditional bullying and cyberbullying has been investigated in several studies included in this review, revealing both similarities and differences between the two phenomena. Students' tendency not to report cyberbullying and their reasons for doing so are consistent with findings from studies examining traditional bullying. According to Mishna et al. (2010), these reasons include fear of retaliation or that the bullying could get worse. However, that young people also fear losing Internet or cell phone privileges seems to be rather a concern that occurs only in the context of disclosing cyberbullying.

Looking at the victims' perspective, the different forms of bullying often seemed to occur in parallel. Mitchell et al. (2011), analyzing data from more than 2,000 adolescents ages 10 to 17, found that 96% of youth who experienced online victimization also reported offline victimization during the same time period. Here, the offline victimizations linked most closely to online victimization were sexual victimizations (e.g., sexual harassment) and psychological and emotional abuse. The findings are in line with those found by Gradinger et al. (2009), who also found that most of the cyberbullying victims were also victims of traditional bullying at the same time.

In contrast, other studies, such as by Schultze-Krumbholz and Scheithauer (2009), concluded that cyberbullying is more common compared to traditional bullying. That cyberbullying and traditional bullying differ in frequency was also shown by Ortega-Ruiz et al. (2009). According to their study, however, the two phenomena are inversely related: significantly more adolescents were targeted by traditional bullying (two in ten) than by cyberbullying (one in ten). One in five participants reported being affected by both types of bullying.

Likewise, the consequences for victims of both forms appear to have similarities. For example, Schneider et al. (2012) found that the level of distress was highest for victims of both cyberbullying and school bullying. Additionally, Ortega-Ruiz et al. (2009) observed similar emotional responses to cyberbullying *via* the Internet and indirect bullying as a special type of traditional bullying (e.g., threats or insults). Emotions cited by victims included anger, stress, or fear. Although Low and Espelage (2013) also mention that cyberbullying seems to have significant overlaps with non-physical bullying (i.e., verbal and relational bullying), longitudinal analyses showed, according to them, less overlap between the different forms.

A connection between bullying at school (in the sense of traditional bullying) and cyberbullying is also assumed by Ševčíková and Šmahel (2009), who hypothesize that the reason for this could be the non-anonymous relationship between perpetrators and victims.

Predictors of Cyberhate Socio-Demographic Variables

In their study involving more than 700 Finnish youth aged 15–18 who used Facebook as a social medium, Oksanen et al. (2014) analyzed the extent to which this age group is exposed to and victimized by online hate material. Two-thirds, and thus the majority, of the youths stated that they had already encountered online hate material, with 21% of the respondents having been victims of online hate material themselves. Furthermore, the authors observed that 70% of the participants more accidentally came across the online hate material, while 22% of the youth intentionally searched for this type of content. As a result of their analysis, they state that none of the sociodemographic variables (e.g., age, gender, living with parents) were found to be significant predictors of exposure to online hate material. Further, they link victimization by online hate material to various social and psychological factors, such as negative offline experiences.

These results are consistent with those found by Wachs et al. (2021), who reported neither significant differences in gender between girls (19%) and boys (15.4%) as related to cyberhate victimization, nor age differences between victims and non-victims of cyberhate.

Although socio-demographic data do not appear to be influential in the context of cyberhate, it is worth noting that, first, in the study by Oksanen et al. (2014) online hate material most frequently targeted at ethnicity/nationality (50%) and religious belief/faith (43%), and second, victims of online hate material were more likely to report material in which sex/gender was targeted compared to non-victims.

Individual and Contextual Factors

Study findings presented by Oksanen et al. (2014) suggest that certain states of agitation may increase the likelihood to be a victim of online hate material. For example, in contrast to those who did not perceive themselves as victims of online hate material, youth who did perceive themselves as victims were more likely to report being worried. This aspect was taken up by the research group in a later paper based on the same data set (Räsänen et al., 2016). In this, they concluded that worrying about

becoming a victim of online hate material increases the likelihood of actually becoming a victim online.

According to the results of Oksanen et al. (2014), the hate material that victims saw online often directed at sexual orientation (68%), physical appearance (61%) and disability (31%). In addition, the authors observed that adolescents who were exposed to hate material online were more active online, not studying and their attachment to family was low. Specifically to the role played by being very active online, Ybarra et al. (2011) had achieved similar findings, and suggest that general technology use is an important factor in predicting risk for violent exposures (eg, hate sites) and experiences online.

The role of the family was again addressed in the study by Räsänen et al. (2016), but the analysis of the effects of covariates in the proposed model excluded the possibility to prove a correlation between living with parents and cyberhate victimization (even though basic statistics showed that not living with parents doubled the risk of online victimization).

Moreover, the authors report that the likelihood of becoming a victim of hate online is higher among youth who visit harmful sites on the Internet, the likelihood increases for those who deliberately search for this type of content, and the odds of victimization are almost four times higher for those producing hate materials. In line with this, Pauwels and Schils (2016) found that measures of extremism through new social media are associated with self-reported political violence. This relationship is most pronounced when users actively search for extremist content, with self-reported political violence being linked to various offline associations (e.g., racist and delinquent peers).

Looking at individual and contextual factors, Wachs et al. (2020) analyzed the influence of family affluence on cyberhate victimization. As the results of their study show, no differences were found between victims of cyberhate with low family affluence (33.9%), middle family affluence (31.7%), and high family affluence (35%).

Finally, Räsänen et al. (2016) found that, in contrast to the hypothesis that the number of friends on Facebook would increase exposure to hate material, this factor seems not to increase the likelihood of victimization. Slightly different is the influence of friends on cyberhate exposure, since Oksanen et al. (2014) found that 8% of the interviewed adolescents encountered hate material *via* a link from a friend.

Overlap of Cyberhate With Other Forms of Aggressive Behavior

As Wachs et al. (2020) showed, adolescents use similar coping strategies for dealing with cyberhate as they do for dealing with cyberbullying. Therefore, the authors not only assume conceptual and empirical overlaps between the two phenomena, but also that both forms have a similar impact on adolescents' behavior and emotions.

Based on the results of their study involving 1,900 French students, Blaya et al. (2020) noted that percentages of young people involved as victims or perpetrators were much higher offline than online.

In addition, they found that as victimization at schools (offline behavior) increased, the likelihood of exposure to hate

material online (online behavior) also increased. Furthermore, their findings indicate that students who insult or threaten others at school also spread hate messages against others online. The way cyberhate is related to other forms of aggressive behavior is shown by the further results of the study. Thus, a weaker relationship was observed between cyberhate victimization and cyberhate perpetration, a moderate relationship was observed between school bullying victimization and cyberhate perpetration, and a moderate relationship was observed between cyberhate victimization and school bullying perpetration.

That exposure to hate material can be associated with offline physical victimization was also noted by Oksanen et al. (2014). In their later work, Räsänen et al. (2016) again consider this factor, stating that the odds of online hate victimization are higher if the user has already experienced online victimization.

Impact of Cyberbullying

Victimization through cyberbullying can result in very different effects. For example, some studies linked online victimization to psychological distress (Mitchell et al., 2011; Schneider et al., 2012), with the odds of distress appearing to remain constant over time (Ybarra et al., 2011). Other studies have shown that victims reported depressive symptoms (Ortega-Ruiz et al., 2009; Schneider et al., 2012; Low and Espelage, 2013), which, like somatic symptoms (e.g., bellyaches and stomach cramps), were most common when victims experienced both traditional bullying and cyberbullying compared with non-victims (Gradinger et al., 2009). According to Mitchell et al. (2011) greater trauma symptomatology, including depression in addition to anger and anxiety, was slightly but significantly linked to online victimization.

Stress is also a reaction that was observed associated with cyberbullying. For instance, Ortega-Ruiz et al. (2009) found that victims who were more affected by cyberbullying (both *via* the Internet and mobile phone) felt more stressed than occasional victims. In the case of cyberbullying *via* the Internet, more women (13.7%) than men (2%) reported feeling stressed. This finding is in line with the conclusion drawn by Baldry et al. (2019) who describe post-traumatic stress as a psychophysiological condition “resulting from stressful traumatic events such as school bullying and cyberbullying”.

In the context of cyberbullying, it was noted that some adolescents do not seem to be bothered by online attacks. This was shown, for example, in the studies by Ortega-Ruiz et al. (2009) and Mishna et al. (2010), the latter also finding a gender difference between males (55.6%) and females (28.6%), but only for cyberbullying *via* mobile phone. At the same time, however, various negative emotions have been reported that appear to be the result of cyberbullying. Across different studies, victims referred to feeling afraid and/or scared (Ortega-Ruiz et al., 2009; Mishna et al., 2010; Patchin and Hinduja, 2010). Others reported feeling alone, defenseless and worried, with females (30.6%) more likely than males (5.6%) to be worried by cyberbullying *via* cell phone (Ortega-Ruiz et al., 2009). Longobardi et al. (2020) observed a negative correlation of cyber victimization and perceived subjective happiness, which is consistent with the findings of Mishna et al. (2010), in which victims reported feeling

sad. Feelings of embarrassment and upset (Ortega-Ruiz et al., 2009; Mishna et al., 2010) were also reported, with the number who felt very or extremely angry as a result of victimization remaining constant over a 36-month observation period (Ybarra et al., 2011). In addition, adolescents expressed feeling angry, as seen in the studies by Ortega-Ruiz et al. (2009) and Mishna et al. (2010), the latter group showing that more females (37%) than males (18%) reported feeling angry when bullied *via* the Internet.

As studies show, adolescents who experienced multiple forms of bullying and victimization appear to be at higher risk for poor adjustment. According to Gradinger et al. (2009), perpetrators who performed bullying online and offline were at highest risk for externalizing adjustment problems (e.g., reactive or instrumental aggression), whereas victims who experienced bullying online and offline were at highest risk for internalizing adjustment problems (e.g., depressive and somatic symptoms). In addition, adolescents who performed and experienced bullying online and offline were at highest risk for both externalizing and internalizing adjustment problems. In this context, it is worth mentioning the study by Waasdorp et al. (2018), which links victimization (including the experience of cyberbullying) to adjustment and social-emotional problems in addition to childhood obesity. And also the study by Mitchell et al. (2011) can be highlighted here, in which online victimization was strongly associated with delinquency (e.g., physically harming other children or adults, intentionally damaging things that belong to others, cheating on tests, skipping school) during the period studied.

Similarities between perpetrators and victims of cyberbullying do not only seem to exist with regard to adjustment problems. Thus, Patchin and Hinduja (2010) found a moderate relationship between both low self-esteem and cyberbullying offending and between low self-esteem and cyberbullying victimization. At the same time, however, victims of cyber- and/or school bullying are the ones who show an increased risk for suicidal behaviors. Zaborskis et al. (2018), analyzing data from a cross-national survey conducted in 2013 and 2014, showed that victims were at higher risk of suicidal thoughts, plans, and attempts regardless of the type of bullying they experienced. Among young people from Lithuania and Luxembourg (in addition, young people from Israel were interviewed), the association between cyberbullying and suicidal behavior was even greater compared to bullying that happened at school. Consistent with these findings Schneider et al. (2012) report suicide attempts of among victims of online and offline bullying, with cyberbullying victims (9.4%) more affected than school bullying victims (4.2%).

Impact of Cyberhate

According to Oksanen et al. (2014), there is an inverse relationship between cyberhate victimization and general psychological well-being, with victims of hate material more likely to be unhappy. In the context of emotional health, which can thus be affected by cyberhate victimization, coping strategies emerge into focus. In their study, Wachs et al. (2020) addressed the question of how adolescents deal with cyberhate. Taking into account differences in gender, age, socioeconomic status, and victimization status of the youth, six different coping strategies

were confirmed. To mitigate the negative effects of cyberhate, adolescents primarily used constructive coping strategies, namely Technical coping (i.e., blocking a person), Assertiveness (telling the person to stop), and Close support (distracting oneself by spending time with friends). According to the authors, the fact that young people responded in this way indicates high levels of digital literacy which they know how to use, as well as high levels of self-efficacy. The remaining three strategies included Helplessness/Self-blame (not knowing what to do), Retaliation (do it back), and Distal advice (go to the police). Considering gender and age, girls were more likely to use all coping strategies (except Retaliation), and younger adolescents were more likely to use Technical coping strategies than older adolescents. Socioeconomic status was relevant to the extent that Distal advice and Technical coping were more common among adolescents with lower socioeconomic status than among peers with higher socioeconomic status.

DISCUSSION

The papers selected for this review provide precise indications on the factors that characterize the phenomena of cyberbullying and cyberhate, both in terms of predictive variables and their impact on adolescents. As the focus of this review was on cyberbullying and cyberhate as two interlinked instances of cyber-aggression, we compare the predictors and effects of one or the other phenomenon, and show how many of these factors characterize both cyberbullying and cyberhate, thus highlighting the level of overlap between them.

We have decided to analyse the predicting variables and consequences of the two phenomena separately, instead of framing them under the broader umbrella of cyber-aggression. The comparison follows this analysis. In such a way, we have the possibility to catch a new perspective on the investigation of these two phenomena.

Before presenting the results of this comparison, it should also be pointed out that the selected articles do not always provide a whole picture of the individual factors, and for this reason it has been necessary to enrich the description of some of them with additional sources of literature.

Overlapping Between Predictors of Cyberbullying and Cyberhate for Adolescents

Amongst the common factors that can predict perpetration and victimization in both cyberbullying and cyberhate, the role of the family has drawn the attention of many scholars. Specifically, the role of the family has been analyzed from different perspectives, sometimes favoring the aspects related to the emotional parent-child relationship, sometimes considering the family as a proxy of the social guardianship that can be a deterrent to cyber-aggression. The latter is particularly present in papers presenting studies that have borrowed the Routine Activity Theory (Cohen and Felson, 1979), one of the main theories of criminology, to explain the phenomena of cyber-aggression.

The importance of the parent-child relationship is highlighted both in the literature on cyberhate, with Oksanen et al. (2014) who found that weak attachment to family significantly predicts exposure to online hate content, and in the literature on cyberbullying, with Ang and Goh (2010) who highlighted the importance of positive caregiver-child relationships in reducing cyberbullying behavior among adolescents. These findings are consistent with similar studies on bullying and cyberbullying; among others, Murphy et al. (2017) found that attachment to parents is a deterrent to both becoming bullies and becoming victims of bullying, factors; Wang et al. (2009) found that bullying and cyber-bullying were similarly related to low parental support.

Surprisingly, the protective role of the families against the risks of cyberbullying, which is highlighted in numerous studies in the literature, including the impressive work by Li et al. (2021) with almost 215,000 school-aged children across 41 countries, has not been explicitly targeted in the papers on cyberbullying selected for this review, with the only exception of the paper by Low and Espelage (2013), who found a positive association between parental monitoring and higher levels of cyber-bullying perpetration (only for white adolescents). The guardianship offered by the parents against the risks of online hate exposure and victimization has been analyzed in Oksanen et al. (2014) and Räsänen et al. (2016). These studies did not find a correlation between family guardianship and online hate material exposure (Oksanen et al., 2014), nor was it possible to prove a correlation between living with parents and cyberhate victimization (Räsänen et al., 2016).

Although this would seem to indicate that the protective role of the family is a further overlapping factor between the phenomena of cyberbullying and cyberhate, as no significant link with either phenomenon was found in the literature reviewed, these results should be further commented on in light of the fact that they seem to contradict numerous studies on cyber-aggression (Li et al., 2021). Of particular significance, Räsänen et al. state in their paper: “Therefore, simply living with one’s parents does not appear to ensure guardianship. Thus, it is difficult to interpret the lack of significance of this variable” (p. 14).

A possible interpretation could be found by investigating the mode and quality of guardianships exercised by parents. In this regard, it is worth mentioning the recent study by Wachs et al. (2021) that has furtherly analyzed the relationships between cyberhate victimization and the form of parental mediation, and found that instructive parental mediation is negatively associated with cyberhate victimization, while restrictive parental mediation determines the opposite effect. This confirms, moreover, the results of a previous study on the protective role of the family (Papatraianou et al., 2014), where the authors highlight the importance of instructive parental mediation: “Strong family relationships within the context of a young person’s home can also help young people overcome online adversity, along with family permissions to use technology in a safe way”.

Closely related to the role of adolescent-family relations on cyberbullying and cyberhate, and with similar outcomes, is the role played by relations with friends. The literature analyzed does not provide consistent results, nor is it possible to give an

unambiguous interpretation by extending the literature analysis to articles not included in this review. Specifically, Mishna et al. (2010) found that friends are the most frequent targets of cyberbullying attacks (in 52% of the cases studied), and this percentage increases to 84% in an earlier study by Ybarra and Mitchell (2004). Although less evident, the negative role of friends is also confirmed in relation to cyberhate, where Oksanen et al. (2014) found that friends are one of the sources from which adolescents receive links to hate material (in 8% of the cases analyzed), thus favoring their exposure to cyberhate. If these results were generalisable, it could be assumed that as the number of friends increases, the risks of Cyberbullying victimization or cyberhate exposure and victimization should increase. However, Räsänen et al. (2016) found that an increase in the number of friends on Facebook did not correspond to an increase in the risk of cyberhate exposure; the same result was reached by Kaakinen et al. (2018). Similarly, Wang et al. (2009) found that having more friends is not associated with cyberbullying.

The synthesis of these results could be found by prioritizing the analysis of the quality of relationships with friends or of the behaviors usually carried out by friends, rather than focusing on the number of friends. Bossler et al. (2012) underline that friends who caused most online harassment were those who committed various forms of computer deviance. The quality of the relationship with friends in relation to the phenomena of cyberbullying is underlined by subsequent works (not selected for this review), which recognize the protective role of friends against cyber-aggression: Papatraianou et al. (2014) have pointed out how strong and supportive friend relationships can support female adolescents' resilience toward online risks and aggression; similarly, Zych et al. (2019) have verified that the quality of relationships with friends is a strong protective factor against cyberbullying. Nevertheless, other scholars have only partially confirmed these results. For example, Bedrosova et al. (2022), who analyzed these aspects with samples of adolescents in the Czech Republic, Poland and Slovakia, found that friendship support was negatively related to cyberhate in the Czech Republic and Poland, but not in Slovakia and, even more surprisingly, friendship support was negatively related to cyberbullying only in the Czech Republic. Similar results were found by Kaakinen et al. (2018) who analyzed, with samples of American, British, German and Finnish adolescents and young adults, how cognitive social capital in the offline context (i.e., trust and sense of belonging in a group of friends) influences cyberhate victimization. In addition to the finding that the number of Facebook friends was not associated with online hate victimization reported above, the authors found that trust and sense of belonging in a group of friends was negatively associated with online hate victimization in all samples, but not for the Finnish one. With everything considered, we therefore encourage further studies on the role of friends in relation to cyberbullying and cyberhate.

The constructs related to sexuality (sexual orientation; sexual identification; etc.) represent a further element of overlapping between cyberbullying and cyberhate, being predictors of perpetration, victimization and exposure to online hate material (Mishna et al., 2010; Schneider et al., 2012; Oksanen et al., 2014). This is not surprising, given that the sexual sphere has always

been a reason for discrimination, both at an individual level and with regard to groups that feel the need to unite in order to fight against discriminatory stereotypes that societies cannot ignore (Russell et al., 2001; Robin et al., 2002; Williams et al., 2003).

Overlapping Between the Impact of Cyberbullying and Cyberhate on Adolescents

The dimension that offers the major number of insights on the overlap between the consequences of cyberbullying and cyberhate on adolescents' individual well-being and emotions. In fact, the negative effects of cyber-aggression on emotional perception was found by authors who analyzed overall subjective happiness, direct emotional responses to experiences and long term emotional states induced by cyberbullying victimization and perpetration, as well as by cyberhate victimization or exposure (Ortega-Ruiz et al., 2009; Mishna et al., 2010; Ybarra et al., 2011; Longobardi et al., 2020; Wachs et al., 2020). In particular, this confirms Wachs et al. (2021) argument that the impact of cyberhate and cyberbullying on adolescents' emotions may be similar. Specifically to this point, it is worth mentioning Catherine Blaya, one of the authors of the EU report on the relation between cyberhate and kids (Machackova et al., 2020), who points out that "the emotional consequences are significant not only for victims but also for witnesses even though they are not targeted by the posted hateful contents. Both groups report experiencing anger and hate following their exposure or victimization" (as reported in Bedrosova, 2020). This confirms that the boundary between exposure to cyberhate and cyberhate victimization regarding its impact on users' emotions is extremely blurred (Machackova et al., 2020).

Strictly related to the effects of cyberbullying and cyberhate on adolescents' emotions, many studies have reported negative effects of cyberbullying victimization and perpetration on individuals' wellbeing, mainly consisting in depressive symptoms, somatic symptoms, post-traumatic stress symptoms and psychological distress (Gradinger et al., 2009; Ortega-Ruiz et al., 2009; Mitchell et al., 2011; Schneider et al., 2012; Fales et al., 2018). Although the selected literature does not provide similar information for adolescents who were exposed to or victims of cyberhate, an online survey administered to 1,512 adolescents (13–18 years.) in 2016 in UK revealed that young people who had been exposed to online hate content reacted to it with anger (37%), sadness (34%) and shock (30%) feelings (UK Safer Internet Centre, 2016). A distinct study involving young people in six countries slightly older than adolescents (18–25 years.) achieved similar results; respondents who had been exposed to online hate speech content reported almost the same negative emotional feelings as the adolescents in the UK survey: anger, sadness and shame (Reichelmann et al., 2020). Hence, findings from both studies on the consequences of cyberhate exposure and victimization identified symptoms which are common to the ones reported by the literature on cyberbullying, thus highlighting a further area of overlapping between cyberbullying and cyberhate.

Adolescents' coping strategies for cyberhate have been analyzed by Wachs et al. (2020), while the literature on cyberbullying selected for this review does not address the issue of how adolescents deal with cyberbullying attacks. Nevertheless, the overlap has been highlighted by Wachs et al., who found out that adolescents use similar coping strategies for dealing with cyberhate as they do for dealing with cyberbullying. Specifically, the conclusions achieved by the authors are similar to those pointed out by other authors who have studied adolescents' coping strategies for cyberbullying (Livingstone et al., 2011; Machackova et al., 2013; Sticca et al., 2015).

Papers in this review underline the role of online activities in cyberbullying and cyberhate phenomena. As expected, the more adolescents spend their time online, the more they are involved in cyberbullying and cyberhate exposure (Ybarra et al., 2011; Oksanen et al., 2014; Räsänen et al., 2016). In this sense, the frequency in using Internet online tools is a predictor for both cyberbullying and cyberhate experiences. Ybarra et al. (2011) extend this concept, confirming that technology use in general is a predictor of both cyberbullying experiences and cyberhate exposure.

Cyberbullying and Cyberhate: Distinguishing Features

The analysis of the literature shows that the concepts of cyberbullying and cyberhate are in part overlapping, but have some characteristics that distinguish them from each other. In particular, by examining the results related to adjustment problems and the ideation of suicide, some important differences can be observed.

Adjustment problems and suicide have been targeted by some of the papers on cyberbullying selected for this review. Specifically, Waasdorp et al. (2018) found that adolescents who have been victims of cyberbullying appear to be more likely to experience adjustment problems; Gradinger et al. (2009) revealed that both bullies and victims are at high risk of adjustment problems, especially if they are involved in both face-to-face and cyberbullying experiences; Schneider et al. (2012) and Zaborskis et al. (2018) identified attempted suicide as a consequence of cyberbullying.

These themes are not present in the literature on cyberhate, probably because they reflect a deep psychological discomfort that can lead to extreme gestures such as suicide, a discomfort that emerges when the victim of the attack is the individual adolescent rather than a group of people (even if the adolescent identifies with the group). Previous research on the consequences of discrimination (online and offline) on adolescents' mental well-being can support this assertion. Discrimination is, in fact, a transversal theme to cyberbullying and cyberhate, where in the first case it is a tool aimed at hurting the individual, while in the second case it is a manifestation of hatred against a group of individuals (based on gender, race, religion, etc.), with which the adolescent may or may not recognize himself. Studies analyzing the effects of discrimination against individuals confirm how the risk of adjustment problems and the number of suicidal ideations and attempts increase for adolescents who

have been discriminated against (Sinclair et al., 2012). The results are different if the discrimination is directed at a group. Particularly relevant is the work of Tynes et al. (2008) who present the results of a cross-sectional survey with 264 US high school students aged 14–18 years old to examine the impact of online racial discrimination on adolescents' psychological well-being. The authors distinguish between individual and vicarious discrimination: the former includes acts of discrimination that are explicitly directed at the individual, similar to what occurs in cyberbullying. Vicarious discrimination refers to discrimination acts directed at same-race adults and peers in the adolescent's life, similar to what occurs in cyberhate. This study confirmed that individual racial discrimination is significantly related to depression and anxiety. However, vicarious discrimination does not correlate with measures of psychological adjustment, thus confirming our assertion that attacks on groups of individuals typical of cyberhate are experienced less dramatically by adolescents than attacks experienced by cyberbullying victims.

Another aspect that differentiates the two types of cyber-aggression analyzed in this review is that certain individual and personal characteristics often related to physical appearance (obesity, overweight, disability) can be predictors of cyberbullying victimization (Mishna et al., 2010; Waasdorp et al., 2018), but are not present among the predictors of cyberhate victimization. This is inherent in the defining characteristics of the two phenomena, since cyberbullying is an aggressive behavior against a person, whereas cyberhate is against a group of individuals, and therefore individual and personal physical factors have much less relevance. Exceptions are those physical traits that can be associated with ethnic groups (e.g., skin color; eye shape, etc.) and are often used as the basis for discriminatory phenomena. This would seem to include the physical appearance which Oksanen et al. (2014) include among the predictors of cyberhate victimization, although they do not specify which specific traits of physical appearance they refer to.

The analysis of the literature has shown that gender is a predictor for cyberbullying perpetration and victimization (Gradinger et al., 2009; Ortega-Ruiz et al., 2009; Mishna et al., 2010; Schneider et al., 2012; Low and Espelage, 2013), with a few exceptions (Mehari and Farrell, 2018; Longobardi et al., 2020). In contrast, gender does not appear to be a predictor for cyberhate exposure and victimization (Oksanen et al., 2014; Räsänen et al., 2016; Wachs et al., 2020). This result seems surprising in light of the fact that the gender is one of the categories targeted by haters (such as race, religion, etc.), but at the same time it confirms the lack of consistent findings in the literature on the relationships between gender and cyberhate perpetration, exposure and victimization (Bauman et al., 2021).

Similarly, it was not possible to determine an overlap between cyberbullying and cyberhate with regard to race/ethnicity, cultural context, language spoken at home (here considered as a proxy of ethnicity). In fact, the literature on cyberhate clearly indicates that race is a predictor of cyberhate exposure and victimization (Oksanen et al., 2014), as could be expected, since race appears as one of the discriminating factors that lead haters to attack groups of people on the basis of the color of their skin, their ethnicity, and their culture. On the contrary, the

analysis of the selected literature (Mishna et al., 2010; Ybarra et al., 2011; Schneider et al., 2012; Low and Espelage, 2013) does not clearly show a causal link between race and cyberbullying experiences. This confirms what has already been pointed out in previous studies. In particular, in a previous review in which the relationships between cyberbullying, race/ethnicity and mental health outcomes were analyzed, the authors indicated that young whites are bullied more than their non-white peers, but clarify that it is not possible to establish whether this is a direct relationship (consequence of race and ethnicity), or rather is due to other factors that differentiate youth of color and their white peers in terms of technology ownership, social media preferences, and socioeconomic backgrounds (Edwards et al., 2016).

Finally, the impact of self-esteem, empathy and high levels of relational aggression on cyberhate experiences has not been sufficiently analyzed in the literature, so that it is not possible to make comparisons with what emerged for cyberbullying (Schultze-Krumbholz and Scheithauer, 2009; Ang and Goh, 2010; Patchin and Hinduja, 2010).

CONCLUSIONS

The first and most evident results from this review are that the cyberhate issue related to adolescents is less investigated than cyberbullying, and most of the papers dealing with one or the other phenomenon lacks of a holistic perspective, rooted in the broader concept of cyber-aggression, which makes it possible to approach cyberbullying and cyberhate as two distinct but often interconnected phenomena. In particular, the literature on cyberbullying lacks references to cyberhate, whereas the papers on cyberhate sometimes refer to literature on cyberbullying.

Nevertheless, by comparing the predictors and outcomes of cyberbullying and cyberhate, important overlapping factors between the two concepts emerge.

The most evident overlapping factors, as highlighted by this review, are the importance of the parent-child relationship to reduce the risk of cyber-aggression; the constructs related to sexuality (sexual orientation; sexual identification; etc.) as predictors of both phenomena; the protective role of the families against cyberbullying and cyberhate attacks, provided that parents offer instructive mediation while restrictive parental mediation determines the opposite effect; the role of good quality friendship relationships as deterrent against cyberbullying and cyberhate attacks; the impact of cyberbullying and cyberhate on adolescents' emotions as well as their consequences on individuals' wellbeing, mainly consisting in depressive symptoms, somatic symptoms, post-traumatic stress symptoms and psychological distress; the same coping strategies put in practice by victims of the two phenomena.

In addition to the factors common to cyberbullying and cyberhate, the literature highlights some of the characteristics that distinguish each of the two phenomena. In particular, differences concern the adjustment problems and the ideation of suicide, which have been found in studies on cyberbullying but

not on cyberhate; individual and personal characteristics, often related to physical appearance (obesity, overweight, disability), as predictors of cyberbullying victimization only; the gender as a predictor for cyberbullying perpetration and victimization, while it does not appear to be a predictor for cyberhate exposure and victimization; the lack of a well-defined overlap between cyberbullying and cyberhate with regard to race/ethnicity, cultural context, language spoken at home (here considered as a proxy of ethnicity); the impact of self-esteem, empathy and high levels of relational aggression on cyber-aggression, even though this issue has not been sufficiently analyzed in the literature on cyberhate.

We argue that the results of this review can stimulate future research on cyberbullying and cyberhate where the two phenomena are analyzed as two interlinked instances of cyber-aggression, while respecting their distinctive features. Moreover, further research should investigate the effectiveness of prevention and intervention programs based on the shared commonalities and reciprocal influence of cyberbullying and cyberhate (e.g., the same coping strategies should be assessed against their capacity to empower adolescents regarding cyberhate and cyberbullying), according to a holistic approach to the general problem of cyber-aggression in adolescence.

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.

AUTHOR CONTRIBUTIONS

GF: conceptualization, formal analysis, and paper drafting and revising. DT: conceptualization and paper drafting and revising. LS: conceptualization, formal analysis, data curation, methodology, and paper drafting. VS: formal analysis, data curation, methodology, and paper drafting and revising. SE: conceptualization and paper revising. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been developed in the framework of the project COURAGE—A social media companion safeguarding and educating students (No. 95567), funded by the Volkswagen Foundation in the topic Artificial Intelligence and the Society of the Future.

ACKNOWLEDGMENTS

The authors are particularly grateful to Lena Hilbig, Dr. Johanna Schäwel and Jonathan Mehl for their suggestions and advice.

REFERENCES

- AACAP. (2018). *The American Academy of Child and Adolescent Psychiatry. Social Media and Teens*. Available online at: https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Social-Media-and-Teens-100.aspx (accessed March 2018).
- Al-Hassan, A., and Al-Dossari, H. (2019). Detection of hate speech in social networks: a survey on multilingual corpus. In: *Computer Science and Information Technology (CS and IT)*. India: AIRCC Publishing Corporation. p. 84–100. doi: 10.5121/csit.2019.90208
- Anderson, M., and Jiang, J. (2018). Teens, social media and technology 2018. *Pew Res. Center*. 31, 1673–1689.
- Ang, R. P., and Goh, D. H. (2010). Cyberbullying among adolescents: The role of affective and cognitive empathy, and gender. *Child Psychiatry Human Develop.* 41, 387–397. doi: 10.1007/s10578-010-0176-3
- Anti-Defamation League. (2010). *Responding to Cyberhate, Toolkit for Action (Preprint)*. Available at: <https://www.adl.org/sites/default/files/documents/assets/pdf/combating-hate/ADL-Responding-to-Cyberhate-Toolkit.pdf> (accessed June 15, 2021).
- Atran, S., and Ginges, J. (2015). “Devoted actors and the moral foundations of intractable intergroup conflict” in *The moral brain: a multidisciplinary perspective*, eds. J. Decety and T. Wheatley (Boston Review), 69–85.
- Backe, E. L., Lilleston, P., and McCreary-Sills, J. (2018). Networked individuals, gendered violence: a literature review of cyberviolence. *Violence Gender*. 5, 135–146. doi: 10.1089/vio.2017.0056
- Baldry, A. C., Sorrentino, A., and Farrington, D. P. (2019). Post-traumatic stress symptoms among Italian preadolescents involved in school and cyber bullying and victimization. *J. Child Family Stud.* 28, 2358–2364. doi: 10.1007/s10826-018-1122-4
- Barlett, C. P., Gentile, D. A., Dongdong, L., and Khoo, A. (2019). Predicting cyberbullying behavior from attitudes: A 3-year longitudinal cross-lagged analysis of Singaporean youth. *J. Media Psychol.* 31. doi: 10.1027/1864-1105/a000231
- Bauman, S., Perry, V. M., and Wachs, S. (2021). The rising threat of cyberhate for young people around the globe. *Child Adolesc. Online Risk Exposure*. 149–175. doi: 10.1016/B978-0-12-817499-9.00008-9
- Bedrosova, M. (2020). European children’s experiences of cyberhate. *Parenting for a Digital Future. Blog Entry. Published by LSE—The London School of Economics and Political Science*. Available online at: <https://blogs.lse.ac.uk/parenting4digitalfuture/2020/06/24/euko-cyberhate/> (accessed June 24, 2020).
- Bedrosova, M., Machackova, H., Šerek, J., Smahel, D., and Blaya, C. (2022). The relation between the cyberhate and cyberbullying experiences of adolescents in the Czech Republic, Poland, and Slovakia. *Comput. Human Behav.* 126, 107013. doi: 10.1016/j.chb.2021.107013
- Blaya, C., and Audrin, C. (2019). Toward an Understanding of the Characteristics of Secondary School Cyberhate Perpetrators. *Front. Educ.* 4, 46. doi: 10.3389/educ.2019.00046
- Blaya, C., Audrin, C., and Skrzypiec, G. (2020). School Bullying, perpetration, and cyberhate: Overlapping issues. *Contemp. Sch. Psychol.* 1–9. doi: 10.1007/s40688-020-00318-5
- Bossler, A. M., Holt, T. J., and May, D. C. (2012). Predicting online harassment victimization among a juvenile population. *Youth Soc.* 44, 500–523. doi: 10.1177/0044118X11407525
- Chetty, N., and Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression Violent Behav.* 40, 108–118. doi: 10.1016/j.avb.2018.05.003
- Chisholm, J. F. (2006). Cyberspace violence against girls and adolescent females. *Annals N. Y. Acad. Sci.* 1087, 74–89. doi: 10.1196/annals.1385.022
- Cohen, L., and Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *Am. Sociol. Rev.* 44, 588–608. doi: 10.2307/2094589
- Corcoran, L., Guckin, C. M., and Prentice, G. (2015). Cyberbullying or cyber aggression?: A review of existing definitions of cyber-based peer-to-peer aggression. *Societies*. 5, 245–255. doi: 10.3390/soc5020245
- Costello, M., Barrett-Fox, R., Bernatzky, C., Hawdon, J., and Mendes, K. (2020). Predictors of viewing online extremism among America’s youth. *Youth Soc.* 52, 710–727. doi: 10.1177/0044118X18768115
- Council of Europe. (1997). *RECOMMENDATION No. R (97) 20 OF THE COMMITTEE OF MINISTERS TO MEMBER STATES ON “HATE SPEECH”*. Available online at: <https://rm.coe.int/1680505d5b> (accessed June 15, 2021).
- Council of Europe. (2003). *Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems*. 28. *European Treaty Series—No. 189*. Available online at: <https://rm.coe.int/168008160f> (accessed June 15, 2021).
- Edwards, L., Kontostathis, A. E., and Fisher, C. (2016). Cyberbullying, race/ethnicity and mental health outcomes: a review of the literature. *Media Commun.* 4, 71–78. doi: 10.17645/mac.v4i3.525
- ElSherief, M., Nilizadeh, S., Nguyen, D., Vigna, G., and Belding, E. (2018). Peer to Peer Hate: Hate Speech Instigators and Their Targets. *Proc. Intern. AAAI Conf. Web Social Media*. 12.
- Englander, E. (2017). Defining Cyberbullying. *Pediatrics*. 140, S148–S151. doi: 10.1542/peds.2016-1758U
- Eurostat (2020). *Community Survey on ICT Usage in Households and by Individuals*. Available online at: <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210630-1> (2020).
- Fales, J. L., Rice, S., Aaron, R. V., and Palermo, T. M. (2018). Traditional and cyber-victimization among adolescents with and without chronic pain. *Health Psychol.* 37, 291–300. doi: 10.1037/hea0000569
- Gerstenfeld, P. B., Grant, D. R., and Chiang, C. P. (2003). Hate online: a content analysis of extremist Internet sites. *Anal. Social Issues Public Policy*. 3, 29–44. doi: 10.1111/j.1530-2415.2003.00013.x
- Goerzig, A., Wachs, S., and Wright, M. (2019). Cyberhate and cyberbullying: joint propensity and reciprocal amplification. Full panel: victims and perpetrators of hate speech. In: *Annual Scientific Meeting of the International Society of Political Psychology*. Lisbon (2019). (Unpublished).
- Grading, P., Strohmeier, D., and Spiel, C. (2009). Traditional bullying and cyberbullying: Identification of risk groups for adjustment problems. *Z. Psychol.* 217, 205–213. doi: 10.1027/0044-3409.217.4.205
- Grigg, D. W. (2010). Cyber-aggression: Definition and concept of cyberbullying. *Aust. J. Guid. Couns.* 20, 143–156. doi: 10.1375/ajgc.20.2.143
- Harriman, N., Shortland, N., Su, M., Cote, T., Testa, M. A., and Savoia, E., Youth (2020). Exposure to hate in the online space: an exploratory analysis. *Int. J. Environ. Res. Public Health*. 17, 8531. doi: 10.3390/ijerph17228531
- Hawdon, J., Bernatzky, C., and Costello, M. (2019). Cyber-routines, political attitudes, and exposure to violence-advocating online extremism. *Social Forces*. 98, 329–354. 2014-ZA-BX-0014 doi: 10.1093/sf/soy115
- Kaakinen, M., Keipi, T., Oksanen, A., and Räsänen, P. (2018). How does social capital associate with being a victim of online hate? Survey evidence from the United States, the United Kingdom, Germany, and Finland. *Policy Internet*. 10, 302–323. doi: 10.1002/poi3.173
- Kemp, S. (2020). *Digital 2020: July Global Statshot Report Hootsuite and We Are Social*. Available at: <https://datareportal.com/reports/digital-2020-july-global-statshot> (accessed September 13, 2020).
- Kemp, S. (2021a). *Digital 2021: April Global Statshot Report Hootsuite and We Are Social*. Available online at: <https://datareportal.com/reports/digital-2021-april-global-statshot> (accessed June 10, 2021).
- Kemp, S. (2021b). *The Social Media Habits Of Young People In South-East Asia*. Available online at: <https://datareportal.com/reports/digital-youth-in-south-east-asia-2021> (accessed July 28, 2021).
- Lee, E., and Leets, L. (2002). Persuasive storytelling by hate groups online: Examining its effects on adolescents. *Am. Behav. Sci.* 45, 927–957. doi: 10.1177/0002764202045006003
- Li, Q., Luo, Y., Hao, Z., Smith, B., Guo, Y., and Tyrone, C. (2021). Risk factors of cyberbullying perpetration among school-aged children across 41 countries: a perspective of routine activity theory. *Int. J. Bull. Prevent.* 3, 168–180. doi: 10.1007/s42380-020-00071-6
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Götzsche, P. C., and Ioannidis, J. P. A. (2009). The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J. Clin. Epidemiol.* 62, 1–34. doi: 10.1016/j.jclinepi.2009.06.006
- Livingstone, S., Gorzig, A., and Olafsson, K. (2011). *Disadvantaged children and online risk*. London: EU Kids Online, London School of Economics and

- Political Science. Available online at: <http://eprints.lse.ac.uk/39385/> (accessed November 10, 2011).
- Longobardi, C., Settanni, M., Fabris, M. A., and Marengo, D. (2020). Follow or be followed: exploring the links between Instagram popularity, social media addiction, cyber victimization, and subjective happiness in Italian adolescents. *Child. Youth Serv. Rev.* 113, 104955. doi: 10.1016/j.chilyouth.2020.104955.
- López, C. A., and López, R. M. (2017). "4.8 hate speech, cyberbullying and online anonymity" in *Online Hate Speech in the European Union A Discourse Analytic Perspective* (Stavros Assimakopoulos Fabienne H. Baider Sharon Millar), 80–83.
- Low, S., and Espelage, D. (2013). Differentiating cyber bullying perpetration from non-physical bullying: Commonalities across race, individual, and family predictors. *Psychol. Violence*. 3, 39–52. doi: 10.1037/a0030308
- MacAvaney, S., Yao, H. R., Yang, E., Russell, K., Goharian, N., and Frieder, O. (2019). Hate speech detection: challenges and solutions. *PLoS One*. 14. doi: 10.1371/journal.pone.0221152
- Machackova, H., Blaya, C., Bedrosova, M., Smahel, D., and Staksrud, E. (2020). *Children's Experiences With Cyberhate*. EU Kids Online.
- Machackova, H., Cerna, A., Sevcikova, A., Dedkova, L., and Daneback, K. (2013). Effectiveness of coping strategies for victims of cyberbullying. *Cyberpsychol. J. Psychosoc. Res. Cyberspace*. 7:1–2. doi: 10.5817/CP2013-3-5
- Mardianto, H., Anurawan, F., Chusniyah, T., Rahmawati, H., Ildil, I., and Pratama M (2019). Cyber aggression of students: the role and intensity of the use of social media and cyber wellness. *Int. J. Innov. Creat. Change*. 5, 567–582. doi: 10.22216/jbe.v1i1.3349
- Mehari, K. R., and Farrell, A. D. (2018). Where does cyberbullying fit? A comparison of competing models of adolescent aggression. *Psychol. Violence*, 8, 31–42. doi: 10.1037/vio0000081
- Mishna, F., Cook, C., Gadalla, T., Daciuk, J., and Solomon, S. (2010). Cyber bullying behaviors among middle and high school students. *Am. J. Orthopsychiatry*. 80, 362–374. doi: 10.1111/j.1939-0025.2010.01040.x
- Mitchell, K. J., Finkelhor, D., Wolak, J., Ybarra, M. L., and Turner, H. (2011). Youth internet victimization in a broader victimization context. *J. Adolesc. Health*. 48, 128–134. doi: 10.1016/j.jadohealth.2010.06.009
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and P. R. I. S. M. A., Group (2009). Preferred reporting items for Systematic reviews and meta-analyses: the prisma statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed.1000097
- Murphy, T. P., Laible, D., and Augustine, M. (2017). The influences of parent and peer attachment on bullying. *J. Child Family Stud.* 26, 1388–1397. doi: 10.1007/s10826-017-0663-2
- Nagle, J. (2018). Twitter, cyber-violence, and the need for a critical social media literacy in teacher education: a review of the literature. *Teach. Teach. Educ.* 76, 86–94. doi: 10.1016/j.tate.2018.08.014
- O'Keeffe, G. S., Clarke-Pearson, K., and Council on Communications and Media. (2011). The impact of social media on children, adolescents, and families. *Pediatrics*. 127, 800–804. doi: 10.1542/peds.2011-0054
- Oksanen, A., Hawdon, J., Holkeri, E., Näsi, M., and Räsänen, P. (2014). Exposure to online hate among young social media users. *Sociol. Stud. Child. Youth*. 18, 253–273. doi: 10.1108/S1537-466120140000018021
- Ortega-Ruiz, R., Elípe, P., Mora-Merchán, J. A., Calmaestra, J., and Vega, E. (2009). The emotional impact on victims of traditional bullying and cyberbullying a study of spanish adolescents. *Zeitschrift für Psychologie* 217:197–204. doi: 10.1027/0044-3409.217.4.197
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 372, 71. doi: 10.1136/bmj.n71
- Papatrianou, L. H., Levine, D., and West, D. (2014). Resilience in the face of cyberbullying: An ecological perspective on young people's experiences of online adversity. *Pastoral Care in Education*. 32, 264–283. doi: 10.1080/02643944.2014.974661
- Patchin, J. W., and Hinduja, S. (2010). Cyberbullying and Self-Esteem. *J. School Health*. 80: 614–621. doi: 10.1111/j.1746-1561.2010.00548.x
- Pauwels, L., and Schils, N. (2016). Differential online exposure to extremist content and political violence: Testing the relative strength of social learning and competing perspectives. *Terror. Pol. Violence*. 28:1, 1–29. doi: 10.1080/09546553.2013.876414
- Peterson, J. K., and Densley, J. (2017). Cyber violence: What do we know and where do we go from here?. *Aggr. Violent Behav.* 34, 193–200. doi: 10.1016/j.avb.2017.01.012
- Pyzalski, J. (2012). From cyberbullying to electronic aggression: Typology of the phenomenon. *Emot. Behav. Difficult.* 17, 305–317. doi: 10.1080/13632752.2012.704319
- Räsänen, P., Hawdon, J., Holkeri, E., Keipi, T., Näsi, M., and Oksanen, A. (2016). Targets of online hate: Examining determinants of victimization among young Finnish Facebook users. *Violence Victims*. 31:4, 708–725. doi: 10.1891/0886-6708.VV-D-14-00079
- Reichmann, A., Hawdon, J., Costello, M., Ryan, J., Blaya, C., Llorent, V., et al. (2020). Hate knows no boundaries: online hate in six nations. *Deviant Behav.* 42, 1100–1111. doi: 10.1080/01639625.2020.1722337
- Robin, L., Brener, N. D., Donahue, S. F., Hack, T., Hale, K., and Goodenow, C. (2002). Associations between health risk behaviors and opposite-, same-, and both-sex sexual partners in representative samples of Vermont and Massachusetts high school students. *Arch. Pediatrics Adolesc. Med.* 156, 349–355. doi: 10.1001/archpedi.156.4.349
- Russell, S. T., Franz, B. T., and Driscoll, A. K. (2001). Same-sex romantic attraction and experiences of violence in adolescence. *Am. J. Public Health*. 91, 903–906. doi: 10.2105/AJPH.91.6.903
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. 9(7):671–675. doi: 10.1038/nmeth.2089
- Schultze-Krumbholz, A., and Scheithauer, H. (2009). Bullying and cyberbullying offending among US youth: the influence of six parenting dimensions. *J. Child Family Stud.* 27, 1–20. doi: 10.1027/0044-3409.217.4.224
- Seglow, J. (2016). Hate speech, dignity and self-respect. *Ethical Theory Moral Pract.* 19, 1103–1116. doi: 10.1007/s10677-016-9744-3
- Ševčíková, A., and Šmahel, D. (2009). Online harassment and cyberbullying in the Czech Republic: Comparison across age groups. *J. Psychol.* 217, 227–229. doi: 10.1027/0044-3409.217.4.227
- Shapiro, L. A. S., and Margolin, G. (2014). Growing up wired: Social networking sites and adolescent psychosocial development. *Clin. Child Family Psychol. Rev.* 17, 1–18. doi: 10.1007/s10567-013-0135-1
- Sinclair, K. O., Bauman, S., Poteat, V. P., Koenig, B., and Russell, S. T. (2012). Cyber and bias-based harassment: Associations with academic, substance use, and mental health problems. *J. Adolesc. Health*. 50, 521–523. doi: 10.1016/j.jadohealth.2011.09.009
- Smith, A. (2009). *Radicalization—A Guide for the Perplexed* (National Security Criminal Investigations, Trans.). Canadian: Royal Canadian Mounted Police.
- Sticca, F., Machmutow, K., Stauber, A., Perren, S., Palladino, B. E., Nocentini, A., et al. (2015). The coping with cyberbullying questionnaire: development of a new measure. *Societies*. 5, 515–536. doi: 10.3390/soc5020515
- Tennakoon, H. (2021). Can the 'Dark Triad' traits be a predictor of cyber hate speech. *Academia Lett.* 2. doi: 10.20935/AL1965
- Tynes, B., Giang, M. T., Williams, D. R., and Thompson, G. N. (2008). Online racial discrimination and psychological adjustment among adolescents. *J. Adolesc. Health*. 43, 565–569. doi: 10.1016/j.jadohealth.2008.08.021
- UK Safer Internet Centre. (2016). *Creating a Better Internet for All: Young People's Experiences of Online Empowerment + Online Hate*. Available online at: <https://childnetsic.s3.amazonaws.com/ufiles/SID2016/Creating%20a%20Better%20Internet%20for%20All.pdf> (accessed February 9, 2021)
- Waasdorp, E. T., Hopkins, J., and Bradshaw, C. P. (2018). Examining variation in adolescent bystanders' responses to bullying. *School Psychol. Rev.* 47, 18–33. doi: 10.17105/SPR-2017-0081.V47-1
- Wachs, S., Costello, M., Wright, M. F., Flora, K., Daskalou, V., Maziridou, E., et al. (2021). "DNT LET'EM H8 U!": Applying the routine activity framework to understand cyberhate victimization among adolescents across eight countries. *Comput. Educ.* 160, 104026. doi: 10.1016/j.compedu.2020.104026
- Wachs, S., Gámez-Guadix, M., Wright, M. F., Görzig, A., and Schubarth, W. (2020). How do adolescents cope with cyberhate? Psychometric properties and socio-demographic differences of a coping with cyberhate scale. *Comput. Hum. Behav.* 104, 106167. doi: 10.1016/j.chb.2019.106167
- Wang, J., Iannotti, R. J., and Nansel, T. R. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *J. Adolesc. Health*. 45, 368–375. doi: 10.1016/j.jadohealth.2009.03.021
- Whittaker, E., and Kowalski, R. M. (2015). Cyberbullying via social media. *J. School Violence*. 14, 11–29. doi: 10.1080/15388220.2014.949377

- Williams, T., Connolly, J., Pepler, D., and Craig, W. (2003). Questioning and sexual minority adolescents: High school experiences of bullying, sexual harassment and physical abuse. *Can. J. Commun. Mental Health*. 22, 47–58. doi: 10.7870/cjcmh-2003-0013
- Yang, C., Sharkey, J. D., Reed, L. A., and Dowdy, E. (2020). Cyberbullying victimization and student engagement among adolescents: Does school climate matter? *Sch. Psychol.* 35, 158. doi: 10.1037/spq0000353
- Ybarra, M. L., and Mitchell, K. J. (2004). Online aggressor/targets, aggressors, and targets: a comparison of associated youth characteristics. *J. Child Psychol. Psychiatry*. 45, 1308–1316. doi: 10.1111/j.1469-7610.2004.00328.x
- Ybarra, M. L., Mitchell, K. J., and Korchmaros, J. D. (2011). National trends in exposure to and experiences of violence on the Internet among children. *Pediatrics*. 128, e1376–e1386. doi: 10.1542/peds.2011-0118
- Zaborskis, A., Ilionsky, G., Tesler, R., and Heinz, A. (2018). The association between cyberbullying, school bullying, and suicidality among adolescents. *J. Crisis Interven. Suicide Prevent.* 40, 100–114. doi: 10.1027/0227-5910/a000536
- Zych, I., Farrington, D. P., and Ttofi, M. (2019). Protective factors against bullying and cyberbullying: a systematic review of meta-analyses. *Aggress. Violent Behav.* 45, 4–19. doi: 10.1016/j.avb.2018.06.008

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 Fulantelli, Taibi, Scifo, Schwarze and Eimler. 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:

Douglas F. Kauffman,
Medical University of the
Americas—Nevis, United States

Reviewed by:

Cuicui Sun,
Sichuan Normal University, China
Xueling Yang,
Southern Medical University, China
Ying Zhang,
Capital University of Economics and
Business, China

*Correspondence:

Huang Zuo
zuohuang_6442@163.com
Zehui Zhan
zhanzehui@m.scnu.edu.cn
Lei Mo
molei@m.scnu.edu.cn

†These authors have contributed
equally to this work and share first
authorship

Specialty section:

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

Received: 24 March 2022

Accepted: 26 May 2022

Published: 15 June 2022

Citation:

Gao R, Huang S, Yao Y, Liu X, Zhou Y,
Zhang S, Cai S, Zuo H, Zhan Z and
Mo L (2022) Understanding
Zhongyong Using a Zhongyong
Approach: Re-examining the
Non-linear Relationship Between
Creativity and the Confucian Doctrine
of the Mean.

Front. Psychol. 13:903411.
doi: 10.3389/fpsyg.2022.903411

Understanding Zhongyong Using a Zhongyong Approach: Re-examining the Non-linear Relationship Between Creativity and the Confucian Doctrine of the Mean

Ruixiang Gao^{1,2†}, Shiqi Huang^{3,4†}, Yujie Yao^{1,2†}, Xiaoqin Liu⁴, Yujun Zhou⁵, Shijia Zhang^{1,2}, Shaohua Cai⁶, Huang Zuo^{7*}, Zehui Zhan^{5*} and Lei Mo^{1,2*}

¹ Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, Guangzhou, China, ² School of Psychology, South China Normal University, Guangzhou, China, ³ School of Human-Environment Studies, Kyushu University, Fukuoka, Japan, ⁴ School of Foreign Studies, South China Normal University, Guangzhou, China, ⁵ School of Information Technology in Education, South China Normal University, Guangzhou, China, ⁶ Center for Teacher Development, South China Normal University, Guangzhou, China, ⁷ Institution for Teachers' Professional Ethics and Virtues Building (South China Normal University), Ministry of Education, Guangzhou, China

Zhongyong, a central theme of Confucian thought, refers to the “doctrine of the mean,” or the idea that moderation in all things is the optimal path. Despite considerable interest in the relationship between *zhongyong* and creativity, especially in China, studies of this relationship have not yielded consistent results. Based on a review of the literature, we hypothesized that this inconsistency arises from the dual nature of *zhongyong* itself, which has both a positive side, promoting creativity, and a negative side, inhibiting creativity. We also hypothesized that the negative side of *zhongyong* takes the form of excessive *zhongyong*. Indeed, the observations that every coin has two sides and that too much of a good thing is as bad as too little are core principles of *zhongyong* in traditional Chinese culture. To test these hypotheses, we conducted two empirical studies (measuring explicit and implicit *zhongyong* personality, respectively) to examine the relationships between positive and negative *zhongyong* and creativity (measured in terms of creative personality, divergent thinking, and convergent thinking). The results of both studies revealed an interaction between positive *zhongyong* and negative *zhongyong*, indicating that only a moderate level of *zhongyong* is conducive to creativity; both deficiency and excess are harmful. We discuss the implications of these results, suggesting that a *zhongyong* approach can help to clarify non-linear relationships between things, and recommending to re-assess the creativity of Chinese culture from a neutral and objective outlook. This paper deepens understanding of *zhongyong* and offers clear insights into creativity from an in-depth cultural perspective.

Keywords: *Zhongyong*, traditional Chinese culture, creativity, every coin has two sides, united contradiction perspective, too much of a good thing, implicit association test

INTRODUCTION

The evolution of humankind is a story of creativity. From the creative use and invention of tools to solve survival problems to the use of innovative strategies to obtain developmental advantages, human activities have always been characterized by creativity (Li, 2017). In the twenty-first century, with the rapid growth of material civilization, governments worldwide are attaching increasing importance to the stimulation and cultivation of creativity on a national level. China, one of the world's four oldest civilizations, is no exception. However, despite its early progress in scientific discovery, China was swiftly overtaken by the West in terms of modern scientific and industrial development (the “Joseph Needham puzzle”). Similarly, the “Qian Xuesen’s question”—“why can’t our universities cultivate such outstanding talent?”—reflects many Chinese people’s concern about a lack of creativity in modern China. In recent years, cross-cultural studies have added weight to this concern. For example, Hu et al. (2004) sampled 2,277 Chinese and British teenagers aged 11 to 18 and found that the British teenagers showed significantly higher levels of scientific creativity than the Chinese teenagers did. Zha et al. (2006) studied 55 American and 56 Chinese doctoral candidates and found that although the Chinese participants had significantly better mathematical skills than the American participants, they scored significantly lower for creative potential. The authors argued that the latter result was related to China’s highly collectivist culture. Similar discoveries have been made regarding artistic creativity (Niu and Sternberg, 2001; Yi et al., 2013). Together, these findings suggest that cultural differences affect creative performance.

To study the influence of Chinese culture on creativity, it is necessary to study *zhongyong* (“中庸”). *Zhongyong*, the doctrine of the golden mean, is a core concept in traditional Chinese culture, which originated in Confucian thought. *Zhongyong* refers to the principle of the desirable middle between two extremes. It was once regarded as a virtue of the highest order and shaped Chinese people’s values, beliefs, and mindsets. However, since ancient times, the connotations of the term *zhongyong* have shifted. To many modern Chinese people, *zhongyong* is a derogatory term, synonymous with “having no independent view and fixed position” (折衷主义), “avoiding speaking up even when necessary” (不出头), “seeking peace without principles” (和稀泥), “always saying yes just to avoid offending” (老好人), and “being equivocal” (模棱两可). Many now believe that following the principle of *zhongyong* will lead to “mediocrity” (平庸), “ordinariness” (庸碌), and “the accomplishment of nothing” (无所作为) (Liu, 2019). The shift in understanding of *zhongyong* between ancient and modern times has led to inconsistencies in the conceptualization, operational definition, and measurement of *zhongyong*, which in turn have led to inconsistencies in research findings. To date, studies exploring the relationship between *zhongyong* and creativity have not yielded consistent results: some have found a significant positive correlation, some a significant negative correlation, and others no significant correlation.

We argue that the conceptualization of *zhongyong* in previous studies has not itself followed the doctrine of *zhongyong*,

which holds that we should treat things dialectically. As the saying goes, “every coin has two sides.” Thus *zhongyong* itself cannot be entirely positive; it must also have a negative side. *Zhongyong* also teaches that “too much of something is as bad as too little,” suggesting that its influence on creativity may not be linear. We therefore re-conceptualize *zhongyong* into two categories, positive *zhongyong* and negative *zhongyong*, and conceptualize the latter as an excessive manifestation of the former. To control for social desirability bias in the results, we also adopt for the first time an implicit association test (IAT) to measure *zhongyong*. In the current paper, we present two empirical studies (which measure explicit and implicit *zhongyong* identity, respectively) examining the relationship between positive and negative *zhongyong* and creativity (which is measured as creative personality, divergent thinking, and convergent thinking).

ANALYTICAL FRAMEWORK

Connotations of Zhongyong at the National and Global Levels

Although the Chinese word *zhongyong* (中庸) was coined by Confucius, the concept of *zhongyong* existed long before Confucius; it can be traced back to the Five Classics (五经) of ancient China. The core meaning of the word *zhongyong* lies in its first Chinese character, *zhong* (中), which refers to the broad idea of “correctness.”

At first, *zhongyong* was seen mainly as the moral requirement for a ruler to be fair and just in political management. Of the three virtues advocated in “The Great Plan” of the *Book of Documents* (《尚书·洪范》), correctness/straightforwardness (正直) was listed in the first place. “The Announcement About Drunkenness” in the *Book of Documents* (《尚书·酒诰》) stated that correctness is achieved “when you can maintain a constant, watchful examination of yourselves, and your conduct is in accordance with correct virtue” (克永观省, 作稽中德). This was the first time that correctness had been explicitly deemed a virtue of rectitude and righteousness. Only when a ruler is honest, sincere, and upright without partiality can he govern well (Legge, 1879).

The full word *zhongyong* first appeared in “Yong Ye” of *The Analects* (《论语·雍也》), in which the Master said: “Perfect is the virtue which is according to the Constant Mean! Rare for a long time has been its practice among the people” (中庸之为德也, 其至矣乎! 民鲜久矣). Although Confucius did not define *zhongyong* directly, he exemplified the doctrine with reference to “five excellent things” (五美): “when the person in authority is beneficent without great expenditure; when he lays tasks on the people without their repining; when he pursues what he desires without being covetous; when he maintains a dignified ease without being proud; when he is majestic without being fierce” (君子惠而不费, 劳而不怨, 欲而不贪, 泰而不骄, 威而不猛; in “Yao Yue” of *The Analects* (《论语·尧曰》) (Legge, 1861). Confucius transformed *zhongyong* from being a moral requirement for rulers to being a moral requirement for everyone in society who sought to better themselves (君子),

defining it as behavior that conforms to the requirements of propriety (礼).

Confucius' disciples later compiled "The State of Equilibrium and Harmony" (《中庸》) as a chapter of *The Classic of Rites* (《礼记》), explaining that, "While there are no stirrings of pleasure, anger, sorrow, or joy, the mind may be said to be in the state of Equilibrium. When those feelings have been stirred, and they act in their due degree, there ensues what may be called the state of Harmony. This Equilibrium is the great root from which grow all the human actings in the world, and this Harmony is the universal path which they all should pursue" (喜怒哀乐之未发, 谓之中; 发而皆中节, 谓之和。中也者, 天下之大本也。和也者, 天下之达道也). In other words, *zhongyong* is the achievement of equilibrium and harmony (中和) by holding a middle ground (执中). This means that people's feelings, desires, thoughts, and behaviors should be controlled within a reasonable range, an idea that became the broadest understanding of *zhongyong* (Legge, 1885).

Cheng Yi, a neo-Confucianist of the Northern Song dynasty (A.D. 960–1127), wrote: "What is not extreme is called 'zhong', and what is not changing is called 'yong'" (不偏之谓中, 不易之谓; Zhu, 2012). Master Zhu Xi of the Southern Song dynasty (A. D. 1127–1279) explained that "'Zhong' means impartial and moderate, and 'yong' means constant" (中者, 不偏不倚、无过不及之名。庸, 平常也) in *Variorum of the Four Books* (《四书章句集注》; Zhu, 2012). *Zhong* (中), therefore, has the meanings of moderation, appropriateness, and propriety (which is the valued standard in Confucianism), while "yong" (庸) means holding such a state and remaining unchanged. Therefore, *zhongyong* can also be interpreted as continuously holding the middle ground or always adhering to the appropriate rules of conduct. This is considered to reflect the morality of a superior person (君子).

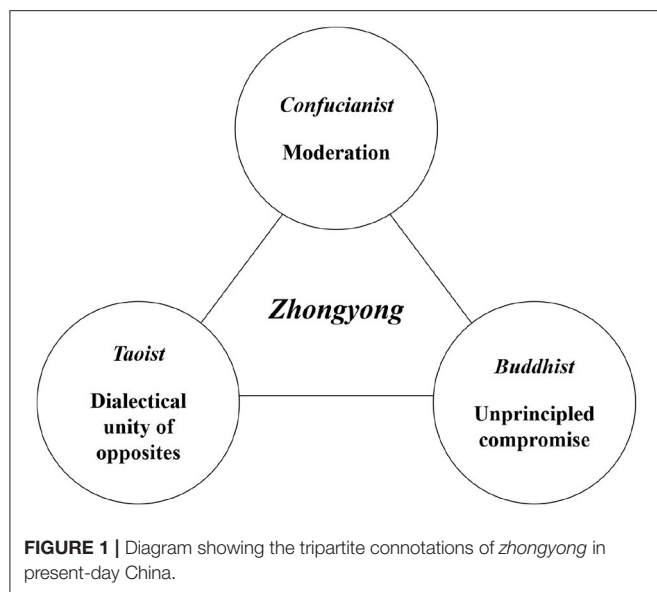
Over time, with the integration of Confucianism, Taoism, and Buddhism, the Confucian concept of *zhongyong* gained new connotations.

The call for moderation in *zhongyong* echoed the concept of the unity of opposites in Taoism, and the connotations of *zhongyong* were expanded to include a perspective on the world in addition to the original moral aspect. Taoism, with its well-known yin–yang diagram (太极图), has an even more profound history than Confucianism, although it was never declared to be the state ideology of ancient China by the emperor. Laozi, the teacher of Confucius, who is said to have composed the *Dao De Jing* (《道德经》), explained the *dao* (道), the law of mutual transformation between yin and yang (阴阳转化), as follows: "it is that existence and non-existence give birth the one to (the idea of) the other; that difficulty and ease produce the one (the idea of) the other; that length and shortness fashion out the one the figure of the other; that (the ideas of) height and lowness arise from the contrast of the one with the other; that the musical notes and tones become harmonious through the relation of one with another; and that being before and behind give the idea of one following another" (有无相生, 难易相成, 长短相较, 高下相倾, 音声相和, ①@Š) (Legge, 1891). This idea can be traced back to the *Book of Changes* (《易经》), particularly its later supplementary chapter "The Great Treatise" (《系辞》),

written by Confucius, which states: "The successive movement of the inactive and active operations constitutes what is called the course (of things). That which ensues as the result (of their movement) is goodness; that which shows it in its completeness is the natures (of men and things)" (一阴一阳之谓道, 继之者善也, 成之者性也) (Legge, 1899). This shows that Confucianism inherited the outlook of dialectical unification from Taoism and folded it into the *zhongyong* moral principle. A *zhongyong* person, therefore, was expected to carefully consider every aspect of a problem when dealing with it and to adopt an objective, neutral standpoint when reconciling different opinions. In "The State of Equilibrium and Harmony" in *The Classic of Rites* (《礼记·中庸》), Confucius provided an example of such a person: "There was Shun: He indeed was greatly wise! Shun loved to question others, and to study their words, though they might be shallow. He concealed what was bad in them and displayed what was good. He took hold of their two extremes, determined the Mean, and employed it in his government of the people. It was by this that he was Shun!" (舜其大知也与! 舜好问而好察迩言, 隐恶而扬善, 执其两端, 用其中于民, 其斯以为舜乎!) (Legge, 1885).

Conversely, the attitude of aloofness in Buddhism lent to *zhongyong* an apparently negative aspect of unprincipled eclecticism that was against Confucius' original intention. In Buddhist thought, escaping the material world is the ultimate goal. According to the "Three Marks of Existence" (Sanskrit *ti-lakkhana*), the basic characteristics of the world are impermanency and constant change (Sanskrit *anicca*). The failure to recognize this, and the practice of clinging to things as if they were permanent, results in dissatisfaction, discomfort, anxiety, frustration, sorrow, pain, suffering, and misery (Sanskrit *duhkha*). The path to Buddhahood is through a careful examination of the constantly changing constituents of a person or an object, by which the practitioner gradually comes to the conclusion that there is no abiding substance in the existence of human beings, other forms of life ("no-self," Sanskrit *anatta*), or non-living things (emptiness, Sanskrit *Sunyata*). The practitioner ultimately reaches a state of liberation from the cycles of rebirth and transcends suffering (*nirvana*; Anderson, 2013). Buddhism's pessimistic outlook on life may have served as comfort for Confucian students who were unsuccessful in their official careers and who either intentionally or unintentionally misinterpreted *zhongyong* as compromise, indifference, or equivocation, even though this deviated from the moral principles promoted by Confucianism. In "Zi Lu" of *The Analects* (《论语·子路》), the Master said: "The superior man is affable, but not adulatory; the mean man is adulatory, but not affable" (君子和而不同, 小人同而不和) (Legge, 1861). According to "Jin Xin II" of *The Works of Mencius* (《孟子·尽心下》), Confucius called these hypocrites *xiangyuan* (乡原), saying that those who blurred the line between right and wrong for the sake of patching up a quarrel were thieves of virtue (德之贼) (Legge, 1985). However, this incorrect interpretation became increasingly popular and is even more popular today.

Today, *zhongyong* has tripartite connotations (see **Figure 1**): (a) moderation, its original meaning, from Confucianism; (b) the dialectical unity of opposites, from Taoism; and (c) a tendency



to compromise, from Buddhism (with the specifically derogatory implication of unprincipled compromise). As the tradition of Confucian propriety (礼) has gradually died out in modern China since the Attack on the Four Olds (破四旧, i.e., old thoughts, old culture, old habits, and old customs) in the Great Cultural Revolution, Confucianist moderation has become increasingly de-emphasized. This is partly attributable to its misinterpretation as unprincipled compromise. Of the three sets of connotations of *zhongyong*, the Taoist conception of *zhongyong* as dialectical thinking is now the most common.

Pang, 1980, 2000 argued that the dialectical unity of opposites provides methodological guidance for achieving moderation in daily practice. Specifically, as a way of thinking about things, *zhongyong* involves two opposite sides of an object, A and B (一体两面, “one body and two sides”). A and B mutually generate and restrict each other (相生相克), representing the contradictory relationship of the unity of opposites. Pang identified four forms of *zhongyong*: (1) “complementing A with B,” or using the opposite B to supplement A’s deficiency; (2) “A but not A,” or removing the negative aspects of A to prevent A from becoming extreme; (3) “neither A nor B,” or being impartial and avoiding too much or too little of anything; and (4) “both A and B,” or the combination (or dynamic balance) at different stages and on different occasions.

The ideology of *zhongyong* is not unique to China. The ancient Greek philosopher Aristotle proposed a doctrine of the golden mean that was identical to the Confucian *zhongyong*. Aristotle argued that virtuous habits of action were often an intermediate condition between two extremes, one of excess and the other of deficiency, and that too much or too little was always wrong. He stated: “virtue must have the quality of aiming at the intermediate. I mean moral virtue; for it is this that is concerned with passions and actions, and in these there is excess, defect, and the intermediate. For instance, both fear and confidence and appetite and anger and pity and in general pleasure and pain may

be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, toward the right people, with the right motive, and in the right way, is what is both intermediate and best, and this is characteristic of virtue” (Aristotle, 1999, p. 27). The philosophy of dialectical materialism, developed by the founders of Communist philosophy Karl Marx and Frederick Engels, conveyed a world outlook similar to that of Taoism and became widespread in modern China. “The central idea of dialectics is the unity and struggle of opposites, that is, contradictory tendencies that are tied together and cause things to change and develop ... [T]hey (Marx and Engels) did not start from scratch. They borrowed ideas from a long history of dialectical thought that dates back at least 25 centuries in Europe and was developed independently in China and India.” The Chinese sources of this dialectical thought were the *Book of Changes* (《易经》), the *Dao De Jing* (《道德经》), and *Mohism* (《墨子》); International Communist Workers Party (ICWP), 2013). However, although the West developed philosophies similar to *zhongyong*, in Chinese culture this philosophy became deeply embedded in value systems, ways of thinking, and daily conduct. This may be because the Chinese “virtue-oriented” educational model focuses more on perfecting the self *via* social ideology, while the Western model emphasizes exploring the outside world (Li, 2012; Gao et al., 2022).

Measurement of *Zhongyong* in Previous Research

The systematic psychological study of *zhongyong* began in the late 1990s, with research by Yang and her colleagues. These scholars attempted for the first time to conceptualize *zhongyong* in psychological research, proposing that it was a practical thinking system that people used to decide how to choose, execute, and correct specific action plans when dealing with routine actions. Their system consisted of eight main sub-constructs: the unity of humans and nature (天人合一), bipolar thinking (两极思维), consequential thinking (后果思维), waiting to see what happens (静观其变), not going to extremes (不走极端), considering the overall situation (顾全大局), being reasonable (合情合理), and retreating to advance (以退为进). Based on this conceptualization, they developed a Zhong-Yong Practice Cognition Scale containing 16 forced-choice items (Yang and Chiu, 1997). Since then, Yang has continued to develop the psychological conceptualization of *zhongyong* into a cultural meaning system that can enter into dialog with Western social psychology (Yang, 2008). Yang constructed a conceptualization diagram of the *zhongyong* practice cognition system (CDZPCS) as a blueprint for the study of the role of *zhongyong* in Chinese people’s lives (Yang, 2009). She also provided a research roadmap based on this blueprint (Yang, 2010) and applied traditional psychological research methods to test the construct validity of the CDZPCS (Yang and Lin, 2012; Yang et al., 2014).

However, Yang’s system was argued to be overcomplicated and was later revised by her collaborators. Chiu (2000) concretized *zhongyong* thinking at the level of action and divided it into three dimensions: “taking harmony as the action goal” (以“和为行动目标”), “recognizing the complex interrelationships between

things” (认清事物复杂的相互关系), and “carrying out actions with reference to a middle ground” (以“执中”开展行动). Based on this reconceptualization, the original scale was reduced to 14 items and then tested in five Chinese communities. However, this new scale was questioned in relation to its quantitative methods, reliability, and validity (Wu and Lin, 2005; Huang et al., 2012). Huang et al. (2012) revised the 16-item scale into a 9-item Zhong-Yong Belief-Value Scale (ZYBV), including two dimensions of self-convergence (自我收敛) and vision elevation (拔高视野). This version has been widely used in research. Other researchers have focused on the basic meaning of *zhongyong*, that is, “master the extremes, but deploy the mean” (执两端而允中). In other words, when dealing with a controversial issue, one should consider a range of perspectives in detail and make decisions that take into account both the overall situation and the self. Huang et al. (2012) also established the Zhong-Yong Thinking Style Scale (ZYTS), which featured three dimensions: “multi-dimensional thinking” (多方思考), “holism” (整合性), and “harmoniousness” (和谐性). This scale is currently one of the most widely used tools for measuring *zhongyong* (Wu and Lin, 2005). Du and Yao (2015) argued that the scales designed by Chiu (2000) and Wu and Lin (2005) were based on ideas of *zhongyong* extracted from the theoretical literature and therefore belonged to “classic *zhongyong*.” They argued that because the concept of *zhongyong* has changed over time, these scales did not measure *zhongyong* as it was perceived and applied by Chinese people today. They therefore investigated Chinese enterprise employees’ perceptions of *zhongyong* using an open questionnaire, then proposed four dimensions of *zhongyong*: “mean and congruence” (执中一致), “personal cultivation” (慎独自修), “no ambition” (消极成就), and “passive avoidance” (消极规避). They identified mean and congruence as the core and went on to investigate the relationship between mean and congruence and collectivism.

In addition to using questionnaires for the static measurement of *zhongyong* as a relatively stable personality trait or thinking habit, some researchers have used other research paradigms, such as situational question priming, to investigate *zhongyong*. Studies have shown that the thinking of East Asians is more context-dependent than that of Westerners (e.g., Ji et al., 2000), and adaptability is also an important feature of *zhongyong*. Based on the four forms of *zhongyong* proposed by Pang (1980), Zhou et al. (2019), Zhang et al. (2020) summarized two main forms of *zhongyong* thinking: eclectic thinking and integrated thinking. “Eclectic thinking” refers to the cognition of individuals who fail to recognize the essential problems underlying contradictory information, leading them to compromise to alleviate superficial or temporary contradictions, which ultimately makes it difficult to solve the problems. This type of thinking occurs in two situations: either a problem is beyond the ability of the solver or the solver avoids making cognitive effort due to laziness. “Integrated thinking” refers to the ability to understand the essence of a problem and synthesize seemingly contradictory information to solve the problem. Integrated thinking takes place at a higher level than eclectic thinking. On this basis, Zhou et al. (2019), Zhang et al. (2020) pioneered a causal experimental approach to studying *zhongyong* by developing

situational problem materials to prime eclectic thinking and integrated thinking. Studies have incorporated Western ideas about paradoxical thinking (Miron-Spektor et al., 2011; Leung et al., 2018), a concept similar to *zhongyong* thinking.

Relationship Between Zhongyong and Creativity

Creativity is a multidimensional construct that captures the ability of an individual to solve a problem in a novel way (Jiao et al., 2017; Li et al., 2017; Lin et al., 2018). The measurement of creativity typically reflects one of two definitions of creativity: the first is as a personality trait, and is usually measured by self-report questionnaires; while the second is as a set of cognitive capabilities (e.g., divergent thinking and convergent thinking), and is usually measured by ability tests with differing degrees of difficulty. *Zhongyong*, as a cultural factor that impacts creativity, has received increasing research interest; however, studies examining the relationship between *zhongyong* and creativity have not yielded consistent results. Several studies (Zhang and Gu, 2015; Yang and Zhang, 2018) have found that the *zhongyong* thinking of employees in enterprises is positively correlated with their innovative behaviors. *Zhongyong* thinking has also been shown to positively predict individual (Liao and Dong, 2015) and team (Chen et al., 2018) innovative behaviors. Some studies, however, have found that *zhongyong* might hinder innovation. For example, a negative correlation has been found between the ZYTS scores of Chinese art students and their creative personality scores (Liu et al., 2015). Other studies have found that the relationship between the two is not linear. For example, Yao et al. (2010) found that scores on the ZYTS scale moderated the relationship between self-evaluated creativity and leader-evaluated innovative behaviors: there was no significant correlation between the two in the high-*zhongyong* group of participants, but there was a significant correlation in the low-*zhongyong* group, implying that *zhongyong* hindered the transformation of creative ideas into innovative action. Du et al. (2018) found that *zhongyong* value orientation promoted incremental innovation but inhibited radical innovation. All of these studies used questionnaires to measure creativity and *zhongyong* thinking. Most of the studies that found negative or no correlations between *zhongyong* and creativity (e.g., Yao et al., 2010; Liu et al., 2015) used the ZYBV scale, while most of those that found a positive correlation (e.g., Liao and Dong, 2015; Zhang and Gu, 2015) used the ZYTS scale. We suggest that this is because the ZYBV scale uses a forced-choice method, making it less prone to social desirability bias.

Other researchers have used non-questionnaire methods to measure creativity. Chang and Yang (2014) used participants’ performance on a redundant-target detection task as indicators of creativity, finding that high *zhongyong* thinkers, as identified using the ZYTS scale, processed information more efficiently and in a more integrated fashion than low *zhongyong* thinkers. Similarly, Wang et al. (2013) used participants’ eye tracking performance when viewing banner advertisements as indicators of creativity, finding that high-*zhongyong* thinkers, as identified using the ZYBV scale, exhibited a more efficient and flexible

perceptual style when switching between global processing and local processing. Huang et al. (2014) further found that when primed with emotional words, the high-*zhongyong* group of participants, as selected using the ZYBV scale, showed significantly more global precedence (i.e., stepping back to see the whole picture). However, when the priming was absent, there was no reliable relationship between *zhongyong* and global processing speed. This implied that *zhongyong* served as an emotional regulator that affected individuals' cognitive processing strategies, affirming Confucius' statement that "while there are no stirrings of pleasure, anger, sorrow, or joy, the mind may be said to be in the state of Equilibrium" (喜怒哀乐之未发, 谓之中). Recently, researchers used participants' performance on a divergent thinking test (the Alternative Uses Task, AUT), a convergent thinking test (the Remote Associates Test, RAT), and insight problem-solving tests (Chinese idiom puzzle problems, brain-teaser problems, and market investment problems) as indicators of creativity and found that there was no significant correlation between scores on the ZYTS and ZYBV scales and these indicators (Zhang et al., 2020). When the participants were primed with a *zhongyong* conditional problem, those primed with integrated thinking performed better in the RAT and the market investment problems than those primed with eclectic thinking and the control group (Zhang et al., 2020; Zhou et al., 2021). The improved RAT performance was supported by EEG data (Zhou et al., 2019), suggesting that the RAT and the integrated thinking priming tasks involve the same neural mechanism. Researchers examining the Western counterpart of *zhongyong*, paradoxical thinking, have typically viewed it as a mental template for approaching contradictory yet interrelated elements to enable change and gain new insights (Gordon, 1961; Fletcher and Olwyler, 1997; Lewis, 2000; Martin, 2009; Ingram et al., 2016; Miron-Spektor and Erez, 2017). Miron-Spektor et al. (2011) found that priming paradoxical frames promoted participants' creative thinking. They argued that this was because the sense of conflict caused by the paradoxical relationship led to a willingness to embrace different perspectives and to integrate these different perspectives by generating new linkages among them, thus promoting creativity. However, Leung et al. (2018) found that people who endorsed a middle-ground approach were less likely to find integrative solutions and thus received fewer of the creative benefits of paradoxical frames.

Based on the above review, it is evident that there is both strong support for and considerable doubt about the idea that *zhongyong* promotes creativity. The main argument supporting the promotion of creativity is that a person who practices integrative thinking is better at viewing problems from a global perspective and adopting flexible strategies to integrate different or even contradictory opinions, and thus is more able to produce new ideas that bridge differences and achieve harmony [see Chang et al., 2014 for a detailed review of the relationship between *zhongyong* and the six dimensions of wisdom, as per Grossmann et al. (2010, 2013): compromise, recognition of the limits of knowledge, flexibility, perspective-taking, recognition of change, and resolution of conflict]. The primary argument in favor of *zhongyong* hindering creativity is that a *zhongyong* person with middle-ground, eclectic thinking tends to avoid

conflict and seek interpersonal harmony, thus compromising easily without challenging authority or social norms. This is a typical argument for the view that collectivist cultures inhibit creativity (Hofstede et al., 2010). While researchers have been puzzled by these seemingly contradictory results, we believe that the results precisely reflect the essence of *zhongyong*. The dialectical thought of *zhongyong* tells us that there is no absolute good or bad and that everything has a positive and a negative side. *Zhongyong* itself is no exception. The positive side of *zhongyong*, integrative thinking, is conducive to creativity, while the negative side, eclecticism, is harmful to creativity. The inconsistent results of previous studies have arisen from a confusion of positive *zhongyong* with negative *zhongyong*—simply reflecting the ability of *zhongyong* to unify contradictory things. With further analysis, we also argue that the pursuit of *zhongyong* is characterized by moderation, while the negative aspects of *zhongyong*, such as unprincipled compromise, arise from an excessive amount of *zhongyong*, indeed. For example, making a decision without considering different points of view is inadvisable, but if someone is *too* cautious and hesitant when combining viewpoints to make a decision, he/she is criticized for exercising eclectic thinking rather than praised as a dialectic thinker. Therefore, to understand the research results of *zhongyong*, we must therefore adopt *zhongyong* thinking.

The Current Studies

To testify the above assumptions, we first distinguished between positive *zhongyong* and negative *zhongyong*. To define and describe the two, we used 12 Chinese idiomatic expressions, six for each. For positive *zhongyong*, the expressions were as follows: "taking the whole situation into account" (顾全大局), "cherishing peace and harmony" (ÔÔ°ÏÏ¹Ö), "properly following rules for advancing and retreating" (进退有度), "knowing when to bend or stand upright" (能屈能伸), "leaving some leeway" (留有余地), and "being impartial" (不偏不倚). For negative *zhongyong*, the idiomatic expressions were as follows: "being worldly-wise to avoid getting into trouble" (明哲保身), "being a yes-man" (好好先生), "flattering both sides" (两面讨好), "swaying both ways" (左右逢源), "making concessions against one's will for a semblance of peace" (委曲求全), and "being untrustworthy" (生性圆滑). To select these 12 Chinese idioms, three experts in the authorship first brainstormed two initial lists of *zhongyong*-related expressions, one for positive *zhongyong* and the other for negative *zhongyong*. Then, another five experts were invited to supplement new items to the lists, which resulted in 18 expressions for positive *zhongyong* and 14 for negative *zhongyong*. Next, 95 university students were recruited to evaluate the representativeness, familiarity, and semantic valence (positive or negative) of these 32 expressions. Eventually, based on the evaluation, two experts in the authorship decided that six idioms for positive *zhongyong* and six for negative *zhongyong* were selected as the final expressions.

Based on these definitions, we conducted two empirical studies to examine the relationship between positive and negative *zhongyong* and creativity (creative personality, divergent thinking, and convergent thinking). As the descriptive terms for negative *zhongyong* were pejorative, in Study 2 we adopted an

implicit association test (IAT) to minimize social desirability bias. The IAT, introduced by Greenwald et al. (1998), is a measure widely used in social psychology to detect attitudes and beliefs that people may not be willing or able to report. The instrument measures each participant's reaction time and accuracy rate, assuming that they are reflective of automatic associations between mental representations. As positive *zhongyong* and negative *zhongyong* might partially overlap because of our operational definitions, we expected to see an interaction between the two. We predicted that when a participant's level of negative *zhongyong* was low, which means that overall *zhongyong* had not reached an extreme level, positive *zhongyong* would positively predict creativity; that is, *zhongyong* is conducive to creativity as long as it is not excessive. However, when their level of negative *zhongyong* was high, which means that *zhongyong* overall had reached an extreme level, positive *zhongyong* would have either no correlation or a negative correlation with creativity; in other words, too much *zhongyong* is detrimental to creativity. If these two parts of hypothesis could be verified, we could reach a conclusion that only a moderate level of *zhongyong* is beneficial for creativity.

STUDY 1

In Study 1, we developed and used a positive and negative *zhongyong* personality scale to explore the non-linear relationship between positive and negative *zhongyong* and creativity (creative personality, divergent thinking, and convergent thinking).

Participants and Procedure

Before recruiting the participants, we used G*Power 3.1 to estimate the minimum sample size suitable for interaction analysis in a three-predictor regression model. According to Faul et al. (2007), when setting the effect size (f^2) at 0.15 (medium level), the error probability (α) at 0.05 (as a common practice), the power ($1 - \beta$) at 0.8 (as a common practice), the number of tested predictors at 1, and the total number of predictors at 3, the calculated minimum sample size was 55.

In reality, one hundred and fourteen undergraduate students (27 male, 87 female) aged between 18 and 22 years ($M = 19.23$, $SD = 0.99$) from two classes of a university-wide selective course took part in the study during a classroom-based session. None of the participants had experience of taking tests similar to those involved in our study. After they consented to participate, the students were asked to complete the AUT, the RAT, the creative personality scale, the ZYTS scale, the ZYBV scale, and our *zhongyong* personality scale (in that order) using their mobile telephones. The instruments were administered via Wenjuanxing, a Chinese online questionnaire survey platform. The total duration for all instruments was ~25 mins.

Methods

Measuring Creativity

Three types of tasks were used to measure creativity: a Chinese version of the Williams Scale, an Alternative Uses Task (AUT, for

measuring divergent thinking) drawn from the Torrance Tests of Creative Thinking, and a Chinese Remote Associates Test (RAT, for measuring convergent thinking test) that we compiled for the purposes of this study. The first instrument, the Chinese Williams Scale, was used to measure creative personality. It consists of four dimensions and 50 items (8 reverse-scored): 14 for curiosity, 13 for imagination, 12 for complexity, and 11 for risk-taking (Lin and Wang, 1999). In the current sample, the 3-point Likert scale we used had a Cronbach's α of 0.88. The second instrument, the AUT, was used to measure divergent thinking, or the ability to think of solutions to a problem from various angles. In this test, participants are given 5 mins and asked to list as many ways as possible that a common item (a cardboard box) could be used (Torrance, 1966). The participants' answers were scored according to three criteria, namely fluency, flexibility, and originality, with their overall score calculated as the mean of the three. This scale has already been shown to have high inter-rater reliability. In a random sample of 89 in another study we conducted, two coders achieved inter-rater reliability values (calculated using a Pearson correlation) of 0.99, 0.95, and 0.89 for each of the three components, respectively. In the current research (both Studies 1 and 2), the answers were scored by one of these two coders. The third instrument, the Chinese RAT, was used to measure convergent thinking, or the ability to apply established rules and logical reasoning to narrow down the possible solutions to a problem. Our instrument was a modified version of Wu and Chen (2017) instrument and contained 50 items suitable for college students. Each item comprised three clue characters, e.g., “原” (plain), “鞋” (shoe), and “野” (wild). The participants had 10 secs to come up with the answer, which was a target character, e.g., “草” (grass), that had a semantic connection with all three clues and created three actual two-character words, e.g., “草原” (grassland), “草鞋” (straw sandal), and “野草” (weed). The pass rate for this instrument was 0.23.

Measuring Zhongyong

Three scales were used to assess *zhongyong*: the 13-item ZYTS scale (Wu and Lin, 2005), the nine-item ZYBV scale (Huang et al., 2012), and the *zhongyong* personality scale that we compiled for the purposes of this study. The ZYTS scale used a 7-point Likert scale, with 7 indicating “extremely like me,” and had no reverse-scored items. The ZYBV scale presented the participants with two contradictory statements and prompted them to choose the one they agreed with. The participants then evaluated the degree to which they agreed with the statement on a 7-point Likert scale, with 7 indicating complete agreement. If the participant chose a non-*zhongyong* statement, the score of the corresponding item was reversed. These two scales were used as criteria in our study and had Cronbach's α values of 0.91 and 0.61, respectively. The *zhongyong* personality scale that we compiled was based on Gough (1979) Creative Personality Scale. This instrument consisted of an adjective checklist that the participants marked to indicate whether each adjective described them well. The list consisted of 32 adjectives, 6 positive *zhongyong* terms, 6 negative *zhongyong* terms (see section 2.4), and 20 terms not relevant to *zhongyong* (such as “hardworking” and “careful”) that were not

TABLE 1 | Descriptive statistics and correlation coefficients for the six measures.

	<i>M</i>	<i>SD</i>	Positive <i>zhongyong</i>	Negative <i>zhongyong</i>	ZYTS	ZYBV	Creative personality	AUT	RAT
Positive <i>zhongyong</i>	5.04	1.24	1						
Negative <i>zhongyong</i>	2.60	1.61	0.27**	1					
ZYTS	5.71	0.76	0.45**	0.14	1				
ZYBV	5.18	0.69	0.39**	0.15	0.28**	1			
Creative personality	2.17	0.24	−0.12	−0.23*	0.14	−0.01	1		
AUT	23.70	9.70	−0.03	0.06	0.25**	−0.08	0.18	1	
RAT	11.67	3.57	0.24**	−0.06	0.17	0.30**	0.06	0.06	1

* $p < 0.05$, ** $p < 0.01$.

used in this study. The participants scored between 0 and 6 for both positive and negative *zhongyong*.

Results

Descriptive statistics and Pearson correlation coefficients for the participants' scores on the six tasks and scales of interest are shown in **Table 1**. Positive *zhongyong* had a positive and significant correlation with negative *zhongyong*, ZYTS, and ZYBV, while negative *zhongyong* had no significant correlation with ZYTS or ZYBV. This supported our argument that the instruments previously used to assess *zhongyong* have considered only its positive side, neglecting its negative side. Positive *zhongyong* had a positive and significant correlation with RAT, while negative *zhongyong* had a negative and significant correlation with creative personality. This supported our argument that *zhongyong* has two opposite sides, one promoting creativity and the other inhibiting it.

Table 2 displays the frequency distributions of positive *zhongyong* and negative *zhongyong*, dividing scores into two groups: [0,3] and [3,6]. We can see that the majority of the participants scored high on positive *zhongyong* and low on negative *zhongyong*. Only one participant scored high on negative *zhongyong* and low on positive *zhongyong*, which suggests that someone with a negative *zhongyong* personality must first have developed a positive *zhongyong* personality, and therefore, supports our argument that negative *zhongyong* is an excessive form of *zhongyong*.

As positive *zhongyong* and negative *zhongyong* are positively correlated, we expected to see an interaction in their effects on creativity. Specifically, we expected negative *zhongyong* to moderate the correlation between positive *zhongyong* and creativity. Therefore, after controlling common method bias by Harman's single-factor test (the percentage of variance for the first common factor was 28.87% <40%) and centralizing the data, we computed three regression models, one each for creative personality, the AUT, and the RAT, using positive *zhongyong*, negative *zhongyong*, and their product term as independent variables (see **Table 3**). However, we found only a marginally significant ($p = 0.07$) interaction in the regression model for the RAT. We then conducted a simple slope test (see **Figure 2**) to examine the moderating effect of negative *zhongyong* and found

TABLE 2 | Cross-tabulation of frequency distributions for positive and negative *zhongyong*.

		Negative <i>zhongyong</i>		Total
		[0,3]	[3,6]	
Positive <i>zhongyong</i>	[0,3]	14	1	15
	[3,6]	69	30	99
Total		83	31	114

that the resulting tendency was in line with our expectation that at low levels of negative *zhongyong*, positive *zhongyong* would positively predict creativity, but at high levels of negative *zhongyong*, positive *zhongyong* would negatively predict or fail to predict creativity.

DISCUSSION

The results of Study 1 lent some support for our expected relationship but did not reach a statistically significant level. A possible reason for this is that the *zhongyong* personality scale that we compiled may not have revealed the participants' true *zhongyong* personality because of social desirability bias. For this reason, we used an implicit method to measure *zhongyong* in Study 2.

STUDY 2

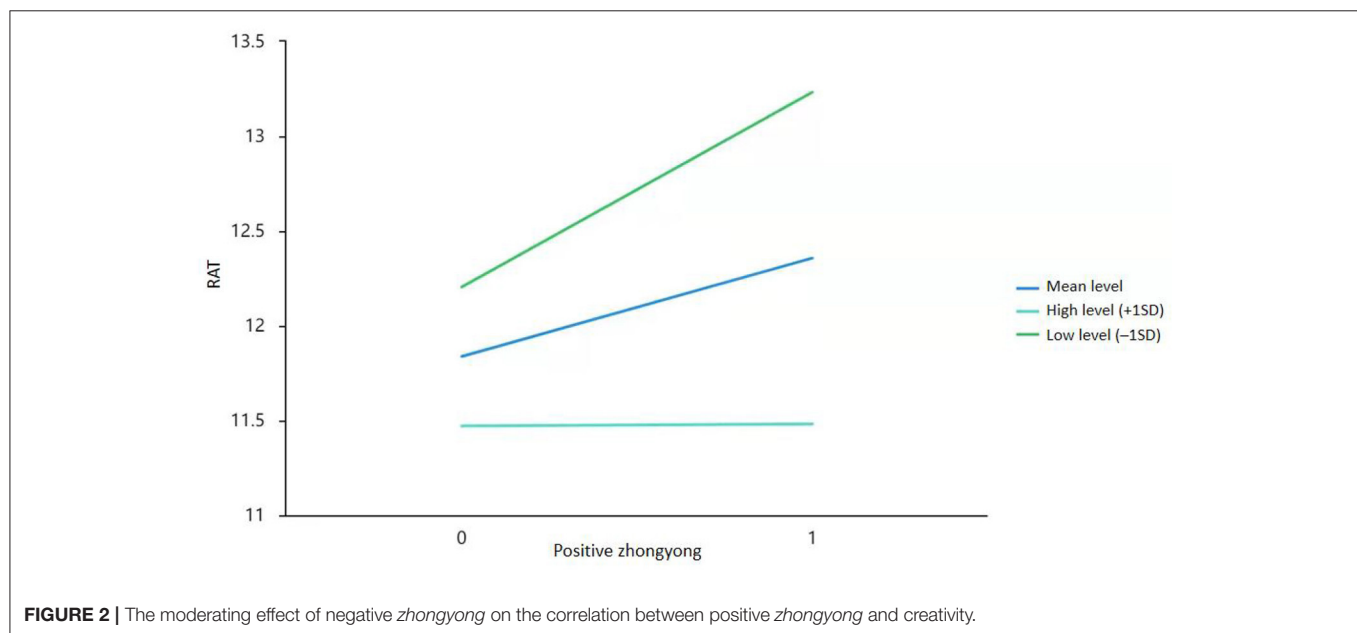
To minimize social desirability bias, we used the IAT method in Study 2 to assess implicit *zhongyong* personality and to check the findings of Study 1.

Participants and Procedure

Before recruiting the participants, we again used G*Power 3.1 to estimate the minimum sample size suitable for a 2×2 between-subjects *F* test. According to Faul et al. (2007), when setting the effect size (f) at 0.25 (medium level), the error probability (α) at 0.05 (as a common practice), the power ($1 - \beta$) at 0.8 (as a common practice), the number of groups at 4, and the degree of freedom at 1, the calculated minimum sample size was 128.

TABLE 3 | Regression models for creative personality, AUT, and RAT on positive and negative *zhongyong* (including their interaction).

	Creative personality			AUT			RAT		
	β	t	p	β	t	p	β	t	p
Positive <i>zhongyong</i>	−0.003	−0.027	0.978	−0.103	−0.911	0.364	0.181	1.668	0.098
Negative <i>zhongyong</i>	−0.231	−2.368	0.020*	0.087	0.871	0.386	−0.103	−1.077	0.284
Interaction	0.110	1.115	0.267	−0.100	−0.988	0.325	−0.177	−1.829	0.070

* $p < 0.05$.

In reality, another 144 undergraduate students (30 male, 114 female) aged between 18 and 28 ($M = 19.62$, $SD = 1.49$) from three classes of a university-wide selective course took part in the study in a classroom-based session. None of the participants had experience of taking tests similar to those used in our study. After they had consented to participate, the students were asked to complete the AUT, the RAT, and the creative personality scale in sequence using their mobile telephones on Wenjuanxing, a Chinese online questionnaire survey platform. They were asked to complete the *zhongyong* IAT using their mobile telephones on DiggMind, a behavioral experiment platform similar to E-prime but suitable for mobile devices. The total duration of these tests was ~30 mins.

Methods

The AUT, the RAT, and the creative personality scale used in this study were identical to those used in Study 1. The *zhongyong* IAT was a modification of the self-esteem IAT (Greenwald and Farnham, 2000). During the first round, words denoting the concept of “self” (such as “me,” “my,” and “myself”) and the six positive *zhongyong* words used the same response button, while words associated with the concept of “other” (using words such as “them,” “their,” and “themselves”) and the six negative *zhongyong* words shared a different response button. Following

this, the categorization task was reversed: the self was grouped with negative *zhongyong* words, and the other with positive *zhongyong* words (see **Table 4** for the complete procedure). In the case of incorrect responses, 600 ms of reaction time was added. If the self–positive pair in Block 4 took less time than the other–positive pair in Block 7, this was taken to indicate that the participant was implicitly demonstrating a positive *zhongyong* personality; if the self–negative pair in Block 7 took less time than the other–negative pair in Block 4, this was taken to indicate that the participant was implicitly demonstrating a negative *zhongyong* personality.

Results

Table 5 shows a cross-tabulation of the distribution of positive and negative *zhongyong* personality in the participants, as measured by the IAT. The results were consistent with those of Study 1: the majority of the participants demonstrated a positive *zhongyong* personality but not a negative *zhongyong* personality, while few of the participants possessed a negative *zhongyong* personality without also displaying a positive *zhongyong* personality. This again supports our conception of negative *zhongyong* as an excessive form of *zhongyong*, because only when one has first developed a positive *zhongyong* personality, can

TABLE 4 | Procedure for conducting the *zhongyong* IAT.

Order of block	Number of trials	Function	Left button (number of trials)	Right button (number of trials)
1	6	Practicing	Positive (3)	Negative (3)
2	6	Practicing	Self (3)	Other (3)
3	12	Practicing	Self (3) + positive (3)	Other (3) + negative (3)
4	24	Formal	Self (6) + positive (6)	Other (6) + negative (6)
5	6	Practicing	Negative (3)	Positive (3)
6	12	Practicing	Self (3) + negative (3)	Other (3) + positive (3)
7	24	Formal	Self (6) + negative (6)	Other (6) + positive (6)

TABLE 5 | The distribution of the number of participants who demonstrated implicit positive and negative *zhongyong* personality, as measured by the IAT.

		Implicit negative <i>zhongyong</i>		Total
		No	Yes	
Implicit positive <i>zhongyong</i>	No	12	2	14
	Yes	113	17	130
Total		125	19	144

their *zhongyong* personality further reach the excessive amount manifesting in the form of negative *zhongyong*.

After homogeneity of variance was tested (p -values 0.756, 0.352, and 0.934 for creative personality, AUT, and RAT, respectively), three 2 (positive *zhongyong*: no/yes) \times 2 (negative *zhongyong*: no/yes) non-repeated ANOVAs were computed for creative personality, AUT, and RAT. The results revealed two significant interactions between positive *zhongyong* and negative *zhongyong* in relation to creative personality and the RAT (see **Table 6**). We then analyzed the simple effects for both interactions (see **Figure 3**). This analysis further supported our prediction that for participants who did not have an implicit negative *zhongyong* personality, their implicit positive *zhongyong* personality would positively predict their creativity, while for participants with an implicit negative *zhongyong* personality, their positive *zhongyong* personality would negatively predict their creativity.

Discussion

Based on the results of Studies 1 and 2, we can conclude that negative *zhongyong* is an excessive form of *zhongyong* that inhibits creativity. We can also conclude that only a moderate level of *zhongyong* is conducive to creativity, with both deficiency and excess being harmful. This interaction effect was found for creative personality and the RAT but not for the AUT, which is consistent with previous findings (Zhou et al., 2019, 2021; Zhang et al., 2020). This implies that the RAT and *zhongyong* thinking have a similar cognitive mechanism, but the AUT does not. Zhu et al. (2019) found a threshold-setting effect of convergent thinking; that is, only when convergent thinking capacity reached a certain level did divergent thinking begin to play a role in

scientific creativity. This suggests that *zhongyong* thinking may be important to scientific creativity.

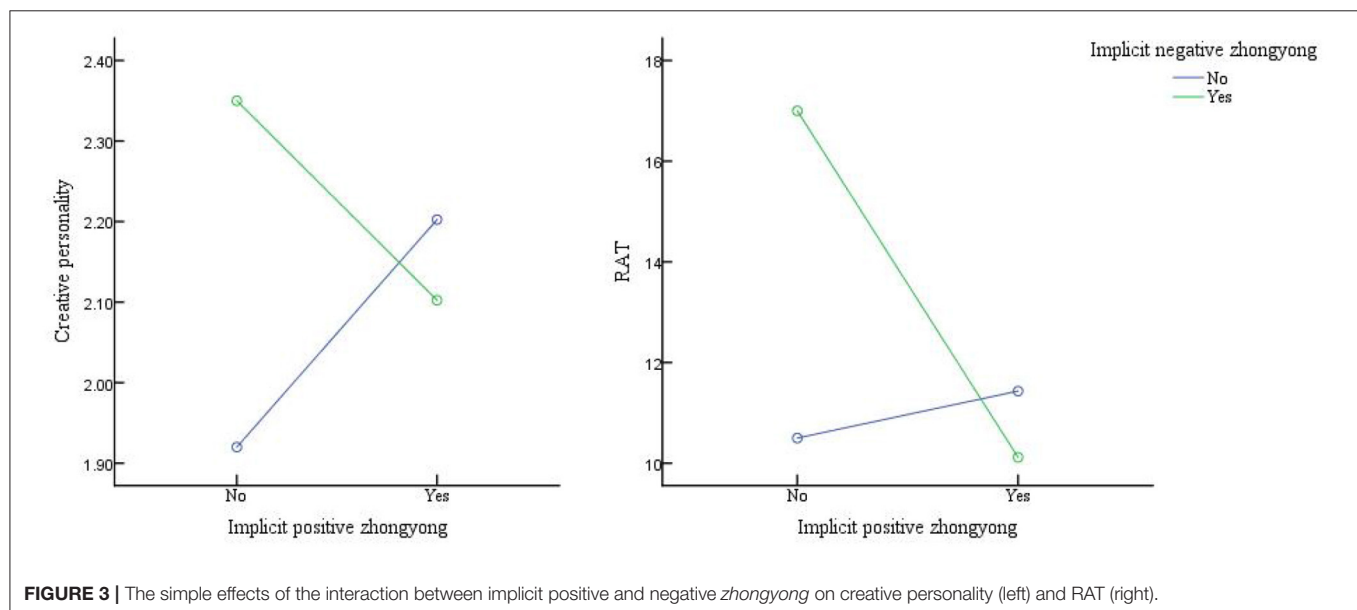
GENERAL DISCUSSION

In this paper we present the results of two empirical studies conducted to examine the non-linear relationship between *zhongyong* and creativity. Based on these studies, we conclude that excessive *zhongyong* is detrimental to creativity and moderate *zhongyong* is conducive to creativity. These findings shed light on the previously inconsistent findings regarding the relationship between *zhongyong* and creativity.

Our study has two major implications. First, our work highlights the possibility of non-linear relationships between constructs. Both *zhongyong* in China and the doctrine of the mean in Europe emphasize moderation, which means that the relationship between positive antecedents and ideal consequences is not necessarily monotonic and that the “more is better” attitude may be misguided. The “too much of a good thing” effect, i.e., the inverse U-shaped relationship, has aroused more and more attention. In the field of psychology, this effect has been observed in relation to individual personality traits (e.g., Bozionelos et al., 2014; Nieß and Biemann, 2014; Vergauwe et al., 2018), skills (e.g., Zettler and Lang, 2015), and demographic variables such as age (e.g., von den Driesch et al., 2015) and family socioeconomic status (e.g., Ren and Xin, 2013). In the fields of economics and management, researchers have also found this effect in resource ownership (e.g., Rotolo and Messeni Petruzzelli, 2013; Shao et al., 2013; Ren and Chadee, 2017; Fisman et al., 2020), positive and negative work experience (e.g., Carette et al., 2013; Lee et al., 2013; Stouten et al., 2013; Lam et al., 2014; Rapp et al., 2014; Astakhova, 2015; Burnett et al., 2015; Zhang and Long, 2016; Mo et al., 2019), employee autonomy (e.g., Lee et al., 2017), emotional expression rules (e.g., Christoforou and Ashforth, 2015), and group diversity (e.g., Ali et al., 2014; Wei et al., 2015; Vicentini and Boccadelli, 2016; Dayan et al., 2017). These studies adopted a new perspective of curvilinear relationships or interaction effects to revisit debates in previously published literature, gaining new insights either (a) by identifying an inflection point after which the positive effects turned negative as a result of breaking up the balance between gains and losses, or (b) by introducing a new factor as a moderator, where the product of the independent variable and the moderating variable

TABLE 6 | The results of analyses of variance for the effects of positive and negative *zhongyong* as well as their interaction on creative personality, AUT, and RAT.

	Creative personality				AUT				RAT			
	<i>F</i>	<i>df</i>	<i>p</i>	η_p^2	<i>F</i>	<i>df</i>	<i>p</i>	η_p^2	<i>F</i>	<i>df</i>	<i>p</i>	η_p^2
Positive <i>zhongyong</i>	0.03	1	0.858	0.00	0.97	1	0.326	0.01	5.13	1	0.025*	0.04
Negative <i>zhongyong</i>	2.90	1	0.091	0.02	2.04	1	0.155	0.01	3.89	1	0.050*	0.03
Interaction	7.49	1	0.007**	0.05	0.92	1	0.340	0.01	8.85	1	0.003**	0.06

p* < 0.05; *p* < 0.01.**FIGURE 3 |** The simple effects of the interaction between implicit positive and negative *zhongyong* on creative personality (left) and RAT (right).

was instrumental in influencing the dependent variable (for reviews, see Grant and Schwartz, 2011; Pierce and Aguinis, 2013; Haans et al., 2016; Xing et al., 2018). The study of *zhongyong* itself is no exception. The inconsistency in previous findings occurred because scholars did not adopt *zhongyong* thinking; they failed to use a *zhongyong* approach to understand *zhongyong* itself. The present study helps us to understand the positive and negative sides of *zhongyong* and to realize that a moderate level of *zhongyong* is conducive to stimulating creativity, while too much or too little is useless.

Second, we argue for a reconceptualization of creativity in Chinese culture. In the past, Chinese people have been labeled “not creative,” partly because the modern industrial and scientific revolutions did not originate in China. Many scholars have tried to determine the cultural reasons for this lack of creativity, writing books with titles such as *Why Asians Are Less Creative Than Westerners* (Ng, 2001) and *Liberating the Creative Spirit in Asian Students* (Ng, 2004). These books have proposed that *zhongyong* thinking, collectivism, hierarchy, obedience to authority, self-inhibition, and mechanical learning in Confucian cultures hinder the development of creativity. Sometimes even Chinese people themselves are not confident in their creativity. They may question, criticize, or even completely deny

long-established Confucian ideas and collectivist values in Chinese traditional culture, believing that they inhibit the development of creativity. However, as our understanding of creativity increases, the incompleteness of these views becomes apparent. Creative thinking includes both divergent thinking and convergent thinking (Guilford, 1967) and creative products should be novel and practical (Mayer, 1999). Innovation is not only in the minds of individuals but also depends on collaboration within groups and sometimes even the cooperation of a whole society (Simonton and Ting, 2010). Several studies have found that Chinese culture is negatively correlated with some aspects of creativity, such as divergent thinking (Kim et al., 2011), product novelty (Hofstede, 2001), and individual independence (De Dreu, 2010). However, Chinese culture has been shown to promote convergent thinking (e.g., Cheung et al., 2016), product practicality (Xie and Paik, 2019), and success in epidemic prevention and control, poverty reduction, and environmental management through large-scale collective action (Han and Huang, 2018). It can be said that creativity in Chinese culture manifests differently, incorporating wisdom that has unique Chinese characteristics but is also of universal value. We must adopt a neutral, objective perspective to re-assess creativity in Chinese culture. *Zhongyong* is at the core of Chinese culture,

and its relationship with creativity is of particular research interest. Miron-Spektor and Erez (2017) discussed the inherently paradoxical nature of creativity from various angles, including that of the coexistence of novelty and practicality in a creative outcome, and this suggests that the essence of *zhongyong* shares aspects of the essence of creativity. We should not only cease to regard *zhongyong* as the opposite of creativity but also dig more deeply into the unique wisdom that *zhongyong* contributes to problem-solving.

Our study has several limitations. First, we were unable to measure positive and negative *zhongyong* on one scale to find an inflection point between positive and negative effects in continuous data. This should be addressed in future research. Second, in Study 2, the small numbers of participants in some of the groups (the non-positive and non-negative *zhongyong* group and the both positive and negative *zhongyong* group) may have led to some statistical bias. However, the non-significant result of the variance homogeneity test indicated that the ANOVA results were basically acceptable.

In summary, the non-linear relationship between *zhongyong* and creativity uncovered by our research sheds light on the inconsistent findings of previous studies. Our results contribute to a more comprehensive understanding of *zhongyong* and offer clear insights into creativity from an in-depth cultural perspective.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical

restrictions. Requests to access the datasets should be directed to molei@m.scnu.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Human Research Ethics Committee for Non-Clinical Faculties in School of Psychology, South China Normal University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by grants from the National Social Science Foundation of China (19ZDA360, 19BKS162), the Key Institute of Humanities and Social Science in the MOE of China (16JJD880025), Major Basic Research and Applied Research Projects of the Guangdong Education Department (#2017WZDXM004), and the Major Project of Social Science in South China Normal University (ZDPY2208).

ACKNOWLEDGMENTS

We would like to thank Yongxing Guo and his team for establishing and providing DiggMind platform for us to collect research data.

REFERENCES

- Ali, M., Ng, Y. L., and Kulik, C. T. (2014). Board age and gender diversity: a test of competing linear and curvilinear predictions. *J. Bus. Ethics* 125, 497–512. doi: 10.1007/s10551-013-1930-9
- Anderson, C. (2013). *Basic Buddhism: A Beginner's Guide: Volume 1 - Origins, Concepts and Beliefs*. Kindle format.
- Aristotle (1999). *Aristotle: Nicomachean Ethics* (W. D. Ross, Tran.). Kitchener: Batoche Books. Available online at: <https://socialsciences.mcmaster.ca/econ/ugcm/3ll3/aristotle/Ethics.pdf> (accessed March 24, 2022).
- Astakhova, M. N. (2015). The curvilinear relationship between work passion and organizational citizenship behavior. *J. Bus. Ethics* 130, 361–374. doi: 10.1007/s10551-014-2233-5
- Bozionelos, N., Bozionelos, G., Polychroniou, P., and Kostopoulos, K. (2014). Mentoring receipt and personality: evidence for non-linear relationships. *J. Bus. Res.* 67, 171–181. doi: 10.1016/j.jbusres.2012.10.007
- Burnett, M. F., Chiaburu, D. S., Shapiro, D. L., and Li, N. (2015). Revisiting how and when perceived organizational support enhances taking charge: an inverted U-shaped perspective. *J. Manage.* 41, 1805–1826. doi: 10.1177/0149206313493324
- Carette, B., Anseel, F., and Lievens, F. (2013). Does career timing of challenging job assignments influence the relationship with in-role job performance? *J. Vocation. Behav.* 83, 61–67. doi: 10.1016/j.jvb.2013.03.001
- Chang, T.-Y., and Yang, C.-T. (2014). Individual differences in Zhong-Yong tendency and processing capacity. *Front. Psychol.* 5, 1316. doi: 10.3389/fpsyg.2014.01316
- Chang, J.-H., Lin, Y.-C., and Huang, C.-L. (2014). “The dialogue between zhongyong and wisdom (in Chinese),” in *Chinese Social Psychological Review*, Vol. 8, ed Y. Y. Yang (Beijing: Social Sciences Academic Press), 212–225.
- Chen, Y., Qi, Z., Chen, Z., and Tian, X. (2018). An analysis of the effect and mechanism of zhongyong thinking on team innovation (in Chinese). *Forecasting* 37, 15–21.
- Cheung, P. C., Lau, S., Lubart, T., Chu, D. H. W., and Storme, M. (2016). Creative potential of Chinese children in Hong Kong and French children in Paris: a cross-cultural comparison of divergent and convergent-integrative thinking. *Think. Skills Creat.* 22, 201–211. doi: 10.1016/j.tsc.2016.09.005
- Chiu, C.-Y. (2000). Assessment of zhongyong thinking: Preliminary findings of a cross-regional study (in Chinese). *Hong Kong J. Soc. Sci.* 18, 33–55.
- Christoforou, P. S., and Ashforth, B. E. (2015). Revisiting the debate on the relationship between display rules and performance: considering the explicitness of display rules. *J. Appl. Psychol.* 100, 249–261. doi: 10.1037/a0036871
- Dayan, M., Ozer, M., and Almazrouei, H. (2017). The role of functional and demographic diversity on new product creativity and the moderating impact of project uncertainty. *Industr. Market. Manage.* 61, 144–154. doi: 10.1016/j.indmarman.2016.04.016
- De Dreu, C. K. W. (2010). Human creativity: Reflections on the role of culture. *Manage. Organiz. Rev.* 6, 437–446. doi: 10.1111/j.1740-8784.2010.00195.x

- Du, J., Qiu, Y., and Ying, J. (2018). Does zhongyong hamper creativity? an empirical multi-level study (in Chinese). *Stud. Sci. Sci.* 36, 378–384. doi: 10.16192/j.cnki.1003-2053.2018.02.021
- Du, J., and Yao, J. (2015). Zhongyong: The connotation and its relationship with collectivism (in Chinese). *Chinese J. Manage.* 12, 638–646.
- Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191.
- Fisman, R., Shi, J., Wang, Y., and Wu, W. (2020). Social ties and the selection of China's political elite. *Am. Econ. Rev.* 110, 1752–1781. doi: 10.1257/aer.20180841
- Fletcher, J., and Olwyler, K. (1997). *Paradoxical Thinking: How to Profit from Your Contradictions*. San Francisco: Berrett-Koehler.
- Gao, R., Zhang, J., Liu, Y., Zeng, J., Wu, D., Huang, X., et al. (2022). A sustainability lens on the paradox of Chinese learners: four studies on Chinese students' learning concepts under li's "virtue-mind" framework. *Sustainability* 14, 3334. doi: 10.3390/su14063334
- Gordon, W. J. J. (1961). *Synectics: The Development of Creative Capacity*. New York, NY: Harper and Row.
- Gough, H. G. (1979). A creative personality scale for the adjective check list. *J. Personal. Social Psychol.* 37, 1398–1405. doi: 10.1037/0022-3514.37.8.1398
- Grant, A. M., and Schwartz, B. (2011). Too much of a good thing: the challenge and opportunity of the inverted U. *Perspect. Psychol. Sci.* 6, 61–76. doi: 10.1177/1745691610393523
- Greenwald, A. G., and Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *J. Personal. Soc. Psychol.* 79, 1022–1038. doi: 10.1037//0022-3514.79.6.1022
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Personal. Soc. Psychol.* 74, 1464–1480. doi: 10.1037//0022-3514.74.6.1464
- Grossmann, I., Na, J., Varnum, M. E., Kitayama, S., and Nisbett, R. E. (2013). A route to well-being: intelligence versus wise reasoning. *J. Experiment. Psychol. General* 142, 944–953. doi: 10.1037/a0029560
- Grossmann, I., Na, J., Varnum, M. E., Park, D. C., Kitayama, S., and Nisbett, R. E. (2010). Reasoning about social conflicts improves into old age. *Proceed. Nat. Acad. Sci. USA* 107, 7246–7250. doi: 10.1073/pnas.1001715107
- Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York, NY: McGraw-Hill.
- Haans, R. F. J., Pieters, C., and He, Z. L. (2016). Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. *Strategic Manage. J.* 37, 1177–1195. doi: 10.1002/smj.2399
- Han, Q., and Huang, X. (2018). *The Contribution of China's Road to the World – Library of Marxism Studies (Volume 2) (in Chinese)*. Beijing: China Renmin University Press.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Thousand Oaks: Sage.
- Hofstede, G., Hofstede, G. J., and Minkov, M. (2010). *Cultures and Organizations: Software of the Mind*. New York, NY: McGraw-Hill. Available online at: https://e-edu.nbu.bg/pluginfile.php/900222/mod_resource/content/1/G.Hofstede_G.J.Hofstede_M.Minkov%20-%20Cultures%20and%20Organizations%20-%20Software%20of%20the%20Mind%203rd%20edition%202010.pdf (accessed March 24, 2022).
- Hu, W., Philip, A., Shen, J., and Lin, C. (2004). The comparisons of the development of creativity between English and Chinese adolescents. *Acta Psychologic. Sinica* 36, 718–731.
- Huang, C.-L., Chung, Y.-C., and Lin, Y.-C. (2014). "Forest before trees when needed: Global processing and zhongyong (in Chinese)," in *Chinese Social Psychological Review*, ed Y.Y. Yang (Beijing: Social Sciences Academic Press), pp. 49–65.
- Huang, C.-L., Lin, Y.-C., and Yang, C.-F. (2012). Revision of the zhong-yong belief-value scale (in Chinese). *Indigenous Psychol. Res. Chinese Soc.* 38, 3–14. doi: 10.6254/2012.38.3
- Ingram, A. E., Lewis, M. W., Barton, S., and Gartner, W. B. (2016). Paradoxes and innovation in family firms: the role of paradoxical thinking. *Entrepreneur. Theory Pract.* 40, 161–176. doi: 10.1111/etap.12113
- International Communist Workers Party (ICWP). (2013). *Dialectics before Marxism. Red Flag Newspaper*. Available online at: <https://icwpredflag.org/wp/wordpress/dialectics/dialectics-before-marxism/> (accessed November 14, 2013).
- Ji, L., Peng, K., and Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *J. Personal. Soc. Psychol.* 78, 943–955. doi: 10.1037/0022-3514.78.5.943
- Jiao, B., Zhang, D., Liang, A., Liang, B., Wang, Z., Li, J., et al. (2017). Association between resting-state brain network topological organization and creative ability: evidence from a multiple linear regression model. *Biologic. Psychol.* 129, 165–177. doi: 10.1016/j.biopsycho.2017.09.003
- Kim, K. H., Lee, H. E., Chae, K.-B., Anderson, L., and Laurence, C. (2011). Creativity and Confucianism among American and Korean educators. *Creativ. Res. J.* 23, 357–3371. doi: 10.1080/10400419.2011.621853
- Lam, C. F., Spreitzer, G., and Fritz, C. (2014). Too much of a good thing: curvilinear effect of positive affect on proactive behaviors. *J. Organization. Behav.* 35, 530–546. doi: 10.1002/job.1906
- Lee, S., Cheong, M., Kim, M., and Yun, S. (2017). Never too much? the curvilinear relationship between empowering leadership and task performance. *Group Organiz. Manage.* 42, 11–38. doi: 10.1177/1059601116646474
- Lee, S., Yun, S., and Srivastava, A. (2013). Evidence for a curvilinear relationship between abusive supervision and creativity in South Korea. *Leadership Q.* 24, 724–731. doi: 10.1016/j.leaqua.2013.07.002
- Legge, J. (1861). "The analects," in *Sacred Books of the East (Volume 1). Chinese Text Project*. Available online at: <https://ctext.org/analects/ens> (accessed March 24, 2022).
- Legge, J. (1879). "Shang Shu," in *Sacred Books of the East (Volume 3). Chinese Text Project*. Available online at: <https://ctext.org/shang-shu/ens> (accessed March 24, 2022).
- Legge, J. (1885). "The Li Ki," in *Sacred Books of the East (Volume 28). Chinese Text Project*. Available online at: <https://ctext.org/liji/ens> (accessed March 24, 2022).
- Legge, J. (1891). "The Tao Te Ching," in *Sacred Books of the East (Volume 39). Chinese Text Project*. Available online at: <https://ctext.org/dao-de-jing/ens> (accessed March 24, 2022).
- Legge, J. (1899). "The book of changes," in *Sacred Books of the East (Volume 16). Chinese Text Project*. Available online at: <https://ctext.org/book-of-changes/ens> (accessed March 24, 2022).
- Legge, J. (1985). *The Works of Mencius*. Oxford: Clarendon Press. Chinese Text Project. Available online at: <https://ctext.org/mengzi/ens> (accessed March 24, 2022).
- Leung, A. K.-y., Liou, S., Miron-Spektor, E., Koh, B., Chan, D., Eisenberg, R., et al. (2018). Middle ground approach to paradox: within- and between-culture examination of the creative benefits of paradoxical frames. *J. Personal. Soc. Psychol.* 114, 443–464. doi: 10.1037/pspp0000160
- Lewis, M. W. (2000). Exploring paradox: toward a more comprehensive guide. *Academy of Management Review* 25, 760–776. doi: 10.5465/amr.2000.3707712
- Li, J. (2012). *Cultural Foundations of Learning: East and West*. New York, NY: Cambridge University.
- Li, J., Zhang, D., Liang, A., Liang, B., Wang, Z., Cai, Y., et al. (2017). High transition frequencies of dynamic functional connectivity states in the creative brain. *Scientific Rep.* 7, 46072. doi: 10.1038/srep46072
- Li, M. (2017). *The mechanism of zhongyong thinking mode on the creative ideas generation and its neural basis*. Master's thesis, Central China Normal University, Wuhan, China.
- Liao, B., and Dong, W. Q. (2015). The study for the relationship among mean thinking, organizational harmony and innovation behavior of knowledge staff (in Chinese). *Sci. Technol. Progr. Policy* 32, 150–154.
- Lin, J., Cui, X., Dai, X., and Mo, L. (2018). Regional homogeneity predicts creative insight: a resting-state fMRI study. *Front. Hum. Neurosci.* 12, 210. doi: 10.3389/fnhum.2018.00210
- Lin, X., and Wang, M. (1999). *Williams Creativity Assessment Packet (in Chinese)*. Taipei, China: Psychological Publishing.
- Liu, H., Wang, F., and Yang, X. (2015). More dialectical thinking, less creativity? the relationship between dialectical thinking style and creative personality: The case of China. *PLoS One* 10, e0122926. doi: 10.1371/journal.pone.0122926
- Liu, C. (2019). The possibility or impossibility of zhongyong: on the dilemma of the empirical psychological studies on zhongyong (in Chinese). *J. Nanjing Norm. Univ.* 05, 65–74.

- Martin, R. L. (2009). *The Opposable Mind: How Successful Leaders Win Through Integrative Thinking*. Boston: Harvard Business Press.
- Mayer, R. E. (1999). "Fifty years of creativity research," in *Handbook of Creativity*, ed R. J. Sternberg (New York, NY: Cambridge University Press), p. 449–460.
- Miron-Spektor, E., and Erez, M. (2017). "Looking at creativity through a paradox lens: deeper understanding and new insights," in *Handbook of Organizational Paradox: Approaches to Plurality, Tensions and Contradictions*, eds W. K. Smith, M. W. Lewis, P. Jarzabkowski, and A. Langley (Oxford: Oxford University Press).
- Miron-Spektor, E., Gino, F., and Argote, L. (2011). Paradoxical frames and creative sparks: enhancing individual creativity through conflict and integration. *Organization. Behav. Hum. Decis. Process.* 116, 229–240. doi: 10.1016/j.obhdp.2011.03.006
- Mo, S., Ling, C.-D., and Xie, X.-Y. (2019). The curvilinear relationship between ethical leadership and team creativity: the moderating role of team faultlines. *J. Bus. Ethics* 154, 229–242. doi: 10.1007/s10551-016-3430-1
- Ng, A. K. (2001). *Why Asians are Less Creative Than Westerners?*. Singapore: Prentice-Hall.
- Ng, A. K. (2004). *Liberating the Creative Spirit in Asian Students*. Singapore: Prentice-Hall.
- Nieß, C., and Biemann, T. (2014). The role of risk propensity in predicting self-employment. *J. Appl. Psychol.* 99, 1000–1009. doi: 10.1037/a0035992
- Niu, W., and Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *Int. J. Psychol.* 36, 225–241. doi: 10.1080/00207590143000036
- Pang, P. (1980). A discussion on zhongyong (in Chinese). *Soc. Sci. China*. 01, 75–100.
- Pang, P. (2000). Zhongyong and Trichotomization (in Chinese). *Literat. Hist. Philosoph.* 04, 21–27. doi: 10.16346/j.cnki.37-1101/c.2000.04.004
- Pierce, J. R., and Aguinis, H. (2013). The too-much-of-a-good-thing effect in management. *J. Manage.* 39, 313–338. doi: 10.1177/0149206311410060
- Rapp, T. L., Bachrach, D. G., Rapp, A. A., and Mullins, R. (2014). The role of team goal monitoring in the curvilinear relationship between team efficacy and team performance. *J. Appl. Psychol.* 99, 976–987. doi: 10.1037/a0036978
- Ren, S., and Chadee, D. (2017). Is guanxi always good for employee self-development in China? examining non-linear and moderated relationships. *J. Vocation. Behav.* 98, 108–117. doi: 10.1016/j.jvb.2016.10.005
- Ren, C., and Xin, T. (2013). Longitudinal study on predicting effect of social economic status on students' performance (in Chinese). *Educat. Res.* 34, 79–87.
- Rotolo, D., and Messeni Petruzzelli, A. (2013). When does centrality matter? Scientific productivity and the moderating role of research specialization and cross community ties. *J. Organ. Behav.* 34, 648–670. doi: 10.1002/job.1822
- Shao, S., Fan, M., and Yang, L. (2013). How resource industry dependence affects economic development efficiency: a test and explanation of conditional "resource curse" hypothesis (in Chinese). *Manage. World* 02, 32–63. doi: 10.19744/j.cnki.11-1235/f.2013.02.005
- Simonton, D., and Ting, S. (2010). Creativity in eastern and western civilizations: the lessons of historiometry. *Manage. Organiz. Rev.* 6, 329–350. doi: 10.1111/j.1740-8784.2010.00188.x
- Stouten, J., van Dijke, M., Mayer, D. M., De Cremer, D., and Euwema, M. C. (2013). Can a leader be seen as too ethical? the curvilinear effects of ethical leadership. *Leadersh. Q.* 24, 680–695. doi: 10.1016/j.leaqua.2013.05.002
- Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking—Norms—Technical Manual Research Edition—Verbal Tests, Forms A and B—Figural Tests, Forms A and B*. Princeton: Personnel Press.
- Vergauwe, J., Wille, B., Hofmans, J., Kaiser, R. B., and De Fruyt, F. (2018). The double-edged sword of leader charisma: understanding the curvilinear relationship between charismatic personality and leader effectiveness. *J. Personal. Soc. Psychol.* 114, 110–130. doi: 10.1037/pspp0000147
- Vicentini, F., and Boccadelli, P. (2016). Career diversity and project performance in the Italian television industry. *J. Bus. Res.* 69, 2380–2387. doi: 10.1016/j.jbusres.2015.10.007
- von den Driesch, T., da Costa, M. E. S., Flatten, T. C., and Brettel, M. (2015). How CEO experience, personality, and network affect firms' dynamic capabilities. *Euro. Manage. J.* 33, 245–256. doi: 10.1016/j.emj.2015.01.003
- Wang, M.-Y., Tang, D.-L., Kao, C.-T., and Sun, V. C. (2013). "Banner evaluation predicted by eye tracking performance and the median thinking style," in *Design, User Experience, and Usability. Health, Learning, Playing, Cultural, and Cross-Cultural User Experience*, ed Marcus, A. DUXU 2013 (Berlin, Heidelberg: Springer), pp. 129–138.
- Wei, X., Liu, Y., and Chen, S. (2015). A meta-analysis of the relationship between team demographic diversity and team performance (in Chinese). *Acta Psychologica Sinica* 47, 1172–1187. doi: 10.3724/SP.J.1041.2015.01172
- Wu, C.-H., and Lin, Y.-C. (2005). Development of a Zhong-Yong thinking style scale (in Chinese). *Indigenous Psychol. Res. Chinese Soc.* 24, 247–300. doi: 10.6254/2005.24.247
- Wu, C.-L., and Chen, H.-C. (2017). Normative data for Chinese compound remote associate problems. *Behav. Res. Methods* 49, 2163–2172. doi: 10.3758/s13428-016-0849-3
- Xie, G., and Paik, Y. (2019). Cultural differences in creativity and innovation: Are Asian employees truly less creative than western employees?. *Asia Pacific Bus. Rev.* 25, 123–147. doi: 10.1080/13602381.2018.1535380
- Xing, L., Sun, J., Yin, K., and Wang, Z. (2018). The Too-Much-of-a-Good-Thing effect and its mechanisms (in Chinese). *Adv. Psychologic. Sci.* 26, 719–730. doi: 10.3724/SP.J.1042.2018.00719
- Yang, C.-F. (2009). A case of attempt to combine the Chinese traditional culture with the social science: the social psychological research of zhongyong (in Chinese). *J. Renmin Univ. China* 23, 53–60.
- Yang, C.-F. (2010). Multiplicity of zhongyong studies (in Chinese). *Indigenous Psychol. Res. Chinese Soc.* 34, 3–96. doi: 10.6254/2010.34.3
- Yang, C.-F., and Chiu, C.-Y. (1997). *The preliminary exploration of zhongyong practice cognition (in Chinese). The Fourth Interdisciplinary Symposium on Chinese Psychology and Behavior Interdisciplinary (CPBI), May 29–31, Taipei, China.*
- Yang, C.-F., and Lin, S. (2012). A construct validity study of C.-F. Yang's zhongyong conceptualization (in Chinese). *Sociological Stud.* 27, 167–186+245. doi: 10.19934/j.cnki.shxyj.2012.04.011
- Yang, C.-F., Yang, Z., and Ding, Y. (2014). "A second validity study of C.-F. Yang's zhongyong conceptualization of (in Chinese)," in *Chinese Social Psychological Review*, ed Y.Y. Yang (Beijing: Social Sciences Academic Press), pp. 18–42.
- Yang, C. F. (2008). "Research on the zhongyong practice cognition: Towards the construction of a set of local psychological knowledge system (in Chinese)," in *On the Approaches of Local Psychological Research*, ed Z. F. Yang (Taipei: Yuan-Liou Publishing Co., Ltd).
- Yang, C., and Zhang, L. (2018). Zhongyong value orientation and employee innovation behavior: a moderated mediating model (in Chinese). *J. Technic. Econ. Manage.* 02, 54–58.
- Yao, X., Yang, Q., Dong, N., and Wang, L. (2010). Moderating effect of Zhong Yong on the relationship between creativity and innovation behavior. *Asian J. Soc. Psychol.* 13, 53–57. doi: 10.1111/j.1467-839X.2010.01300.x
- Yi, X., Hu, W., Scheithauer, H., and Niu, W. (2013). Cultural and bilingual influences on artistic creativity performances: comparison of German and Chinese students. *Creativ. Res. J.* 25, 97–108. doi: 10.1080/10400419.2013.752260
- Zettler, I., and Lang, J. W. B. (2015). Employees' political skill and job performance: an inverted U-shaped relation?. *Appl. Psychol.* 64, 541–577. doi: 10.1111/apps.12018
- Zha, P., Walczyk, J. J., Griffith-Ross, D. A., Tobacyk, J. J., and Walczyk, D. F. (2006). The impact of culture and individualism-collectivism on the creative potential and achievement of American and Chinese adults. *Creativ. Res. J.* 18, 355–366. doi: 10.1207/s15326934crj1803_10
- Zhang, G. X., and Gu, X. Y. (2015). Moderation thinking and employees' creativity (in Chinese). *Sci. Res. Manage.* 36, 251–257. doi: 10.19571/j.cnki.1000-2995.2015.s1.037
- Zhang, H., Li, M., and Zhou, Z. (2020). "Different people, different views? the relationship between zhongyong and creativity (in Chinese)," in *Chinese Social Psychological Review*, ed Y.Y. Yang (Beijing: Social Sciences Academic Press), pp. 12–27+236.
- Zhang, J., and Long, L. (2016). The curvilinear relation between leader's forgiveness and job performance: moderating effects of conscientiousness and procedural justice (in Chinese). *Manage. Rev.* 28, 134–144. doi: 10.14120/j.cnki.cn11-5057/f.2016.04.014
- Zhou, Z., Hu, L., Sun, C., Li, M., Guo, F., and Zhao, Q. (2019). The effect of zhongyong thinking on remote association thinking: an EEG study. *Front. Psychol.* 10, 207. doi: 10.3389/fpsyg.2019.00207

- Zhou, Z., Zhang, H., Li, M., Sun, C., and Luo, H. (2021), The effects of zhongyong thinking priming on creative problem-solving. *J. Creat. Behav.* 55, 145–153. doi: 10.1002/jocb.441
- Zhu, X. (2012). *Variorum of the Four Books – Corpus of Ancient Classic Masterpieces (Chinese Edition)*. Zhejiang: Zhejiang Ancient Books Publishing House.
- Zhu, W., Shang, S., Jiang, W., Pei, M., and Su, Y. (2019). Convergent thinking moderates the relationship between divergent thinking and scientific creativity. *Creativ. Res. J.* 31, 320–328.

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 Gao, Huang, Yao, Liu, Zhou, Zhang, Cai, Zuo, Zhan and Mo. 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.



Influence of Mentorship and the Working Environment on English as a Foreign Language Teachers' Research Productivity: The Mediation Role of Research Motivation and Self-Efficacy

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the
Americas—Nevis, United States

Reviewed by:

Hassan Mohebbi,
European Knowledge Development
Institute (EUROKD), Turkey
Zhonggen Yu,
Beijing Language and Culture
University, China
Zhengdong Gan,
University of Macau, Macao SAR,
China

*Correspondence:

Lawrence Jun Zhang
lj.zhang@auckland.ac.nz

†ORCID:

Yanping Li
orcid.org/0000-0001-6741-7132
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: 29 March 2022

Accepted: 24 May 2022

Published: 15 June 2022

Citation:

Li Y and Zhang LJ (2022)
Influence of Mentorship
and the Working Environment on
English as a Foreign Language
Teachers' Research Productivity:
The Mediation Role of Research
Motivation and Self-Efficacy.
Front. Psychol. 13:906932.
doi: 10.3389/fpsyg.2022.906932

Yanping Li[†] and Lawrence Jun Zhang^{*†}

Faculty of Education and Social Work, The University of Auckland, Auckland, New Zealand

Research productivity is an important criterion for the university to assess teachers. Studies about factors that affect teachers' research productivity are increasing nowadays. It is generally agreed that academics' research productivity depends on how much mentorship is provided to them and how the current working environment is mediated by their research motivation and self-efficacy. Despite the increasing amount of the literature along this line, we know little about what kinds of situations that Chinese university English as a foreign language (EFL) teachers are in and how they regard the importance of mentorship and what roles their working environments would play in affecting their research productivity. To fill the research gap, we utilized the snowball method to collect the survey data from 546 Chinese EFL tertiary teachers. The results show that mentorship is not correlated with research productivity while the working environment has a positive direct correlation with it. Both motivation and self-efficacy mediate the working environment and research productivity significantly. Specifically, only extrinsic motivation has a negative mediation influence on teachers' research productivity; teachers' intrinsic motivation and self-efficacy play a positive mediation role in affecting their research productivity.

Keywords: research motivation, research self-efficacy, research productivity, mediation effects, university teachers, foreign language teachers

INTRODUCTION

International and local university rankings have been a symbol of the university's influence and competitiveness (Morze et al., 2022). Research productivity takes an important position for its weight in the world university ranking. For example, research takes about 40% of the total score in the 2022 Academic Rankings of World Universities. Similarly, in the Times Higher Education ranking system, research and citations (research impact) account for 30% respectively of the overall score in the 2022 rankings. As the core of science, the publications are significant for communication and exchanging current findings, knowledge, and ideas (Fox, 1992). Besides, teachers' professional development relates to not only teaching but also their career advancement

based on publications and other forms of research productivity (Borg and Liu, 2013). Additionally, Zhang (2021a) found that teachers felt the institutional push to do research, and are pressurized to get it published in international journals (Mu and Zhang, 2018). In their study, Zheng and Gao (2016) found that Chinese scholars preferred using Chinese and English in their effort to pursue research excellence by leveraging on languages they are proficient in for research and publications. This is particularly true of scholars, who work in language-related disciplines (e.g., language and literacy education, including foreign language acquisition). Research has actually been an important means to teachers' professional development (Gao et al., 2011). Among diverse professional development choices, research engagement has been strongly suggested as an innovative model for its potential as a powerful transformative force in English as a foreign language (EFL) teachers' work and professional development (Borg, 2010). It is now a requirement for faculty members in research institutions and all types of institutions to publish (Lucas and Murry, 2011), which attracted scholars to study individual-level research productivity and factors that contribute to its increase (Nygaard, 2017). Uwizeye et al. (2021) found that, as individual factors, academic qualifications, gender, motivations, and research self-efficacy had the most consequences on teachers' research productivity in African higher education institutions. Among those institutional factors, research environments, or cultures, are considered as the most influential ones that impact research productivity (Ajajawi et al., 2018). Similarly, mentoring affects the mentee, mentor, and organization positively (Eby and Robertson, 2020). In order to have a better understanding of teachers' research productivity, we focus on those aspects relating to teachers' psychology, namely, research motivation, self-efficacy, and their mediating role in influencing institutional policies on teachers' research productivity, particularly the provision or absence of mentorship and the working environment of EFL teachers in China.

We propose that the mediating mechanism be used to explain how university EFL teachers' motivation and self-efficacy as mediators affect their research productivity. To this end, we have three specific aims: (1) to investigate the mediating role of teachers' research motivation and self-efficacy in the relationship between institutional support and research productivity, (2) to broaden the former studies which focused on the direct effect on research productivity, (3) to provide practical information, especially for administrators in higher education to increase teachers' research productivity.

LITERATURE REVIEW

Many studies have been conducted to identify factors that affect an individual's research productivity (Freedenthal et al., 2008). These factors were categorized into individual and institutional factors (Uwizeye et al., 2021). Individual factors included teachers' motivation and self-efficacy, while institutional factors included mentorship and working environment. Following this line of research, we introduced teachers' mentorship and working

environment mediated by their research motivation and self-efficacy as a mechanism of affecting research productivity.

Research Productivity

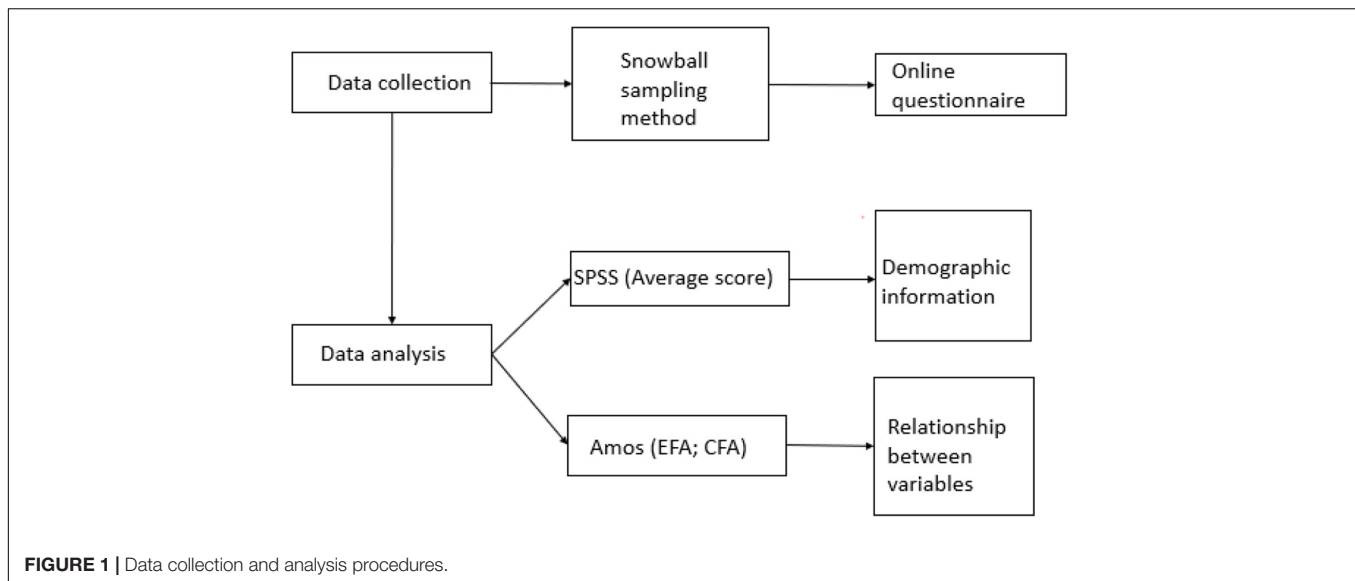
The appearance of the construct of research productivity could date back to the early 1970s (Creswell, 1985), and it is defined as the number of publications in academic refereed journals and/or scholarly books as well as presentations in conferences which usually have the chief function of productivity measure for promotion and tenure in university (Dundar and Lewis, 1998). However, different institutions and disciplines vary in measuring productivity (McGill and Settle, 2012; Paul and Mukhopadhyay, 2022). Generally, for university teachers, the number of research publications in top-ranked journals over the past 10 years (Long et al., 2009), reports, monographs, book chapters, book reviews, books, and the amount of research funding awarded are often used to assess their research productivity (Blackburn and Lawrence, 1995). Despite research productivity being crucial for teachers and universities, studies on research productivity amongst university teachers and sole and joint productivity analysis in academia are rather limited (Jang and Shin, 2011). Borg (2007) has also called for empirical research on the EFL teachers' research engagement. Additionally, Heng et al. (2022) reported that the academics' research engagement was limited in the marginalized global south nations, as well as countries such as China, and such a situation exacerbated during the pandemic. Therefore, they appealed that research be conducted to better understand how the fragile research environment would affect teachers' research, and enhance their research productivity.

Institutional Support and Research Productivity

Research support is defined as any provided resource that can boost a faculty member's ability to engage in scholarship (McGill and Settle, 2012). Previous research has investigated the main types of institutional support and the relationship between institutional support and research productivity (Freedenthal et al., 2008). According to these studies, institutional support includes three sub-constructs: research mentoring experience (i.e., being mentored in research), research environment, and research education (Jang and Shin, 2011). Among various support, mentorship is the most prominent factor affecting teachers' research productivity (Allen et al., 2018). And a beneficial working environment bolsters teachers' research productivity (Aboagye et al., 2021). Luo and Hyland (2016) have found that a lack of institutional support is one of the main reasons why Chinese university teachers' manuscripts cannot be published and even if they get published, their work is cited less frequently.

Mentorship Support

Mentorship is a kind of institutional support in which a more-experienced member supplies information, support, and guidance to a less-experienced, usually new member of an institution to promote the successful chances of the latter within or beyond the institution (Campbell and Campbell, 1997). Transferring skills and supporting continuous learning,



especially when skills are scarce, are the main functions of mentoring (Nundulall and Dorasamy, 2010). General guidance and skill development training from the relevant technical expertise also enhance research engagement in university teachers (Wilkins, 2011). For instance, Loewen (2019) reported that language teachers are neither paid nor trained to do research. In his meta-analysis of 43 studies on mentoring, Simmering (2007) found that teachers' research output is comparatively low with a lack of mentorship programs. Likewise, mentorship programs will increase research output (Nundulall and Reddy, 2011). Besides, studies have found that the availability of training facilities at universities can improve the publication rates of university teachers (Phillips and Russell, 1994). Similarly, Kelly and Warmbrod (1986) found that the lack of training and reflection of research hindered university teachers' research productivity. In addition, engagement in context-sensitive activities and networking opportunities is vital for university teachers because this can help them secure internal and external funding and be exposed to suitable methods for publishing books and articles (Shaw, 2002).

Working Environment Support

Research environment refers to the behaviors that include, at a minimum, shared values, assumptions, beliefs, rituals, and the valued, worthwhile, and pre-eminent activity with a central focus on the acceptance and recognition of research practices and outcomes (Evans, 2007). It is found that faculty's work environments drove their productivity and prominence (Way et al., 2019). It was necessary for higher education institutions to provide a conducive research environment for academics to stimulate their engagement with research (Tadesse and Khalid, 2022). In reality, studies found that teachers had unsatisfied work environments. With limited available time for research, teachers were also imposed by the excessive workload that results in their lower research productivity as well as fewer opportunities for research training (Griffiths et al., 2010). Similarly, Kelly

and Warmbrod's (1986) qualitative research revealed that heavy teaching load had a negative impact on the university teachers' research productivity. In China, the limited educational resources were adverse to language teachers' professional development (Gao and Xu, 2014).

Across various contexts and disciplines, lack of time acted as a negative mechanism leading to decreased research productivity (Ajjawi et al., 2018). Because of that, some researchers emphasized the necessity of separating research from teaching hours in faculty time allocation (Creswell, 1985), and they argued that universities must arrange schedules that allowed teachers to have sufficient time to gather resources and conduct research (Graves et al., 1982). Besides time support, adequate financial support can positively affect teachers' research productivity (Jung, 2012). The allocation of funding for research was output-driven, usually in the form of academic publications (Nundulall and Dorasamy, 2010). Realizing that, department heads and chairs have provided institutional support to increase teachers' research engagement, such as supporting research travel and nominating teachers for research honors and awards, as Bland et al. (2005) reported. In a study conducted by Dundar and Lewis (1998) in the United States, it was found that funding-related support and faculty productivity had a positive relationship. Increasing institutional funding for teachers would improve teachers' research productivity (McGill and Settle, 2012). Also, McGill and Settle (2012) discovered that teachers who received more institutional funding were more likely to engage in research. More studies still need to be conducted to explore how institutions can better support university teachers' research quality and productivity (Dundar and Lewis, 1998).

Teachers' Research Motivation and Research Productivity

It is found in recent empirical studies that psychological factors are valuable in explaining research productivity (Hemmings and Kay, 2016). Among these factors, motivation is a prominent

one. The definition of motivation is “the dynamically changing cumulative arousal in a person that initiates, directs, coordinates, amplifies, terminates, and evaluates the cognitive and motor processes whereby initial wishes and desires are selected, prioritized, operationalized, and (successfully or unsuccessfully) acted out” (Dörnyei and Ottó, 1998, p. 65). Motivation can be divided into intrinsic motivation and extrinsic motivation: intrinsic motivation refers to the internal fascination and gratification of the activity itself as the main reasons to attract people to engage in an activity, while extrinsic motivation means incentives or external pressures that attract people to pursue an activity (Reeve, 1995).

Extrinsic Motivation

As for external incentives, substantial incentives have been used to facilitate teachers to do research. Substantial incentives such as income increases and bonuses took a more significant role than non-substance encouragement, such as the certification and honorary title award in Chinese universities (Henley and Nyaw, 1986). For instance, Santo et al. (2009), who investigated the faculty at the School of Education in Midwestern America, found that the department’s lack of financial support limits the teachers’ opportunities to attend academic-related activities, thus leading to low research productivity among them. Meanwhile, in another study, Creamer (1998) identified department heads or deans who considered research and research productivity the center of rewards. More specifically, Brewer and Brewer (1990) found that 42 of the responding deans in their sample believed that the presence of a merit pay system could and/or do increase faculty research productivity. Besides the financial rewards, it is found that teachers had moderately different patterns of research productivity with varying statuses of tenure (McNurlen and West, 2000). However, the findings are opposite, involving no relationship, negative relationship, and positive relationship between teachers’ tenure and research productivity (Chen et al., 2010). For example, Teodorescu (2000) found no effect of tenure on research productivity. Inversely, Chen et al. (2010) noted that the relationship between research productivity and tenure was strong so that universities took great advantage of it to make teachers’ research productive. Whereas Santo et al. (2009) found tenure and research productivity had a negative relationship because teachers worried nothing about obtaining tenure, thus their research motivation to publish decreased. Especially for those tenured teachers, who may be in a semi-retired state but still employed by the university, they hardly had the motivation to do research (Chen et al., 2006). Besides tenure, the promotion also affects teachers’ research.

From the management perspective, the promotion has been considered as one of the effective ways of encouraging productivity among university teachers (Lai, 1990). Outstanding external rewards make faculty members try their best when promotion and tenure decisions are forthcoming, while less effort after promotion; foretelling fluctuations in productivity through time (Hu and Gill, 2000). However, after achieving the title of full professor and as retiring faculty member, the research productivity is ultimately a recession in their later academic life (Hu and Gill, 2000). Realizing this, universities

have been using promotion as an extrinsic motivational tool to boost the research output of university teachers (Chen et al., 2010). Meanwhile, as research output is one of the most important indicators in academic promotion assessment, the promotion has been a robust extrinsic motivator on research productivity (Chen et al., 2006). Similarly, Tien and Blackburn (1996) found that the expected research productivity remains low because of no conferred promotion reward. Higher education institutions could influence academic staff’s research behavior by manipulating the reward structure for promotion (Fox, 1985).

Generally, most universities have specific and clear performance appraisal documents for teachers. However, the workload and research productivity requirements of teachers differ from one university to another. For example, teachers in Chinese higher education institutions are evaluated on their teaching, research, administration, curriculum, and subject construction (Liu and Yu, 2013). To improve the quantity and quality of research, almost every university has *Research and Teaching Office*, which is a traditional department for different levels educational authorities in China to administrate research and pedagogical innovation activities (Gao et al., 2010). Teachers will be in different research-engaged statuses in various institutions because of the different policies on appraising publications. Meanwhile, teachers’ salary depends on their professional titles in combination with their professional performance, in which research productivity accounts for a large proportion (Luthans and Stajkovic, 1999). Additionally, Borg (2009) proposed that English teachers were mainly driven by practical (e.g., solving teaching problems, identifying better teaching ways) and professional (e.g., professional development) concerns to conduct research.

Intrinsic Motivation

Besides extrinsic motivation, intrinsic motivation also plays a vital role in motivating teachers to do research. McKeachie (1982) has postulated that academics publish for the enjoyment of the process of inquiry rather than the external rewards. Differently, while expressing their own willingness to integrate research into teaching, teachers are dismissive of other teachers who lack interest in research (Sato and Loewen, 2019). The sense of satisfaction about discovery, such as defining research goals and outlining paths to achievements, gives academics satisfying emotions, leading them to high work motivation (Stark, 1986). Moreover, intrinsic motivation is significant in the preference for autonomy and independence as well as achieving something on one’s own (Blackmore and Kandiko, 2011). Thus, intrinsic motivation creates teachers engaging in research actively (Trembley et al., 2009). However, teachers with research experience and skills are sometimes demotivated by their perception of the discrepancy between their preferred research and the institutions that encourage them to conduct (Kyaw, 2021).

Teachers’ Research Self-Efficacy and Research Productivity

Research self-efficacy is an individual’s beliefs about his or her ability to carry out research (Morrison and Lent, 2014).

TABLE 1 | Participant information.

Demographic characteristic	N	Valid%
Gender		
Female	380	74.8
Male	128	25.2
Missing	0	-
Age		
≤ 30	67	13.2
31–40	197	38.8
41–50	199	39.2
≥ 51	45	8.9
Missing	0	-
Rank		
Assistant lecturer	59	11.6
Lecturer	226	44.5
Associate professor	195	38.4
Full professor	28	5.5
Missing	0	-
Total	508	-

TABLE 2 | Factor loadings for exploratory factor analysis and internal reliability of the two institutional supports ($n = 508$).

Factor loading				
Factor	Item	1	2	α
Mentorship (M)	M1-item 1	0.482		0.949
	M2-item 2	0.785		
	M3-item 3	0.953		
	M4-item 4	0.955		
	M5-item 5	0.949		
	M6-item 6	0.845		
	M7-item 7	0.643		
	M8-item 8	0.536		
Working Environment (WE)	WE1-item 9		0.626	0.944
	WE2-item 10		0.548	
	WE3-item 11		0.790	
	WE4-item 12		0.832	
	WE5-item 13		0.817	
	WE6-item 14		0.923	
	WE7-item 15		0.893	
	WE8-item 16		0.854	

Items with factor loading of 0.30 or greater are included; α = Cronbach's alpha.

Generally, studies on teachers' research self-efficacy are divided into three categories. First, previous studies have shown that research self-efficacy has a positive relationship with research disposition which consists of research interest and research experience (Bandura and Adams, 1977; Bard et al., 2000; West et al., 2007). Specifically, research interest has a high association with research self-efficacy (Bard et al., 2000; West et al., 2007) and research experience will bolster research self-efficacy (Love et al., 2007). Second, research self-efficacy correlates with research support directly (Jang and Shin, 2011). Research support involves three sub-constructs: research mentoring experience

(i.e., being mentored in research), research training environment, and research education (Jang and Shin, 2011). Specifically, research mentoring experience associates with research self-efficacy positively (Hollingsworth and Fassinger, 2002) as well as research training environment and research education (Holland, 1985; Judge et al., 2007). Third, research self-efficacy will impact research outcome. Research outcome refers to research outcome expectation (Bard et al., 2000) and research productivity (Phillips and Russell, 1994; Kahn and Scott, 1997). Specifically, Bieschke et al. (1998) found there is a strong positive relationship between research self-efficacy and research outcome expectation. However, the correlation between teachers' self-efficacy and research productivity has inconsistent findings. For example, Landino and Owen (1988), Vasil (1992), and Kahn and Scott (1997) have found that self-efficacy correlates positively with university teachers' research productivity, inversely, Pasupathy and Oginga (2014) found that the correlation between research self-efficacy and research productivity is weak. Chineses scientific research fails to establish an international reputation (Lin and Fan, 1990) as their limited multilingual capabilities in scholarly publication (Zheng and Gao, 2016). More recently, studies have shown that teachers' ability to do research is closely linked to their professional identities (Yuan and Zhang, 2020). From the given literature, it can be deduced that the influence of research self-efficacy on research outcome should be further studied to provide more implications for higher education (Jang and Shin, 2011). Besides, the relationship between self-efficacy and academic achievement has been widely studied in the context of the tertiary English as second language (ESL) (i.e., contexts where English is the dominant or first language in education, law and every sphere in society and those who learn it in such native-speaking contexts are ESL learners and users). However, such studies have been scarcely reported in relation to an EFL context (Noorollahi, 2021). To fill this gap, our study investigated how the academics were affected by their self-efficacy beliefs in China, a typical EFL context, where English is seldom used in society as a working language; nor is it even used informally for daily communication. English is offered as a subject in schools and universities and taught as a foreign language (Zhang, 2021b).

Meditation Effect of Motivation and Self-Efficacy on Research Productivity

A supportive research environment could affect teachers' intrinsic motivation in research and their productivity (Peng and Gao, 2019); also, research self-efficacy correlates with research support directly (Jang and Shin, 2011). However, few studies have examined the interaction between individual and institutional characteristics and how individuals handle conflicting goals or priorities (Nygaard, 2017). Academics' priority is affected by their self-efficacy and the perception of organizational priorities (Williams and Kotrlik, 2004). Findings from recent studies imply that there might be mediation or moderation between research productivity and the research environment by diverse organizational and individual-level factors (Ajjawi et al., 2018). For example, Gelso and Lent (2000) indicated that the self-efficacy plays a mediation role between factors, such as research

training environments and ultimate continuum outcomes (e.g., research productivity). However, empirical research into the potential mediational role between institutional factors and research productivity is scarce (Kozhakhmet et al., 2022). Specifically, it is necessary to examine the mediating function of individual characteristics between institutional factors and research productivity (Holtum and Goble, 2006). Based on the above considerations, the following hypotheses are proposed:

H1: Both motivation and self-efficacy mediate institutional support and research productivity.

H2: Either motivation or self-efficacy mediates institutional support and research productivity.

H3: Neither motivation nor self-efficacy mediates institutional support and research productivity.

METHODOLOGY

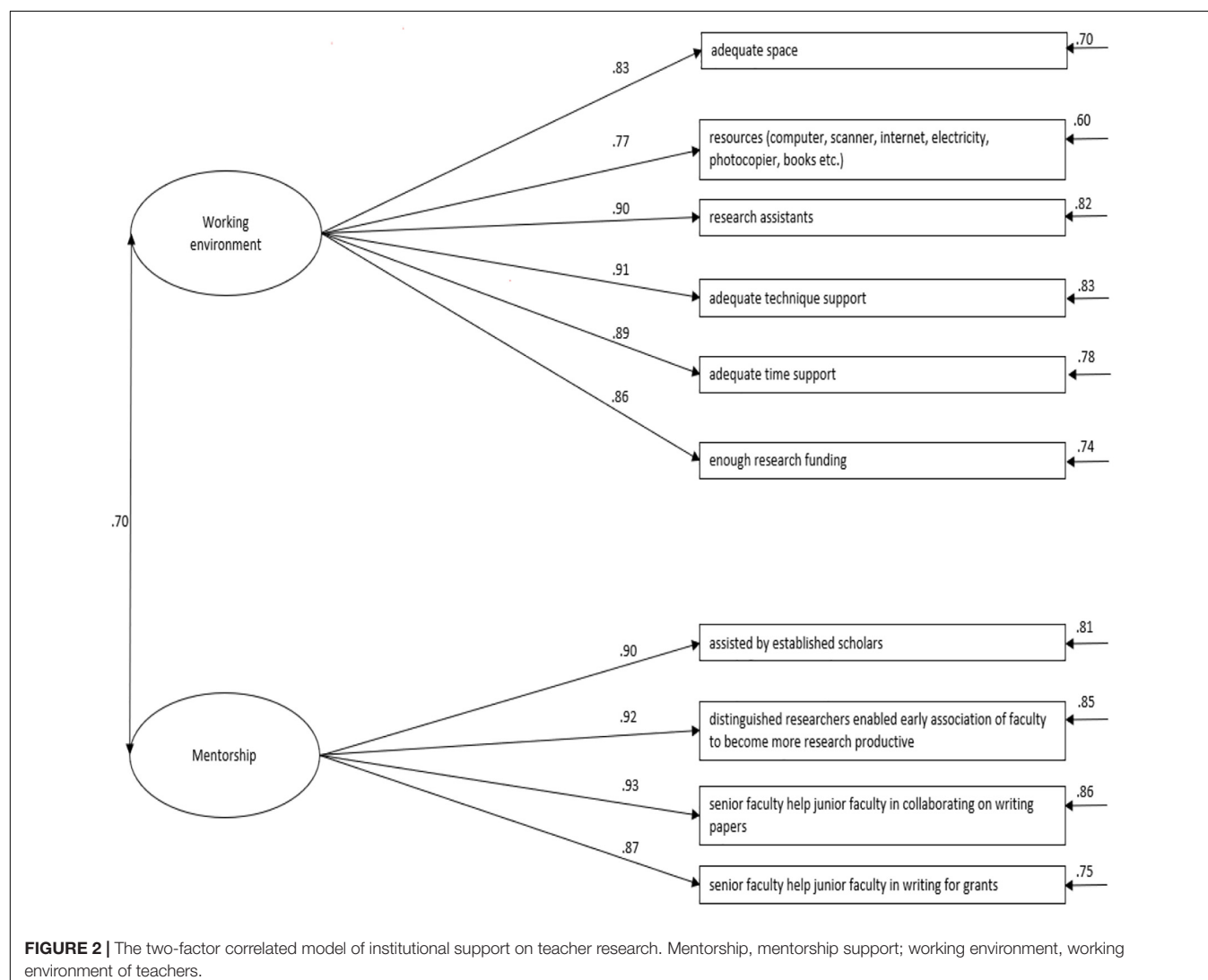
Participants

The data of this study were gathered through a questionnaire survey sent to Chinese university EFL teachers. As long

TABLE 3 | Model fit indices of institutional support model.

Model	Description	χ^2 (df)	χ^2/df	CFI	Gamma hat	RMSEA	90% CI	SRMR	AIC
IS	IS model (Two-factor, 10 items)	130.735 (34)	3.845***	0.98	0.96	0.075	0.062–0.089	0.047	172.74

***Means the result is significant because it is within the range of $2 \leq \chi^2/df \leq 5$.



as they were teaching English at the university level, they were eligible to participate in this survey. The snowball sampling method (Dörnyei, 2007) was adopted in this study for it was the most comprehensive way to collect the representative data in China. With the snowball sampling method, participants who met the criteria of the present study were contacted first then, these teachers were asked to pass on the information to other teachers who might be interested in participating in this study (Dörnyei and Taguchi, 2010). Finally, a total of 536 teachers responded to the questionnaire.

Instrument

The inventory of *Questionnaire on Teachers' Research Productivity (QTRP)* was composed of four subscales respectively named *Questionnaire on Teacher Research Self-efficacy (QTRSE)*, *Questionnaire on Teacher Research Motivation (QTRM)*, and *Questionnaire on Institutional Support for Teacher Research (QISTR)* with 37 items in total to test the influence of these factors on teachers' research productivity as well as a one-factor

scale with three items measuring teachers' research productivity (*Questionnaire on Teachers' Research Productivity*).

Teachers chose the degree level of agreeing with each item on a six-point rating scale. The scale had three negative responses and three positive responses symmetrically, 1 (Strongly Disagree), 2 (Disagree), 3 (Moderately Disagree), 4 (Moderately Agree), 5 (Agree), 6 (Strongly Agree). Teachers were asked to tick the response box corresponding to their beliefs about each item.

Procedures for Data Collection

An invitation email was sent to the heads of school to get their permission for disseminating the questionnaire to their faculty during the first semester in 2019–2020. After the head approved of the researcher's request, an invitation to respond to the questionnaire and the PISs and CIF were delivered to these teachers by the faculty secretary. All the teachers who participated in the survey were told that their responses were confidential.

The participants were requested to specify the numbers of their academic publications in the preceding 10 years (2010–2019) within three categories: (i) scholarly articles in journals, (scholarly articles in journals, specify the unapplied

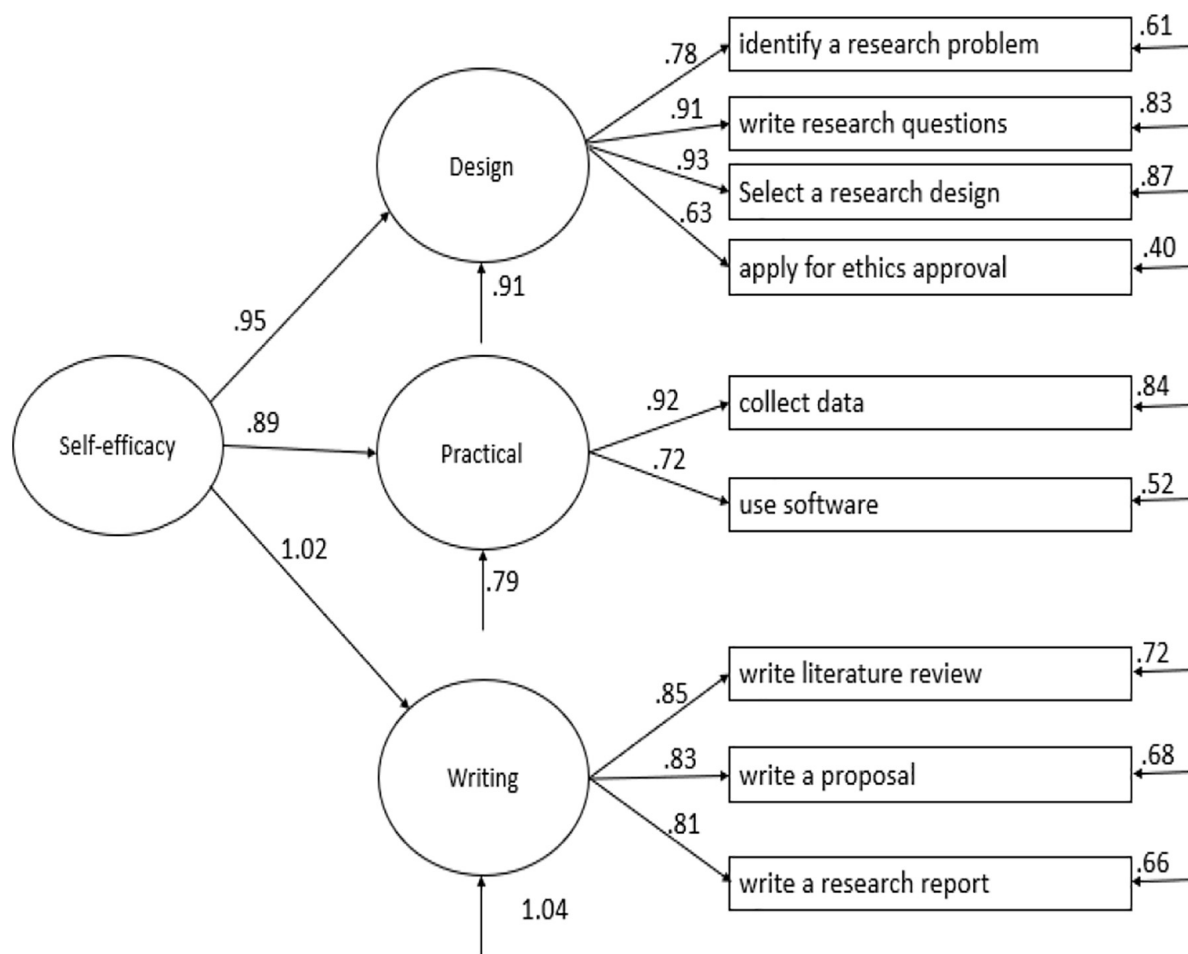


FIGURE 3 | The three-factor correlated model of research self-efficacy. Design, research design skills; practical, practical research skills; writing, writing skills.

projects. The manuscripts that are currently in preparation are excluded from counting toward their total research productivity, as they are not complete scholarly work (Wester et al., 2019). Finally, 536 teachers responded to the questionnaire, among which 508 complete questionnaires were taken as valid and were analyzed.

Analysis

There were two stages in combining the exploratory and confirmatory procedures (Anderson and Gerbing, 1988). Exploratory Factor Analysis (EFA) was used in the self-development scale *Questionnaire on Institutional Support for Teacher Research* to examine the reliability and validation of the construct. The measurement model was determined using EFA by performing maximum likelihood extraction and oblique rotation (Costello and Osborne, 2005). After that, Confirmatory Factor Analysis (CFA) was applied in these four scales: *Teacher Research Self-efficacy*, *Questionnaire on Teacher Research Motivation*, *Questionnaire on Institutional Support for Teacher Research*, and *Questionnaire on Teachers' Research Productivity* to evaluate the factorial and construct validity for each scale within the measurement model. The total sample (546) was randomly divided into two equal halves with one half being used for exploratory factor analysis (EFA; $n = 273$) and the other half for confirmatory factor analysis (CFA; $n = 273$). The EFA and CFA were performed by SPSS 27 and AMOS 27 separately. As recommended by the previous researchers, a model did not need to be rejected if the following conditions were satisfied (Hu and Bentler, 1999; Fan and Sivo, 2005):

- 1) χ^2 per degree of freedom was statistically non-significant (i.e., $\chi^2/df \leq 3.83$),
- 2) comparative fit index (CFI) and Gamma hat > 0.90 ,
- 3) Root mean square error of approximation (RMSEA) < 0.08 , with 90% confidence interval being less than 0.08, and
- 4) Standardized root mean residual (SRMR) < 0.08 .
- 5) The bootstrapping technique was used to test the mediation effect of motivation and self-efficacy.

The concrete data collection and analysis procedures are displayed in **Figure 1**.

RESULTS

This section presents the demographic information of the participants, the results of instrument validation and findings about EFL teachers' research productivity. It also shows how research productivity was affected by the working environment both directly and indirectly, mediated by teachers' intrinsic motivation, extrinsic motivation, and self-efficacy.

Demographic Information of the Participants

This study recruited a total of 536 Chinese EFL teachers on a voluntary basis. As expected, more females (380, 74.8%)

TABLE 4 | CFA regression weights for the two-factor correlated model of institutional support.

Institutional support	Unstandardized estimate	Standardized estimate	C.R.
Item3-M1	1.000 ^a	0.900	a
Item4-M2	0.966	0.920	33.259***
Item5-M3	0.989	0.926	33.839***
Item6-M4	0.927	0.867	28.831***
Item11-WE1	1.000 ^a	0.834	a
Item12-WE2	0.983	0.772	20.724***
Item13-WE3	1.117	0.904	26.771***
Item14-WE4	1.128	0.909	27.014***
Item15-WE5	1.076	0.886	25.815***
Item16-WE6	1.037	0.862	24.629***

*** $p < 0.001$; "a" means the regression weight was fixed at 1.00 for model identification purposes; hence no critical ratio was computed. M, mentorship; WE, working environment.

TABLE 5 | Inter-correlation of the two-factor institutional support.

Factor	Inter-correlation				
	1	2	M	SD	Effect size
1. Mentorship	1		3.42	1.42	0.17
2. Working environment	0.699	1	3.18	1.36	

than males (128, 25.2%) took part in the survey. This is a reflection of the EFL teacher composition in the Chinese education system, where there were more female teachers than male teachers. The age group of teachers between 31 and 40 (38.8%) almost equaled teachers between 41 and 50 (39.2% of the total 536 teachers), constituting 80% of the participants, then the remaining 13.2% and 8.9% of the respondents formed by those under 30 or over 51 separately. Lecturers (226, 44.5%) and associate professors (195, 38.4%) accounted for almost four-fifths of the total 536 participants, and assistant lecturers took up only a tiny percentage, which was twice as many as the full professors in this study. **Table 1** displays the specific participant information.

Measurement Models of Variables

Questionnaire on Institutional Support for Teacher Research

Descriptive statistics showed that the average mean scores of these 16 items ranged from 2.71 (SD = 1.60) to 3.79 (SD = 1.39). The skewness and kurtosis indices were between the cutoff value of $|3.0|$ and $|8.0|$ separately, indicating the normal distribution for the exploratory analysis (Kline, 2016). **Supplementary Appendix 1** shows the descriptive analysis of the *Questionnaire on Institutional Support for Teacher Research* with 16 items.

To conduct exploratory factor analysis (EFA), sampling adequacy was verified with Kaiser-Meyer-Olin by $KMO = 0.956$. Bartlett's test of sphericity ($df = 120$, $p < 0.001$) indicates that correlations between items were sufficiently large for an EFA. Maximum likelihood (ML) estimation was employed on

TABLE 6 | CFA regression weights for the three-factor correlated model of research self-efficacy.

Research engagement	Unstandardized estimate	Standardized estimate	C.R.
Item1-D1	0.960	0.779	23.589***
Item4-D2	1.000 ^a	0.915	a
Item5-D3	1.010	0.938	36.127***
Item9-P1	1.000 ^a	0.846	a
Item10-P2	1.084	0.954	30.092***
Item11-P3	1.045	0.921	28.496***
Item3-W1	1.000 ^a	0.871	a
Item6-W2	0.938	0.823	23.459***

*** $p < 0.001$; “a” means the regression weight was fixed at 1.00 for model identification purpose hence no critical ratio was computed. Design, research design skills; practical, practical research skills; writing, writing skills.

the 16 items *via* oblique rotation with Kaiser Normalization, which analyzed the underlying factors that were assumed to be correlated (Field, 2019). The parallel analysis was used to retain components and evaluate the internal reliability of this questionnaire. No factor was removed in this stage. Two predominant factors with more than three indicators each were extracted, explaining 73.62% of the variance.

The two factors were labeled as Factor 1 Mentorship (63.77% variance); Factor 2 Working Environment (9.84% variance). Cronbach's alpha coefficient for the two factors ranged from 0.944 for Factor 2 to 0.949 for Factor 1. The internal consistency for the two factors met the benchmark value for satisfactory reliability (≥ 0.70), supporting the significant indicator-construct relationship of the instrument. **Table 2** shows the factor loadings and the internal reliability of the two-factor scale.

Confirmatory factor analysis (CFA) was used to test the factor structure. The two-factor structure generated in EFA with maximum likelihood (ML) estimator was examined by CFA. A correlated model was constructed on the basis of the EFA results. To improve the modification indices (MI), some adjustments were made to the original factorial structure in turn. By doing so, six items were deleted. These deleted items were shown as follows: Item 1: In my department I have been, or had been, formally assigned an advisor or mentor to help me in research; Item 2: In my department mentors provide emotional and professional support to junior faculty in times of need; Item 7: My department provides many ways for junior teachers to communicate with experienced

scholars; Item 8: My department provides training for me to get skills and knowledge to do research; Item 9: My department provides latest literature for me to do research; Item 10: My department provides access to external research resources for me to conduct my research. After the removal of the factors, the final model with acceptable fit indices was produced (see **Table 3**). **Figure 2** illustrates the factor structures of the above-tested models. **Table 4** shows regression weights of the two-factor correlated model of institutional support. The mean, standard deviation, effect size for the difference in means, and inter-correlation between the two factors in the final IS Model (Institutional Support Model) are given in **Table 5**. As depicted in **Table 5** and **Figure 2**, the correlation matrix showed that the two factors were significantly correlated with strong degrees in a positive direction. The satisfactory levels of correlations verified that these factors were distinct enough but also under the same theoretical construct of institutional support, confirming the discriminant validity.

The results of EFA and CFA supplied substantial evidence for the factorial structure of the questionnaire, involving mentorship and working environment. The medium strength of the correlations between the two constructs was distinguished but correlated under the same construct of institutional support.

Mentorship

The first dimension was defined as mentorship. In this study, the mentorship was measured by four items (e.g., In my department senior faculty help junior faculty in collaborating on writing papers; In my department senior faculty help junior faculty in writing for grants.) for different aspects of mentorship. However, mentorship was paid insufficient attention in the study of the effect of institutional support on research. Therefore, the exploration of the relationship between mentorship and research engagement was in need. Hopefully, the findings of this study would add up more evidence on the relationship of the two factors.

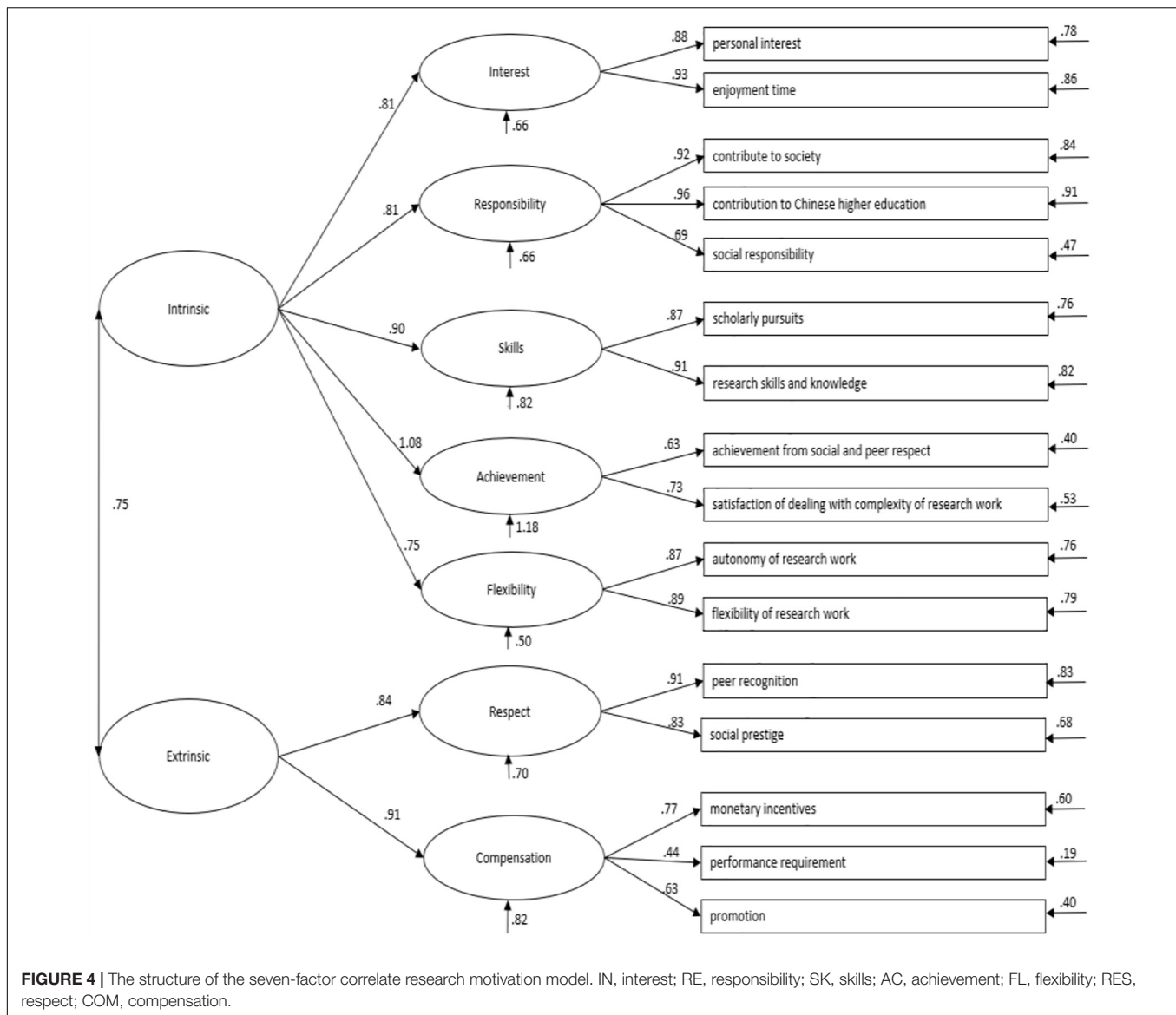
Working Environment

The second dimension, labeled as the working environment, refers to provided time, funding, technical expertise, and assistance support from the institution. In this study, the working environment of the EFL teachers in China was investigated through six items (e.g., My department provides resources such as the computer, scanner, internet, electricity, photocopier, books, etc.) for me to conduct my research; My

TABLE 7 | Model fit indices of research self-efficacy model.

Model	Description	χ^2 (df)	χ^2 /df	CFI	Gamma hat	RMSEA	90% CI	SRMR	AIC
RS	RS model (Three-factor, 8 items)	52.221 (17)	3.072***	0.99	0.98	0.064	0.045–0.084	0.018	90.221

CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual; CI, confidence interval; AIC, akaike information criterion; *** $p < 0.001$.



department provides enough research funding to do research). Previous literature has found that working environments affected research productivity of university teachers (Wilkins, 2011). Therefore, this empirical exploration of EFL teachers' working environment was expected to offer insight into how to provide effective support for improving EFL teachers' research productivity.

Questionnaire on Teacher Research Self-Efficacy

The *Questionnaire on Teacher Research Self-efficacy (QTRSE)* was adapted from *The Self-Efficacy in Research Measure (SERM)* (Phillips and Russell, 1994) to investigate three types of research self-efficacy. As shown in previous literature, the *SERM* has been widely validated in many research settings with sound psychometric properties, this study directly applied CFA to evaluate the validity of the modified instrument *QTRSE* in EFL environments.

Results of confirmatory factor analysis (CFA) produced interesting findings. Based on the previous literature and the theoretical framework of the *SERM*, we hypothesized a three-factor structure as designed in the *SERM* in general academic contexts, which eliminated the Quantitative and Computer Skills factor from the original questionnaire because both quantitative and qualitative skills were employed widely among Chinese EFL teachers, and therefore it unfitted the Chinese context. After that, 10 items were deleted during the preparatory stage as the pilot 60 teachers deemed these items were not suitable in the Chinese context. Finally, there were 15 items loaded onto three factors: Research Design Skills, Practical Research Skills, and Writing Skills. The inspection of modification indices suggested the possibly mis-specified items. The initial model was respecified by removing seven questionable items successively to improve the model fit. Finally, the three-factor model with eight items was defined as the most appropriate research self-efficacy model.

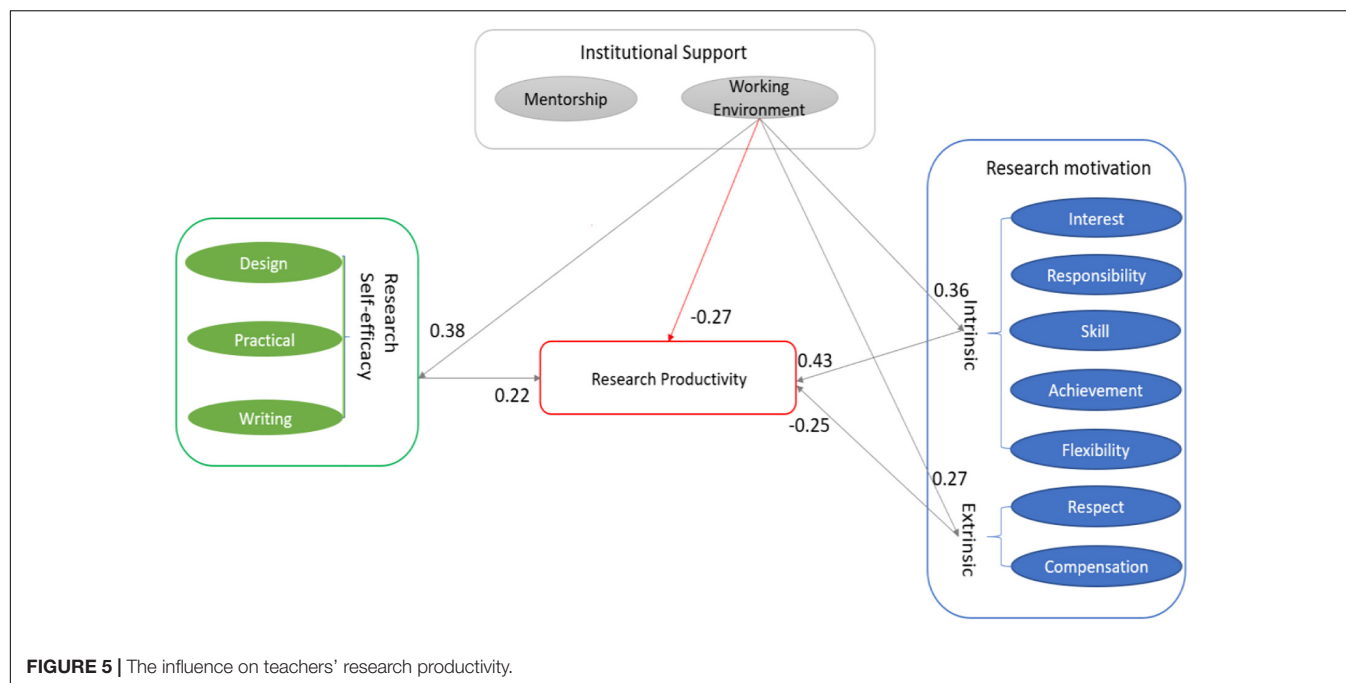


Figure 3 shows the correlation model of research self-efficacy. There was a concern with the desired minimum number of three items under each factor in CFA. However, the two-item per factor model could be justified as reliable when the two variables are highly correlated with each other ($r > 0.70$) (Yong and Pearce, 2013). In the present case, the correlation between items W1 and W2 is 0.72, which suggested that the two items could adequately measure the desired factor of writing skills. The factor structures of the above-discussed model are shown in **Figure 3**. In this correlated model, standardized estimates loadings of all 8-item on the hypothesized latent constructs were higher than 0.70 (see **Table 6**). **Table 7** shows the CFA results of research self-efficacy.

The mean, standard deviation, effect size for the difference in means, and inter-correlation between the three factors in research self-efficacy are given in **Table 8**.

Questionnaire on Teacher Research Motivation

Confirmatory factor analysis was used to evaluate the initial seven-factor 22-item *Questionnaire on Teacher Research Motivation*. The examination of the initial model showed the unaccepted model fit indices. To achieve acceptable model fit, three items were removed from the original model in turn. The final well-fitted model rm model was the most parsimony model with 19 items belong to seven factors. **Table 9** provides a summary of the model fit indices of the final model of research motivation. Generally, one factor should be measured by at least three items in a scale, as the exception, scales measured more than one factor would be identified with minimum two items per factor (Raubenheimer, 2004), meanwhile, the two items should be highly correlated ($r > 0.70$) (Worthington and Whittaker, 2006). In *QTRM*, the correlation between items SK1 and SK2 is 0.79 while for FL1 and FL2 is 0.77,

TABLE 8 | Inter-correlation of the three-factor research self-efficacy.

Factor	Inter-correlation				
	1	2	3	M	SD
1. Design	1			3.99	1.03
2. Practical	0.773	1		3.63	1.14
3. Writing	0.956	0.772	1	3.92	1.01
Effect sizes					
Design compared to				Writing compared to	
Practical	0.33			Practical	
Writing	0.07			0.27	

which implied the two items could adequately measure the desired factor of skills and flexibility. As to AC1 and AC2, the correlation between them is only 0.68, however, this is close to 0.70 showing strong correlation of the two items and arguably acceptable. When speaks the factor loading, if the sample size was over 350, then the factor loading of 0.30 can be acceptable (Yari et al., 2014). Thus, item COM3 (0.37) was accepted in this study.

Figure 4 shows the structure of the Seven-factor Correlate Model. In this correlated model, standardized estimates loadings of all 19-item on the hypothesized latent constructs were almost higher than 0.50 beside COM3 (0.37) (see **Table 10**). The mean, standard deviation, effect size for the difference in means, and inter-correlation between the seven factors in research motivation were given in **Tables 11, 12**.

Questionnaire on Teachers' Research Productivity

Because research productivity is a single factor-dependent variable, the measurement test (CFA) showed a saturated model,

TABLE 9 | Model fit indices of research motivation model.

Model	Description	χ^2 (df)	χ^2/df	CFI	Gamma hat	RMSEA	90% CI	SRMR	AIC
RM	RM model (Seven-factor, 19 items)	505.947 (131)	3.862***	0.95	0.93	0.075	0.068–0.082	0.055	623.947

CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual; CI, confidence interval; AIC, akaike information criterion; *** $p < 0.001$.

TABLE 10 | CFA regression weights for the seven-factor correlated model of research motivation.

Research engagement	Unstandardized estimate	Standardized estimate	C.R.
Item1-IN1	1.000 ^a	0.880	a
Item2-IN2	1.047	0.916	27.780***
Item3-IN3	0.747	0.751	20.422***
Item8-RE1	1.000 ^a	0.921	a
Item9-RE2	1.032	0.954	36.450***
Item10-RE3	0.710	0.683	18.913***
Item4-SK1	1.000 ^a	0.871	a
Item5-SK2	0.955	0.915	20.077***
Item6-AC1	1.000 ^a	0.850	a
Item7-AC2	0.939	0.803	21.129***
Item13-FL1	1.000 ^a	0.849	a
Item14-FL2	1.074	0.912	21.548***
Item15-RES1	1.000 ^a	0.811	a
Item16-RES2	1.189	0.936	25.935***
Item17-RES3	1.141	0.914	25.191***
Item18-COM1	1.000 ^a	0.788	a
Item19-COM2	0.949	0.686	14.554***
Item20-COM3	0.398	0.371	7.731***
Item22-COM4	0.809	0.548	11.536***

*** $p < 0.001$; "a" means the regression weight was fixed at 1.00 for model identification purpose hence no critical ratio was computed. IN, interest; RE, responsibility; SK, skills; AC, achievement; FL, flexibility; RES, respect; COM, compensation.

TABLE 11 | Inter-correlation of the seven-factor research motivation.

Factor	Inter-correlation							M	SD
	1	2	3	4	5	6	7		
1. Interest	1							4.01	1.28
2. Responsibility	0.662	1						4.17	1.16
3. Skills	0.756	0.758	1					4.49	1.21
4. Achievement	0.651	0.774	0.878	1				4.63	1.15
5. Flexibility	0.628	0.605	0.606	0.604	1			4.17	1.21
6. Respect	0.485	0.603	0.551	0.731	0.606	1		4.15	1.10
7. Compensation	0.473	0.580	0.614	0.728	0.663	0.776	1	4.42	0.97

and goodness of fit tests was not available. As a saturated model, it has as many parameters as data points to which it is fitted with zero degrees of freedom (Agresti, 2002). For this kind of model, there was no estimation on model fit indices but focused on only the path coefficient (Steeger and Gondoli, 2013). All of the three observed variables reflected research

productivity positively. Projects had the most robust relationship with research productivity ($r = 0.71$, $p < 0.001$), and articles ($r = 0.67$, $p < 0.001$) impacted research productivity relatively lower than projects. However, articles had a much greater influence on research productivity than conferences ($r = 0.51$, $p < 0.001$).

TABLE 12 | Effect sizes of the seven-factor research motivation.

Factors	Effect sizes	Factor	Effect sizes	Factor	Effect sizes	Factor	Effect sizes	Factor	Effect sizes
Interest compared to		Responsibility compared to		Skills compared to		Achievement compared to		Flexibility compared to	
Responsibility	−0.13	Skills	−0.27	Achievement	−0.12	Flexibility	0.39	Respect	0.02
Skills	−0.38	Achievement	−0.40	Flexibility	0.26	Respect	0.43	Compensation	−0.23
Achievement	−0.51	Flexibility	0	Respect	0.29	Compensation	0.20		
Flexibility	−0.13	Respect	0.02	Compensation	0.06			Respect compared to	
Respect	−0.12	Compensation	−0.23					Compensation	−0.26
Compensation	−0.36								

TABLE 13 | Results of the Structural Equation Modeling (SEM) for research productivity.

Model	Description	χ^2 (df)	χ^2/df	CFI	RMSEA	90% CI	TLI	Gamma hat	SRMR
RP	RP model (Four-factor, 40 items)	1553.041 (678)	2.291***	0.95	0.050	0.047–0.054	0.94	0.92	0.0531

***Means the result is significant because it is within the range of $2 \leq \chi^2/df \leq 5$.

TABLE 14 | The statistical value of the effect analysis on research productivity.

Dependent variables		Research productivity		
Independent variables	Total effect (TE)	Indirect effect (IE)	Direct effect (DE)	
Intrinsic motivation	0.427***	0.00	0.427***	
Extrinsic motivation	−0.253***	0.00	−0.253***	
Self-efficacy	0.217***	0.00	0.217***	
Mentorship	0.130	0.032	0.098	
Working environment	−0.098	0.170***	−0.267***	

*** $p < 0.05$.

Structural Model of Direct and Indirect Effect on Research Productivity

Two of the four exogenous variables significantly affected research productivity. Besides, two factors among all the twelve latent variables directly affected research productivity. The working environment influenced research productivity negatively ($\beta = -0.263$, $p < 0.05$). Inversely, teachers' research self-efficacy ($\beta = 0.351$, $p < 0.05$) positively affected research productivity. The research productivity was explained 13.7% ($SMC = 0.137$, $p < 0.001$) in total by this model, and the effect size of this model was 0.159. **Table 13** shows the model fit index of the RP Model (Research Productivity Model). **Table 14** displays the regression weights.

Mediating Roles of Researchers' Motivation and Self-Efficacy on Research Productivity

The working environment, teachers' research motivation, and self-efficacy directly influenced teachers' research productivity. However, teachers' extrinsic research motivation negatively influenced teachers' research productivity in this study. Based on the existing theoretical framework, this study also examined

the influence of institutional support on teachers' research productivity as mediated by teachers' research motivation and self-efficacy. The result shows that motivation and self-efficacy significantly mediate the relationship between working environment and productivity ($\beta = 0.170$, $p < 0.05$, 95% CI, 0.097–0.260). The result revealed that the working environment is significantly related to teachers' research productivity indirectly mediated by motivation and self-efficacy. **Figure 5** displays both the direct and indirect influence of teachers' research motivation, self-efficacy, mentorship, and the working environment on their research productivity.

DISCUSSION

Research motivation and self-efficacy mediate the indirect relationship between the working environment and research productivity, proving H1. It also added evidence that environmental factors have a powerful influence on individual variables (Lent et al., 2000). The intermediary model shows the following two relationships. Firstly, it is found that teachers' research motivation plays a vital role in bridging the working environment and research productivity. Notably, Nguyen's

(2021) presented that the perceived institutional support directly affects teachers' research motivation, therefore, impacts teachers' research productivity. Specifically, intrinsic motivation positively mediates this relationship, suggesting that improving the working environment encourages teachers to produce more. It is because the provided convenient environment motivates teachers to engage in research more easily and efficiently. Naturally, the research outcomes would increase. Corroborating the findings of Trembley et al. (2009), enhancing intrinsic motivation leads to increasing research productivity. Inversely, extrinsic motivation plays a negative mediation role. Teachers would decline their research engagement whenever they achieve their goals successfully (Santo et al., 2009) or feel hopeless to fulfill the career objectives. This may be caused by teachers considering the rewards for research never appeal to them. Teachers engage less in research because the inadequate working environment declines their extrinsic motivation for research, reducing the research outputs. Another possible explanation is that the insufficient support from the institution makes teachers confront intellectual and financial challenges, and they had to squeeze time to do research (Gao et al., 2010). Those teachers might give up research if they did not internally motivated. Therefore, it can be argued that it is a wise choice for universities to stimulate teachers' intrinsic motivation to do research rather than the external drive. Second, it is revealed that self-efficacy positively mediates the working environment and research productivity. Corroborating Nguyen's (2021) findings, the researcher found, with the advanced working environment, teachers are more confident in doing research. Promoting the working environment increases research self-efficacy and then enhances individuals' research productivity. No matter what and when they need research support, teachers could get smoothly from their department, which would improve their confidence level for research. The confidence level of teachers is in proportion to their research engagement. Undoubtedly, it enhances teachers' research productivity indirectly.

As regards the direct relationship in this study, the current findings suggest that mentorship is not significantly correlated with teachers' research productivity corresponding with Simmering's (2007) finding. That may be because no mentors help teachers publish, and teachers get little help from experienced academics in Chinese higher education institutions. Teachers in Chinese universities are used to doing research alone because of the fiercely competitive relationship among colleagues. The institutional practice of research appraisal gives credit to only the first author, which might be a significant reason that deters Chinese university teachers from collaboration or co-authoring research publications. Hence, mentoring does not have any social role in this case. However, the collaboration of research between scholars within or without the same institutions is a trend worldwide (Paul and Mukhopadhyay, 2022). Therefore, Chinese scholars are inspired to consider cooperating with other scholars in publishing to improve their own rating and the ranking of their institutions. This study also found that the working environment is negatively correlated with teachers' research productivity. The negative association

between working environment and research productivity may imply that when teachers are pressured to publish whereas publication beyond their reach, they might be surrender for having little support from their department. Compared with other universities, the universities in question provide inadequate support for teachers to do research. Firstly, these teachers had to squeeze time to do research and pay money for publishing by themselves. Secondly, the research fundings were limited, and the number of applicants was far more than the projects. At last, the sources of research, such as online databases, conferences, workshops, etc., are insufficient in China. Teachers struggle for publishing when they have to work alone, which restricts their research productivity as seen from the number of research outputs.

CONCLUSION AND IMPLICATIONS

This study examined individual factors as the mediating mechanism, by which institutional factors affected teachers' research productivity. First of all, it was indicated that the working environment was negatively associated with their productivity. Mentorship had no relationship with teachers' research productivity. Second, it was revealed that teachers' research motivation was crucial in bridging working environment and research productivity; similarly, teachers' research self-efficacy had the same mediating function. Hence, the result supported hypothesis 1 that motivation and self-efficacy mediated the working environment and research productivity. Therefore, a supportive work environment where teachers feel comfortable doing research multiplies the effect of self-efficacy and intrinsic motivation on teachers' research productivity. In efforts to improve teachers' research productivity, higher education institutions need to provide adequate institutional support to these university teachers. It is crucial for teachers working in the Chinese context, where they have very few publications and are therefore less frequently cited (Luo and Hyland, 2016), to enhance their work environments to improve research productivity. Teachers gradually increase their confidence in doing research from the improved research environment, which would promote their enthusiasm for research. Institutions would also benefit from providing an appropriate research environment to teachers to increase their research productivity (Gao et al., 2010). This serves another purpose, namely, to boost the institutional ranking of their universities. This means that administrators should offer different support according to the practical situations of their institutions.

Despite the interesting findings, we need to point out the limitations of this study. First, the participants in our study are Chinese EFL teachers, which means that our sample does not represent other groups of teachers. The findings should also be interpreted with this limitation in mind. We encourage colleagues who are interested in the findings to replicate this study in other disciplines or in different contexts. Second, research support in various universities in China is different, and we need to state that these differences could not be reflected through the

questionnaire. Future studies might need to use additional techniques for collecting the data to investigate how these differences are played out. Possible tools could include interviews, observation, or other qualitative methods that describe the different types of research support for teachers in detail and depth. Third, this study only clarifies the mediated function of motivation and self-efficacy. Further studies are needed to examine the mediation role of other individual factors.

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 Auckland Ethics Committee on Human Participants. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Aboagye, E., Jensen, I., Bergström, G., Brämberg, E. B., Pico-Espinosa, O. J., and Björklund, C. (2021). Investigating the association between publication performance and the work environment of university research academics: a systematic review. *Scientometrics* 126, 3283–3301. doi: 10.1007/s11192-020-03820-y
- Agresti, A. (2002). *Categorical Data Analysis*, 2nd Edn. Hoboken, NJ: John Wiley & Sons, Inc. 320–332. doi: 10.1002/0471249688
- Ajjawi, R., Crampton, E. S. P., and Rees, C. E. (2018). What really matters for successful research environments? a realist synthesis. *Med. Educ.* 52, 936–950. doi: 10.1111/medu.13643
- Allen, J. L., Huggins-Hoyt, K. Y., Holosko, M. J., and Briggs, H. E. (2018). African american social work faculty: overcoming existing barriers and achieving research productivity. *Res. Soc. Work Pract.* 28, 309–319. doi: 10.1177/1049731517701578
- Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychol. Bull.* 103, 411–423. doi: 10.1037/0033-2909.103.3.411
- Bandura, A., and Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognit. Ther. Res.* 1, 287–310. doi: 10.1007/BF01663995
- Bard, C. C., Bieschke, K. J., Herbert, J. T., and Eberz, A. B. (2000). Predicting research interest among rehabilitation counseling students and faculty. *Rehabil. Couns. Bull.* 44, 48–55. doi: 10.1177/003435520004400107
- Bieschke, K., Herbert, J., and Bard, C. (1998). Using a social cognitive model to explain research productivity among rehabilitation counselor education faculty. *Rehabil. Educ.* 12, 1–16.
- Blackburn, R. T., and Lawrence, J. H. (1995). *Faculty at Work: Motivation, Expectation, Satisfaction*. Baltimore, MD: The Johns Hopkins University Press.
- Blackmore, P., and Kandiko, C. B. (2011). Motivation in academic life: a prestige economy. *Res. Post Compulsory Educ.* 16, 399–411. doi: 10.1080/13596748.2011.626971
- Bland, C. J., Weber-Main, A. M., Lund, S. M., and Finstad, D. A. (2005). *The Research-Productive Department: Strategies from Departments that Excel*. Bolton, MA: Anker Publishing.
- Borg, S. (2007). Research engagement in english language teaching. *Teach. Teach. Educ.* 23, 731–747. doi: 10.1016/j.tate.2006.03.012
- Borg, S. (2009). English language teachers' conceptions of research. *Appl. Linguist.* 30, 358–388. doi: 10.1093/applin/amp007

AUTHOR CONTRIBUTIONS

YL designed this study, collected the data, and wrote the first draft. LZ contributed to the interpretation of results and revision of the subsequent versions of the manuscript prior to its submission as the corresponding author. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was based on part of the data for the YL's Ph.D. research, funded by a joint doctoral scholarship awarded to YL by The University of Auckland and the China Scholarship Council, Ministry of Education of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.906932/full#supplementary-material>

- Borg, S. (2010). Language teacher research engagement. *Lang. Teach.* 43, 391–429. doi: 10.1017/S0261444810000170
- Borg, S., and Liu, Y. (2013). Chinese college English teachers' research engagement. *TESOL Q.* 47, 270–299. doi: 10.1002/tesq.56
- Brewer, P. D., and Brewer, V. L. (1990). Promoting research productivity in colleges of business. *J. Educ. Bus.* 66, 52–57. doi: 10.1080/08832323.1990.10117453
- Campbell, T. A., and Campbell, D. E. (1997). Faculty/student mentor program: effects on academic performance and retention. *Res. High. Educ.* 38, 727–742. doi: 10.1023/A:1024911904627
- Chen, Y. N., Gupta, A., and Hoshower, L. (2006). Factors that motivate business faculty to conduct research: an expectancy theory analysis. *J. Educ. Bus.* 81, 179–189. doi: 10.3200/JOEB.81.4.179-189
- Chen, Y., Nixon, M. R., Gupta, A., and Hoshower, L. (2010). Research productivity of accounting faculty: an exploratory study. *Am. J. Bus. Educ.* 3, 101–113. doi: 10.19030/ajbe.v3i2.389
- Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* 10, 1–9.
- Creamer, E. G. (1998). *Assessing Faculty Publication Productivity: Issues of Equity*. Graduate School of Education and Human Development. Washington, DC: The George Washington University.
- Creswell, J. W. (1985). *Faculty Research Performance: Lessons from the Science and the Social Science*. (ASHE-ERIC Higher Education Report No. 4). Washington, DC: Association for the Study of Higher Education.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: K.Oxford University Press.
- Dörnyei, Z., and Ottó, I. (1998). *Motivation in Action: A Process Model of L2 Motivation*. Working Papers in Applied Linguistics, Vol. 4. London: Thames Valley University. 43–69.
- Dörnyei, Z., and Taguchi, T. (2010). *Questionnaires in Second Language Research*. 2nd Edn. Abingdon-on-Thames: Routledge. doi: 10.4324/9780203864739
- Dundar, H., and Lewis, H. (1998). Determinants of research productivity in higher education. *Res. High. Educ.* 39, 607–631. doi: 10.1023/A:1018705823763
- Eby, L. T., and Robertson, M. M. (2020). The psychology of workplace mentoring relationships. *Ann. Rev. Organ. Psychol. Organ. Behav.* 7, 75–100. doi: 10.1146/annurev-orgpsych-012119-044924
- Evans, L. (2007). "Developing research cultures and researchers in HE: the role of leadership [Paper presentation]," in *Proceedings of the Annual Conference of the Society for Research into Higher Education (SRHE)*. (Brighton: SRHE). .

- Fan, X., and Sivo, S. A. (2005). Sensitivity of fit indices to misspecified structural or measurement model components: rationale of two index strategy revisited. *Struct. Equ. Modeling* 12, 343–367. doi: 10.1207/s15328007sem1203_1
- Field, A. (2019). *Discovering Statistics Using IBM SPSS Statistics*, 4th Edn. Thousand Oaks, CA: SAGE.
- Fox, M. E. (1985). "Publication, performance, and reward in science and scholarship," in *Higher Education: Handbook of Theory and Research*, Vol. 1, ed. J. C. Smart (New York, NY: Agathon Press), 255–282.
- Fox, M. F. (1992). "Research productivity and the environmental," in *Context research and higher education: The United Kingdom and the United States*, eds T. G. Whiston and R. L. Geiger (Buckingham: The SRHE and Open University Press), 103–111.
- Freedenthal, S., Potter, C., and Grinstein-Weiss, M. (2008). Institutional support for faculty scholarship: a national survey of social work programs. *Soc. Work Res.* 32, 220–230. doi: 10.1093/swr/32.4.220
- Gao, X., and Xu, H. (2014). The dilemma of being English language teachers: interpreting teachers' motivation to teach, and professional commitment in China's hinterland regions. *Lang. Teach. Res.* 18, 152–168. doi: 10.1177/1362168813505938
- Gao, X., Barkhuizen, G., and Chow, A. (2010). "Nowadays, teachers are relatively obedient": understanding primary school English teachers' conceptions of and drives for research in China. *Lang. Teach. Res.* 15, 61–81. doi: 10.1177/1362168810383344
- Gao, X., Barkhuizen, G., and Chow, A. W. K. (2011). Research engagement and educational decentralisation: problematising primary school English teachers' research experiences in China. *Educ. Stud.* 37, 207–219. doi: 10.1080/03055698.2010.506338
- Gelso, C. J., and Lent, R. W. (2000). "Scientific training and scholarly productivity: the person, the training environment, and their interaction," in *Handbook of Counseling Psychology*, eds S. D. Brown and R. W. Lent (Hoboken, NJ: Wiley), 109–139.
- Graves, P., Marchand, J., and Thompson, R. (1982). Economics departmental rankings: research incentives, constraints, and efficiency. *Am. Econ. Rev.* 72, 1131–1141.
- Griffiths, V., Thompson, S., and Hryniewicz, L. (2010). Developing a research profile: mentoring and support for teacher educators. *Prof. Dev. Educ.* 36, 245–262. doi: 10.1080/19415250903457166
- Hemmings, B., and Kay, R. (2016). The relationship between research self-Efficacy, research disposition and publication output. *Educ. Psychol.* 36, 347–361. doi: 10.1080/01443410.2015.1025704
- Heng, K., Hamid, M. O., and Khan, A. (2022). Academics' conceptions of research and the research-teaching nexus: insights from Cambodia. *Int. J. Educ. Dev.* 90, 1–11. doi: 10.1016/j.ijedudev.2022.102569
- Henley, J. S., and Nyaw, M. K. (1986). Introducing market forces into managerial decision-making in Chinese industrial enterprises. *J. Manag. Stud.* 23, 635–656. doi: 10.1111/j.1467-6486.1986.tb00440.x
- Holland, J. L. (1985). *Vocational Preference Inventory Manual*. Lutz, FL: Psychological Assessment Resources.
- Hollingsworth, M. A., and Fassinger, R. E. (2002). The role of faculty mentors in the research training of counseling psychology doctoral students. *J. Couns. Psychol.* 49, 324–330. doi: 10.1037/0022-0167.49.3.324
- Holtum, S., and Goble, L. (2006). Factors influencing levels of research activity in clinical psychologists: a new model. *Clin. Psychol. Psychother.* 13, 339–351. doi: 10.1002/cpp.501
- Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118
- Hu, Q., and Gill, G. T. (2000). Is faculty research productivity: influential factors and implications? *Inf. Resour. Manag. J.* 13, 15–25. doi: 10.4018/irmj.2000040102
- Jang, D., and Shin, I. (2011). The relationship between research self-efficacy and other research constructs: synthesizing evidence and developing policy implications through meta-analysis. *KEDI J. Educ. Policy* 8, 279–301.
- Judge, T., Jackson, C., Shaw, J., Scott, N., and Rich, B. (2007). Self-efficacy and work-related performance: the integral role of individual differences. *J. Appl. Psychol.* 92, 107–127. doi: 10.1037/0021-9010.92.1.107
- Jung, J. (2012). Faculty research productivity in Hong Kong across academic discipline. *High. Educ. Stud.* 2, 1–13. doi: 10.5539/hes.v2n4p1
- Kahn, J. H., and Scott, N. A. (1997). Predictors of research productivity and science-related career goals among counseling psychology doctoral students. *Couns. Psychol.* 25, 38–67. doi: 10.1177/0011000097251005
- Kelly, M. E., and Warmbrod, J. R. (1986). Developing and maintaining productive researchers in agricultural education. *J. Am. Assoc. Teach. Educ. Agric.* 27, 27–32.
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modelling*, 4th Edn. New York, NY: Guilford Press.
- Kozhakhmet, S., Moldashev, K., Yenikeeva, A., and Nurgabdeshev, A. (2022). How training and development practices contribute to research productivity: a moderated mediation model. *Stud. High. Educ.* 47, 437–449. doi: 10.1080/03075079.2020.1754782
- Kyaw, M. T. (2021). Factors influencing teacher educators' research engagement in the reform process of teacher education institutions in Myanmar. *SAGE Open* 11:215824402110613. doi: 10.1177/21582440211061349
- Lai, C. I. (1990). "Comment on the assessment of engineering education in Taiwan," in *The Development Tendency of the Taiwanese Higher Education in the Twenty-first Century-structure, Function, and Organization*, ed. C. P. Chang (Taipei: National Taiwan Normal University), 446–447.
- Landino, R. A., and Owen, S. V. (1988). Self-efficacy in university faculty. *J. Vocat. Behav.* 33, 1–14. doi: 10.1016/0001-8791(88)90030-9
- Lent, R. W., Brown, S. D., and Hackett, G. (2000). Contextual supports and barriers to career choice: a social cognitive analysis. *J. Couns. Psychol.* 47, 36–49. doi: 10.1037/0022-0167.47.1.36
- Lin, B. J., and Fan, L. M. (1990). *Education in Mainland China: Review and Evaluation*. Taipei: Institute of International Relations.
- Liu, F. Y., and Yu, J. L. (2013). "Constructing teacher performance evaluation system about Chinese research universities [Paper presentation]," in *Proceedings of the International Conference on Education, Management and Social Science* 2013.23 doi: 10.2991/icemss.2013.23
- Loewen, S. (2019). The role of research in the second language classroom [Keynote speech]. *Lang. Teach. Res.* Q. 9, 31–38. doi: 10.32038/ltrq.2019.10.03
- Long, R., Crawford, A., White, M., and Davis, K. (2009). Determinants of faculty research productivity in information system: an empirical analysis of the impact of academic origin and academic affiliation. *Scientometrics* 78, 231–260. doi: 10.1007/s11192-007-1990-7
- Love, K. M., Bahner, A., Jones, L., and Nilsson, J. (2007). An investigation of early research experience and research self-efficacy. *Prof. Psychol. Res. Pract.* 38, 314–320. doi: 10.1037/0735-7028.38.3.314
- Lucas, C. J., and Murry, J. W. (2011). *New Faculty: A Practical Guide for Academic Beginners*. London: Palgrave Macmillan. doi: 10.1057/9780230120082
- Luo, N., and Hyland, K. (2016). Chinese academics writing for publication: english teachers as text mediators. *J. Second Lang. Writ.* 33, 43–55. doi: 10.1016/j.jslwr.2016.06.005
- Luthans, F., and Stajkovic, A. D. (1999). Reinforce for performance: the need to go beyond pay and even rewards. *J. Acad. Manag. Exec.* 18, 49–57. doi: 10.5465/ame.1999.1899548
- McGill, M. M., and Settle, A. (2012). Identifying effects of institutional resources and support on computing faculty research productivity, tenure, and promotion. *Int. J. Dr Stud.* 7, 167–198. doi: 10.28945/1581
- McKeachie, W. J. (1982). The rewards of teaching: motivating professors to teach effectively. *New Dir. Teach. Learn.* 10, 7–14. doi: 10.1002/tl.37219821003
- McNurlen, B., and West, C. K. (2000). "Research productivity and tenure. [Paper presentation]," in *Proceedings of the American Educational Research Association*, (New Orleans, LA: American Educational Research Association).
- Morrison, M. A., and Lent, R. W. (2014). The advisory working alliance and research training: test of a relational efficacy model. *J. Couns. Psychol.* 61, 549–559. doi: 10.1037/cou0000030
- Morze, N. V., Buinytska, O. P., and Smirnova, V. A. (2022). Designing a rating system based on competencies for the analysis of the university teachers' research activities. *CEUR Workshop Proc.* 3085, 139–153.
- Mu, C., and Zhang, L. J. (2018). Understanding Chinese multilingual scholars' experiences of publishing research in English. *J. Sch. Publ.* 49, 397–418. doi: 10.3138/jsp.49.4.02

- Nguyen, N. L. (2021). The effects of leader expectation and coworker pressure on research engagement in higher education: the moderating role of achievement value. *J. Appl. Res. High. Educ.* 14, 1114–1126. doi: 10.1108/JARHE-04-2021-0123
- Noorollahi, N. (2021). On the relationship between Iranian English language teaching students-self-efficacy, self-esteem, and their academic achievement. *Lang. Teach. Res. Q.* 21, 84–96. doi: 10.32038/ltrq.2021.21.06
- Nundulall, R., and Dorasamy, N. (2010). Mentorship and sustainable research output: a case study of the University of Johannesburg. *Ind. High. Educ.* 26, 127–137. doi: 10.5367/ihe.2012.0087
- Nundulall, R., and Reddy, K. (2011). Mentorship as a strategy to improve research output at tertiary institutions: a case study of University of Johannesburg. *S. Afr. J. High. Educ.* 25, 1155–1177.
- Nygaard, P. L. (2017). Publishing and perishing: an academic literacies framework for investigating research productivity. *Stud. High. Educ.* 42, 519–532. doi: 10.1080/03075079.2015.1058351
- Pasupathy, R., and Oginga, K. (2014). An investigation of research self-efficacy beliefs and research productivity among faculty members at an emerging research university in the USA. *High. Educ. Res. Dev.* 33, 728–741. doi: 10.1080/07294360.2013.863843
- Paul, P., and Mukhopadhyay, K. (2022). Measuring research productivity of marketing scholars and marketing departments. *Mark. Educ. Rev.* 1–11. doi: 10.1080/10528008.2021.2024077
- Peng, J. E., and Gao, X. (2019). Understanding TEFL academics' research motivation and its relations with research productivity. *SAGE Open* 9, 1–13. doi: 10.1177/2158244019866295
- Phillips, J., and Russell, R. (1994). Research self-efficacy, the research training environment, and research productivity among graduate students in counselling psychology. *Couns. Psychol.* 22, 628–641. doi: 10.1177/0011000094224008
- Raubenheimer, J. (2004). An item selection procedure to maximise scale reliability and validity. *SA J. Ind. Psychol.* 30, 59–64. doi: 10.4102/sajip.v30i4.168
- Reeve, J. M. (1995). *Motivating others: Nurturing inner Motivational Resources*. Boston, MA: Allyn and Bacon.
- Santo, A., Engstrom, M. E., Reetz, L., Schweinle, W., and Reed, K. (2009). Faculty productivity berries and support at a school of education. *Innov. High. Educ.* 34, 117–129. doi: 10.1007/s10755-009-9098-z
- Sato, M., and Loewen, S. (2019). Do teachers care about research? The research–pedagogy dialogue. *ELT J.* 73, 1–10. doi: 10.1093/elt/ccy048
- Shaw, V. N. (2002). Counselling the university professor on the securing of research grants and the publishing of research products. *Education* 123, 395–400.
- Simmering, M. (2007). *Mentoring: Positive Outcomes of Mentoring Relationships: Gender, Diversity, and Mentoring*. Available online at: <http://www.referenceforbusiness.com/management/Mar-No/Mentoring.html> (accessed June 13, 2021).
- Stark, J. S. (1986). “Administrator and faculty views of scholarly performance,” in *Measuring Faculty Research Performance. New Directions for Institutional Research*, ed. J. W. Creswell (Hoboken, NJ: Jossey-Bass). 59–74. doi: 10.1002/ir.37019865006
- Steege, C. M., and Gondoli, D. M. (2013). Mother-adolescent conflict as a mediator between adolescent problem behaviors and maternal psychological control. *Dev. Psychol.* 49, 804–814. doi: 10.1037/a0028599
- Tadesse, E. F., and Khalid, S. (2022). Are teachers and HE are on the same page? Calling for a research–teaching nexus among Ethiopian and Pakistani academics. *J. Appl. Res. High. Educ.* doi: 10.1108/JARHE-09-2021-0348
- Teodorescu, D. (2000). Correlates of faculty publication productivity: a cross national analysis. *High. Educ.* 39, 201–222. doi: 10.1023/A:1003901018634
- Tien, F. F., and Blackburn, R. T. (1996). Faculty rank system, research motivation, and faculty research productivity: measure refinement and theory testing. *J. High. Educ.* 67, 2–22.
- Trembley, M. A., Blanchard, C. M., Taylor, S., and Pelletier, L. G. (2009). Work extrinsic and intrinsic motivation scale: its value for organizational psychology research. *Can. J. Behav. Res.* 41, 213–226. doi: 10.1037/a0015167
- Uwizeye, D., Karimi, F., Thiong'o, C., Syonguvi, J., Ochieng, V., Kiroro, F., et al. (2021). Factors associated with research productivity in higher education institutions in Africa: a systematic review. *AAS Open Res.* 4:26. doi: 10.12688/aasopenres.13211.1
- Vasil, L. (1992). Self-efficacy expectations and causal attributions for achievement among male and female university faculty. *J. Vocat. Behav.* 41, 259–269. doi: 10.1016/0001-8791(92)90028-X
- Way, S. F., Morgan, A. C., Larremore, D. B., and Clauset, A. (2019). Productivity, prominence, and the effects of academic environment. *Proc. Natl. Acad. Sci. U.S.A.* 166, 10729–10733. doi: 10.1073/pnas.1817431116
- West, C. R., Kahn, J. H., and Nauta, M. M. (2007). Learning styles as predictors of self-efficacy and interest in research: implications for graduate research training. *Train. Educ. Prof. Psychol.* 1, 174–183. doi: 10.1037/1931-3918.1.3.174
- Wester, K. L., Gonzalez, L., Borders, L. D., and Ackerman, T. (2019). Initial development of the Faculty Research Self-Efficacy Scale (FaRSES): evidence of reliability and validity. *J. Profess.* 10, 78–99.
- Wilkins, P. (2011). *Research Engagement for School Development. Institute of Education*. London: University of London.
- Williams, H. A., and Kotlik, J. W. (2004). “A mediated hierarchical regression analysis of factors related to career research productivity of human resource education and development postsecondary faculty [Paper presentation],” in *Proceedings of the Academy of Human Resource Development International Conference (AHRD)*, (Austin, TX: AHRD).
- Worthington, R. L., and Whittaker, T. A. (2006). Scale development research. *Couns. Psychol.* 34, 806–838. doi: 10.1177/0011000006288127
- Yari, A., Nadrian, H., Rashidian, H., Nedjat, S., Esmailinasab, N., Doroudi, R., et al. (2014). Psychometric properties of the Persian version of Social Capital Questionnaire in Iran. *Med. J. Islam. Repub. Iran* 28, 17–28.
- Yong, A. G., and Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutor. Quant. Methods Psychol.* 9, 79–94.
- Yuan, R., and Zhang, L. J. (2020). Teacher metacognitions about identities: case studies of four expert language teachers in China. *TESOL Q.* 54, 870–899. doi: 10.1002/tesq.561
- Zhang, L. J. (2021a). Curriculum innovation in language teacher education: reflections on 31 years of the PGDELTA program for teacher continuing professional development. *Chin. J. Appl. Linguist.* 44, 435–450. doi: 10.1515/CJAL-2021-0028
- Zhang, L. J. (2021b). “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. Kulbrandstad, and L. J. Zhang (Bristol, UK: Multilingual Matters). 147–163. doi: 10.21832/9781788928571-011
- Zheng, Y., and Gao, A. X. (2016). Chinese humanities and social sciences scholars' language choices in international scholarly publishing: a ten-year survey. *J. Sch. Publ.* 48, 1–16. doi: 10.3138/jsp.48.1.1

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.



Massive Distance Education: Barriers and Challenges in Shifting to a Complete Online Learning Environment

Ching-Yi Yeh ^{1*} and Chin-Chung Tsai ^{1,2*}

¹ Program of Learning Sciences, National Taiwan Normal University, Taipei, Taiwan, ² Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Benjamin Phipps,
University of New South
Wales, Australia

*Correspondence:

Ching-Yi Yeh
tracyyeh420@gmail.com
Chin-Chung Tsai
tsaicc@ntnu.edu.tw

Specialty section:

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

Received: 26 April 2022

Accepted: 30 May 2022

Published: 23 June 2022

Citation:

Yeh C-Y and Tsai C-C (2022) Massive
Distance Education: Barriers and
Challenges in Shifting to a Complete
Online Learning Environment.
Front. Psychol. 13:928717.
doi: 10.3389/fpsyg.2022.928717

The global pandemic has dramatically changed how the world functions and impacted all sectors of society including all educational institutions. Government and educators respond with immediate online teaching and learning for all students. Massive distance education has been drawn into the picture to provide non-stop learning in most countries worldwide. This study focuses on examining different orders of barriers educators have encountered during the Covid-19 pandemic. The barriers to massive online teaching and learning included the first-order barrier (technological or external barrier), the second-order barrier (internal barrier or teachers' and parents' beliefs), the third-order barrier (design thinking barrier), and the 2.5th order barrier (the classroom management barrier). Both teachers and students are suffering from unstable or limited internet connectivity and it directly hinders students' rights in the massive online education. Teachers are facing the need for sudden pedagogical redesign while parents are enduring the burden of providing all kinds of support for their children's online learning at home. Some learners are experiencing videoconferencing fatigue and struggling with overwhelming resources and an excessive amount of technology time. This study also identifies a group of forgotten learners, the videoconferencing refugees, who have limited access to the Internet and lost their learning opportunities. From a global perspective, shifting to massive online education may be possible with all four orders of barriers being overcome.

Keywords: massive distance education, barriers, videoconferencing fatigue, videoconferencing refugee, online learning, fully online learning

INTRODUCTION

"Lockdown" and "Coronavirus (Covid-19)" have been two of the most popular searched terms on the Internet since 2020. The year 2020 has been an unusual and extremely difficult time for the entire world due to the major outbreak of the covid-19 pandemic (Spinelli et al., 2020). According to the World Health Organization (WHO), there have been more than 300 million confirmed Covid cases globally prior to January 2022. The critical situation has also impacted the educational settings and brought in many unexpected challenges for educators and learners around the world (Dhawan, 2020; Eradze et al., 2021; Muthuprasad et al., 2021). Modern crises such as global warming, pollution, and pandemic, all of which may lead to education fall off in a high-risk society (Pietrocola et al., 2021). Unexpectedly, this global pandemic has affected our living conditions dramatically and changed how we function every day (Spinelli et al., 2020). The sudden dramatic change leaves no

time and space for traditional classroom teaching and learning (Kapasia et al., 2020). School across from K12 to universities around the world such as Australian universities faced a sudden shift to online teaching and learning during the pandemic (Smith and Kaya, 2021). All stakeholders need to keep up with many uncertainties of the pandemic, including policymakers, educators, parents, and learners will now have to quickly respond to teaching, learning, and collaborating online through different online learning platforms and tools (Gonzalez et al., 2020; Junus et al., 2021). During the pandemic, all learning modes are shifted online due to the lockdown of campuses and schools (Dhawan, 2020; Asanov et al., 2021; Muthuprasad et al., 2021). Challenges, barriers, and potential concerns appeared when it comes to shifting all teaching and learning to an online mode. As a result, there is an urgent need for teachers and educators to adopt new ways of teaching (Harsha and Bai, 2020; Junus et al., 2021).

MASSIVE DISTANCE EDUCATION

Distance education across different regions and countries in the world has been utilized to fulfill a small student population's learning need. Students who are unable to attend face-to-face classes can participate in distance education courses (Beldarrain, 2006). Distance education, clearly, should rely on technology-assisted instruction or online education (Beldarrain, 2006). In online education, teachers and individual students interact with each other in online learning environments using the Internet and technology (Beldarrain, 2006), some common technology that supports learning are mobile phones, laptops, tablets, computers, etc. (Singh and Thurman, 2019). Different from distance education which offers small-scale learners experience in the same class at the same time, massive online open courses (MOOCs) intend to offer opportunities and facilitate learning outside of the classroom settings (Kop, 2011), meaning an unlimited number of learners can attend the same course at any time globally (Grover et al., 2013; Joia and Lorenzo, 2021).

Now, as all learners are required to attend classes online worldwide, the status quo becomes a massive distance education setting during the pandemic. Unlike MOOCs that offer certain courses online, thousands of daily routine K-12 classes and higher education courses are pushed online entirely during the pandemic, meaning a remarkable number of learners are currently participating in a massive distance education environment worldwide. With the increasing need for online technology-assisted teaching and learning, teachers and learners encounter some barriers and challenges. Reimers and Schleicher (2020) pointed out the need for redesign in all aspects including the roles of teachers and students, curriculum design, classroom activities, assessment, as well as support for student wellbeing. Furthermore, to maintain non-stop learning in basic education, school districts and universities have considered and included online learning tools such as Zoom or Google Meet (Alameri et al., 2020; Rahiem, 2020; Serhan, 2020; Muthuprasad et al., 2021).

Online learning provides learners and faculty members with another way to maintain education during the pandemic (Olivares et al., 2021). Many instructors in higher education have switched to online teaching and focused on technology integration for students' learning at home in most countries during the pandemic (Alameri et al., 2020). With online learning, the government and school districts hope that students can attend classes without being exposed to any potential viruses and maintain proper social distances (Toquero, 2020; UNESCO, 2020). However, there are some potential difficulties and uncertainties when all teachers and students are forced to learn in the same online mode (Hodges et al., 2020; Kapasia et al., 2020; Zalut et al., 2021). Educators, clearly, encounter barriers and challenges in online teaching and learning settings (Reimers and Schleicher, 2020; Engzell et al., 2021).

BARRIERS TO TECHNOLOGY INTEGRATION IN EDUCATION

Ertmer (1999) proposed a framework elaborating on first-order barriers and second-order barriers for technology integration in education. The first-order barrier includes some external factors that may constrain classroom technology integration, such as lack of adequate access, time, training, and institutional support. These factors are extrinsic to teachers. Furthermore, the author added the second-order barrier, which is more intrinsic to teachers, includes teachers' beliefs in pedagogy, beliefs in technology integration, and teachers' willingness to change (Tsai and Chai, 2012); these are teachers' personal beliefs that may promote or burden the implementation of technology integration in classrooms. In addition, Tsai and Chai (2012) proposed the teachers' design thinking as the third-order barrier to technology integration. Tsai and Chai (2012) further explained that teachers can use design thinking to redesign lessons and offer creative activities to better facilitate different groups of learners' needs.

RE-EXAMINE THE BARRIERS TO MASSIVE DISTANCE EDUCATION

As shown in previous studies, teachers face three orders of barriers in classroom technology implementation. Now, it is time to re-examine the three orders of barriers in massive distance education settings, the first-order barrier—technology integration (extrinsic barrier), the second-order barrier—teachers' personal beliefs (intrinsic barrier), and the third-order barrier—design thinking.

Teachers may run into a variety of difficulties and obstacles as they try to integrate technology into their teaching (Ertmer, 1999). Ertmer (1999) categorized the barriers in two different orders. The first-order barrier in technology integration in classrooms is extrinsic, both preservice and in-service teachers may run into first-order barriers in resources such as available technology, sufficient training, planning time, and relevant administrative support (Ertmer, 1999; Lin et al., 2014). Teachers' technological skills directly impact the effectiveness and quality of online teaching (Danchikov et al., 2021). Teachers are also

reporting barriers to the lack of access to the internet and devices in e-learning implementation (Almanthari et al., 2020). Moreover, teachers do not have sufficient experience in a fully online learning environment (Lase et al., 2021).

In modern education, students' technological skills become essential in obtaining the learning resources (Rasheed et al., 2020). Danchikov et al. (2021) pointed out that students' technological skills affect the effectiveness of their online learning. Limited internet access directly impacts the parents who work from home and homeschool children during the pandemic (Alba and Kang, 2020; Stelitano et al., 2020). Lack of proper devices and stable internet connectivity hinder e-learning at home (Almanthari et al., 2020). In addition, parents have the extra burden to ensure students have all the learning materials ready and set up the technology for their children to attend online classes (Iivari et al., 2020). In the current situation, parents are also experiencing the burden and barrier of providing and maintaining proper online learning technology for their children (Abuhammad, 2020; Aliyyah et al., 2020; Garbe et al., 2020). Students' participation in online learning activities is interrupted due to deficient internet connectivity in rural areas and the slow internet connection frustrates the learners while trying to access the learning platforms and materials (Muthuprasad et al., 2021). Students find it challenging to stay connected and learn online from home (Rahiem, 2020). Therefore, immediately intervention strategies should be considered to help strengthen the communication and collaboration between schools and parents to better facilitate children's learning (Aliyyah et al., 2020; Manca and Meluzzi, 2020). The factors mentioned above are categorized as first-order barriers.

The second-order barrier is intrinsic, including factors such as teachers' personal beliefs about technology integration, willingness to change, and teachers' pedagogical beliefs (Ertmer, 1999; Tsai and Chai, 2012). Unlike traditional classrooms without technology implementation, teachers may be dealing with multiple changes in teaching methods, assessment, and management styles (Kerr, 1996). These second-order barriers are more likely to hinder classroom technology integration as they are deeply rooted in teachers' personal beliefs (Dede, 1998; Ertmer, 1999). Extensive research has also been done in addressing teachers' beliefs and conceptions of the technology-enhanced learning environment (Ellis et al., 2006; González, 2009, 2010; Sherman and Howard, 2012; Schweighofer and Ebner, 2015; Svihla et al., 2015; Saxena, 2017; Durff and Carter, 2019).

During the pandemic, parents and teachers become the essential support for students to foster a proper attitude toward online learning (Manca and Meluzzi, 2020). Parents are having a difficult time becoming teachers at home due to their inability to provide effective distance learning support for their children (Lase et al., 2021). As participants, parents' and students' intrinsic beliefs about online learning may also affect the learning outcome. Parents' assistance in the massive online education becomes crucial; however, they may not see the positive learning outcome of online learning. Students are dealing with learning challenges with limited non-verbal cues in an online learning environment (Khalil et al., 2020). It becomes

challenging for learners to keep learning and stay connected in a comfortable online learning environment (Muthuprasad et al., 2021). In Scull et al. (2020)'s engagement study of Australian university students, they found that students need more guidance on how to ask the right questions and seek help in an online learning environment. In taking distance education courses, major concerns such as time management, motivation, and language skills should be taken into account (Fidalgo et al., 2020). As a result, teachers and school administration at all education levels should seek possible ways to overcome the second-order barrier to further assist the parents and students. These are the crucial factors in the second-order barrier.

In the past, researchers developed very few platforms for totally online teaching. The learning systems and platforms are easier to implement if they could be integrated with traditional face-to-face instruction before the pandemic. Most teachers did not have rich and adequate experiences of totally online education or teaching. Educators are ill-prepared in transitioning into the online teaching setting (Rahiem, 2020). Traditionally, technology integration instruction is designed as a part of face-to-face teaching. Teachers who have experience in teaching in online settings do not encounter as many obstacles as teachers who have none or minimal online teaching experience. Dhawan (2020) stated that teachers will need to find new ways to provide meaningful learning and engage the students in the online setting. Tsai and Chai (2012) pointed out the third-order barrier—the design of teaching strategies. The authors discussed the possibilities of achieving successful technology integration after the first and second-order barriers are being overcome. As the teachers have sufficient technology with adequate pedagogical beliefs, they may still encounter the third-order barrier, which is the need to redesign learning materials to cater to different learners' needs in a completely online setting.

Design thinking skills in education help promote creativity, collaboration, and problem-solving skills (Caruso, 2011; Scheer et al., 2012; Watson, 2015; Henriksen et al., 2017; Lambert et al., 2021; Nguyen et al., 2021). Design thinking is a powerful process to solve problems collaboratively (Deitte and Omary, 2019; Panke, 2019). Tsai and Chai (2012) indicated that design thinking intends to make changes and solve current issues with a creative mindset. With all being said, teachers may face different challenges when it comes to design thinking in online settings (Vallis and Redmond, 2021). Teachers need design thinking skills to overcome the potential challenges of online teaching (Vallis and Redmond, 2021). Muthuprasad et al. (2021) indicated the key factor of a successful online class is interactivity. In addition, they explained that constant meaningful activities help engage the learners in online classes. In massive online education, teachers must create a collaborative online learning environment to enhance the effectiveness of massive online education. Teachers with design thinking skills act as facilitators to provide students with creative learning experiences and guide students to deal with challenges (Noweski et al., 2012; Lambert et al., 2021). The design thinking skill is the third-order barrier.

Apart from all three orders of barriers, Chen et al. (2022) found the 2.5th order barriers: classroom management for totally online teaching. The authors studied the potential barriers to

teachers' use of mobile devices in classrooms. They added that when using the mobile devices to attend online classes, teachers found it challenging to engage the students and maintain their attention. Students were often distracted by other tablet applications or accidentally clicked on the wrong button that led to other matters. This would require the teachers to use a different set of classroom management skills specifically for online teaching. They further explained that as all three barriers are overcome, the barrier of classroom management may still affect teachers' willingness to integrate technology in their classrooms. One of the biggest concerns for K-12 teachers and higher education instructors is that they may not have sufficient educational knowledge for online teaching (Ching et al., 2018). Ghateolbahra and Samimi (2021) indicated teachers need to put in extra effort in dealing with online classroom management to provide meaningful learning. In the totally online teaching environment, classroom management becomes another crucial factor that affects the effectiveness of online teaching and learning. Classroom management in the online education environment leads to a new view in primary schools (Lathifah et al., 2020). Teachers can address the online classroom management challenge by setting up comfortable relationships with the students, paying extra attention to in-class disruption, and having inclusion plans for students with special needs (Baker et al., 2016). Constant monitoring of student practice and effective feedback are also important in an online learning environment (Prilop et al., 2021). The classroom management strategy is the 2.5th order barrier.

VIDEOCONFERENCING FATIGUE AND VIDEOCONFERENCING REFUGEES

When it comes to online learning challenges, the term "digital divide" was highly discussed in the United States and in Europe (Van Dijk and Hacker, 2003). Hargittai (2003) stated that the digital divide is a social issue representing the gap between those who have Internet access and those who do not have any. Rogers (2001) also discussed the extent to which individuals become a disadvantaged group of people in society due to the lack of access to the Internet. During the current pandemic, as all students are pushed to a massive distance learning environment, similar to the digital divide issue, it is unfortunate to see a countless number of students are losing their learning opportunities and falling behind the curriculum in massive distance education. Currently, many countries face the challenges of not having reliable Internet connectivity or sufficient digital devices (Pokhrel and Chhetri, 2021). Some students are facing the problem of restricted or no electricity (Lathifah et al., 2020). On top of internet connectivity concerns, the students are struggling with issues such as technical problems, utilizing online learning strategies, social isolation, stress, etc. (Elmer et al., 2020; Babicka-Wirkus et al., 2021).

As all teaching and learning shift online entirely during the pandemic, an increasing number of students are required to join the online learning environment (Serhan, 2020). Now, with campus lockdown, many higher education institutions and teachers use common online video conferencing tools such as

Google Meets, Microsoft Team, or Zoom (Almendingen et al., 2021; Jindal et al., 2021). Higher education institutions have moved all learning online and utilized web conferencing tools for course content delivery (Bullock et al., 2021). Serhan (2020) states that the use of video conferencing tools is not new in the education field. At the university level, videoconferencing was used during office hours to answer students' questions and concerns regarding course content (Danchikov et al., 2021). During the pandemic, more research attention has been given to the use of video conferencing tools, more specifically, the use of Zoom in K12 to higher education classes (Serhan, 2020; Singhal, 2020; Joia and Lorenzo, 2021; Wiyono et al., 2021). To facilitate learning at home, some school administrators and teachers provide computers, tablets, and Internet access for students to learn online and get through the transition (Serhan, 2020). For students who have sufficient necessary resources at home, they have stable Internet connectivity, computers, laptops, tablets, and even smartphones. These students are the somewhat fortunate ones who have abundant resources. However, these resources can be overwhelming and sometimes disrupting when it is all combined with the online classroom. With the sudden dramatic growth of screen time and intense online learning schedule, the situation can go out of control and easily lead to videoconferencing fatigue for learners in higher education.

Bailenson (2021) defined "Zoom Fatigue" as an exhausting situation where the use of Zoom video conferencing increased dramatically during the pandemic. Fosslien and Duffy (2020) further explain that the intense focus on verbal conversations makes Zoom users debilitating. In K-12 education, all learners have been learning and completing assignments online during the pandemic (García and Weiss, 2020; Serhan, 2020) and there is no foreseeable end date to the current situation. Dorn et al. (2021) point out that students struggle with multiple difficulties in class scheduling, technical issues, and Zoom fatigue. Notably, students are unprepared and overwhelmed with the given information and they may not have all the necessary technological skills to navigate around the online learning platforms. Bullock et al. (2021) mentioned the overuse of technology could lead to extra stress both mentally and physically as more higher education institutions move to online learning. In addition, learners are experiencing a lack of physical interaction such as body gestures and facial expressions or responses in Zoom classes (Peper et al., 2021).

Many students are facing troubles such as attending and logging on to online classes, uploading assignments, getting used to the functions to share their screens or express their opinions, etc. (Goldstein et al., 2020). Moreover, student engagement and learning assessment are unclear in remote learning during the pandemic (Asanov et al., 2021). This group of learners suffer from the overwhelming online learning environment and are potentially experiencing videoconferencing fatigue. Furthermore, they have to stay online all day without any social interaction with their peers (García and Weiss, 2020). Ferri et al. (2020) identified similar situations as social challenges where the students have no social interaction with their teacher or other students and they also lack support from their parents or caretakers who might be working from home in the same space.

There is an urgent need for an appropriate and effective remote teaching and learning plan for both teachers and students.

Educators face the fast change into emergent remote teaching and teachers need to rely on educational tools such as computers/laptops, the Internet, online platforms, social media, etc. to offer a sufficient online learning environment (Svrcek et al., 2022). Besides the challenges and barriers teachers encounter, the student's access to technology at home also affects their learning quality (Muthuprasad et al., 2021). Other than learners who are receiving overwhelming resources and experiencing videoconferencing fatigue, there is another group of learners who are dealing with limited or no access to necessary technology tools or have insufficient internet connectivity and limited access to educational resources (Asanov et al., 2021; Lai and Widmar, 2021). Manca and Meluzzi (2020) pointed out there is a group of underserved students who do not have any access to the necessary technology to attend online classes. In contrast to learners who are experiencing zoom or videoconferencing fatigue, the authors of this paper would call this group of struggling learners the *videoconferencing refugees*.

Videoconferencing refugees, coined by this paper, are the forgotten ones during the pandemic, both emotionally and academically. Unfortunately, they are not just being left behind; they are being forgotten in the online learning environment. As the pandemic transformed learning, the video conferencing refugees lost their fair chances to learn or participate in any learning activities due to insufficient technology, digital resources, or Internet connectivity. Padilla Rodríguez et al. (2021) pointed out that rural area students are suffering from unstable internet connectivity, limited electricity, and unqualified teacher in the online learning environments. The learners do not have any opportunities to continue learning even if they wanted to.

During the pandemic, students from low socioeconomic families are facing the challenge of not having a stable internet connection, proper technology, or parent support (Rahiem, 2020; Azubuike et al., 2021). Learners deal with different levels of online learning difficulties while trying to learn at home (Padilla Rodríguez et al., 2021). Some school districts are providing limited hotspot services for families in rural areas (Lai and Widmar, 2021). The achievement gap among the disadvantaged students widened during the pandemic (Dorn et al., 2021). Kapasia et al. (2020) state that some students are now dealing with an unfavorable learning environment due to the pandemic. From computer technical skills to understanding the academic content in online settings, they have no support and are left with no other choices to maintain basic education; hence, the feeling of isolation and anxiety increases as the lockdown period extends. Online learning can be overwhelming and can lead to extra potential stress on the students (Shanahan et al., 2020). In Shanahan et al.'s study of 768 participants, their results indicate an increasing level of emotional distress during the pandemic. Teachers are reporting that their students are not logging into the online sessions, especially the lower-income students and there is a clear increase in student absences rate in the United States (Goldstein et al., 2020). Videoconferencing refugees do not have stable internet connectivity and access to adequate learning

resources or technology simply lose their fair chance to learn as much as their peers. As mentioned above, there is a remarkable number of students who are now required to learn online worldwide. Other than instructions given by their teachers, the students, and the parents are fighting the battle alone at home.

OVERALL CHALLENGES IN COVID

There are multiple aspects of challenges all stakeholders need to deal with during and after the pandemic. In massive distance education, teachers in K-12 and higher education are facing totally online teaching and lesson redesigning obstacles. Teachers need to find the appropriate ways of expression through the online manner and the usage of different modalities to enrich teaching or course materials. The teaching materials (readings, videos, exercises, etc.) become an influential factor as students are spending more time reading and reviewing the lesson materials on their own at home (Rapanta et al., 2020). Furthermore, the use of various online platforms or the fear of monopolies of certain platforms may affect the quality of totally online teaching and learning. The different platforms offer different interactions among various parties (Kennedy, 2020). K12 to higher education teachers need to know how to effectively use online tools, systems, or modules developed before in such a massive distance education setting. In addition, educators can use new technologies (such as automatic analysis) to detect students' engagement, boredom, frustration, success, or failure in learning. Another challenge is the course redesign for skill-based or internship courses. Educators will have to find ways to accommodate students' learning in skill-based courses in the massive distance education setting. For the skill-based courses, instructors may incorporate immersive virtual reality (VR) as a possible solution to ensure meaningful learning takes place in massive distance education. VR offers faculty and students a meaningful and reality-informed learning experience (Schott and Marshall, 2021). Khalil et al. (2020) mentioned the use of virtual simulation technologies can facilitate clinical practice for medical students. The group chat application is used in sharing information and collaborating globally during the pandemic in medical education; virtual learning can also benefit the fellow-in-training (FIT) in multiple medical areas with proper planning for its implementation (Almarzooq et al., 2020).

As for the learner challenges in massive distance education settings, they may face the expression challenges in using various tools of verbal, written, or visual representations. They will also need extra guidance to reach fluent communication through different modalities such as text, verbal, facial expression, social media, etc. In massive distance education settings, students' self-regulation in both synchronous and asynchronous modes becomes another issue for teachers and parents. Parents' concerns, opinions, and time management toward massive distance education should not be overlooked. In addition, school districts and parents need to consider the investment in the infrastructure of massive distance education and the technical support at different levels including for

learners, teachers, schools, and parents. Overall, the quality of massive distance education and online learning should be reexamined.

FURTHER THOUGHTS IN GLOBAL PERSPECTIVES

As more vaccines are given to citizens worldwide, perhaps we will be seeing the end of this Covid pandemic; yet, will life go back to normal or will this be a constant change? It is time to consider possible ways to overcome the challenging and overwhelming online learning environment to avoid and decrease the level of videoconferencing fatigue. In the transition to massive online education, how do we move forward in higher education? Engzell et al. (2021) mentioned the pandemic has brought significant impact and concerns in students' learning. In their study of national data of The Netherlands students ($n \approx 350,000$), they found that learning from home students make minimal or nearly no learning progress, especially the disadvantaged students. The disadvantaged students are at risk while coping with learning from home. Furthermore, the quality of learning is being impacted without thorough online teaching and learning planning. After the technological and pedagogical barriers are being overcome with proper online classroom management strategies and appropriate design thinking, the quality of massive online education could be raised to a satisfactory level. To build on the pandemic online teaching experience in higher education, we can possibly make extensive use of all the online

learning data collected during the pandemic in the massive distance education to provide more comprehensive support for precision education. The collected data may be studied to improve the quality of massive online education for colleges and universities. As we move along, it is important to keep in mind that videoconferencing refugees in all education levels, from K-12 to university levels should not be held back or forgotten. What will online teaching be like in the near future for higher education institutions? What will online learning be like in the near future for all learners? With all being said, more research should be done to provide a full scope of e-pedagogy in response to such a global crisis.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was financially supported by the Institute for Research Excellence in Learning Sciences 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 and Ministry of Science and Technology, Taiwan (Grant numbers: MOST108-2511-H-003-038-MY3 and MOST 109-2511-H-003-013-MY3).

REFERENCES

- Abuhammad, S. (2020). Barriers to distance learning during the COVID-19 outbreak: a qualitative review from parents' perspective. *Heliyon* 6, e05482. doi: 10.1016/j.heliyon.2020.e05482
- Alameri, J., Masadeh, R., Hamadallah, E., Ismail, H. B., and Fakhouri, H. N. (2020). Students' perceptions of e-learning platforms (Moodle, Microsoft Teams and Zoom platforms) in The University of Jordan Education and its relation to self-study and academic achievement during COVID-19 pandemic. *J. ISSN* 2692, 2800.
- Alba, D., and Kang, C. (2020). So we're working from home. Can the Internet handle it. *The New York Times*, p. 16.
- Aliyyah, R. R., Rachmadtullah, R., Samsudin, A., Syaodih, E., Nurtanto, M., and Tambunan, A. R. S. (2020). The perceptions of primary school teachers of online learning during the COVID-19 pandemic period: a case study in Indonesia. *J. Ethn. Cult. Stud.* 7, 90–109. doi: 10.29333/ejecs/388
- Almanthari, A., Maulina, S., and Bruce, S. (2020). Secondary school mathematics teachers' views on e-learning implementation barriers during the COVID-19 pandemic: the case of Indonesia. *Eurasia J. Math. Sci. Technol. Educ.* 16, em1860. doi: 10.29333/ejmste/8240
- Almarzoq, Z. I., Lopes, M., and Kochar, A. (2020). Virtual learning during the COVID-19 pandemic: a disruptive technology in graduate medical education. *J. Am. Coll. Cardiol.* 75, 2635–2638. doi: 10.1016/j.jacc.2020.04.015
- Almendingen, K., Morseth, M. S., Gjølstad, E., Brevik, A., and Tørris, C. (2021). Student's experiences with online teaching following COVID-19 lockdown: a mixed methods explorative study. *PLoS ONE* 16:e0250378. doi: 10.1371/journal.pone.0250378
- Asanov, I., Flores, F., McKenzie, D., Mensmann, M., and Schulte, M. (2021). Remote-learning, time-use, and mental health of Ecuadorian high-school students during the COVID-19 quarantine. *World Dev.* 138, 105225. doi: 10.1016/j.worlddev.2020.105225
- Azubuikwe, O. B., Adegboye, O., and Quadri, H. (2021). Who gets to learn in a pandemic? Exploring the digital divide in remote learning during the COVID-19 pandemic in Nigeria. *Int. J. Educ. Res. Open* 2, 100022. doi: 10.1016/j.ijedro.2020.100022
- Babicka-Wirkus, A., Wirkus, L., Stasiak, K., and Kozłowski, P. (2021). University students' strategies of coping with stress during the coronavirus pandemic: data from Poland. *PLoS ONE* 16:e0255041. doi: 10.1371/journal.pone.0255041
- Bailenson, J. N. (2021). Nonverbal overload: a theoretical argument for the causes of Zoom fatigue. *Technol. Mind Behav.* 2. doi: 10.1037/tmb0000030
- Baker, C., Gentry, J., and Larmer, W. (2016). A model for online support in classroom management: perceptions of beginning teachers. *Adm. Issues J.: Connect. Educ. Pract. Res.* 6, 22–37. doi: 10.5929/2016.6.1.3
- Beldarrain, Y. (2006). Distance education trends: integrating new technologies to foster student interaction and collaboration. *Distance Educ.* 27, 139–153. doi: 10.1080/01587910600789498
- Bullock, A. N., Colvin, A. D., and Jackson, M. S. (2021). "All Zoomed Out": strategies for addressing Zoom fatigue in the age of COVID-19," in *The Learning Ideas Conference* (Cham: Springer), 61–68. doi: 10.1007/978-3-030-90677-1_6
- Caruso, C. (2011). *The Tools of Engagement: Bridging Design Thinking and Social Media to Enhance and Support Collaborative Learning* (Doctoral dissertation). Carleton University, Ottawa, ON, Canada.
- Chen, C.-H., Siu-Yung Jong, M., and Tsai, C.-C. (2022). A comparison of in-service teachers' conceptions of barriers to mobile technology-integrated instruction and technology-integrated instruction. *Aust. J. Educ. Technol.* 38, 35–50. doi: 10.14742/ajet.7299
- Ching, Y. H., Hsu, Y. C., and Baldwin, S. (2018). Becoming an online teacher: an analysis of prospective online instructors' reflections. *J. Interact. Learn. Res.* 29, 145–168.
- Danchikov, E. A., Prodanova, N. A., Kovalenko, Y. N., and Bondarenko, T. G. (2021). The potential of online learning in modern conditions and

- its use at different levels of education. *Linguistics Cult. Rev.* 5, 578–586. doi: 10.21744/lingcure.v5n51.1442
- Dede, C. (1998). *Learning With Technology: The 1998 ASCD Yearbook*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Deitte, L. A., and Omary, R. A. (2019). The power of design thinking in medical education. *Acad. Radiol.* 26, 1417–1420. doi: 10.1016/j.acra.2019.02.012
- Dhawan, S. (2020). Online learning: a panacea in the time of COVID-19 crisis. *J. Educ. Technol. Syst.* 49, 5–22. doi: 10.1177/0047239520934018
- Dorn, E., Hancock, B., Sarakatsannis, J., and Viruleg, E. (2021). *Covid-19 and Education: The Lingering Effects of Unfinished Learning*. McKinsey and Company.
- Durff, L., and Carter, M. (2019). Overcoming second-order barriers to technology integration in K-5 schools. *J. Educ. Res. Pract.* 9, 17. doi: 10.5590/JERAP.2019.09.1.18
- Ellis, R. A., Steed, A. F., and Applebee, A. C. (2006). Teacher conceptions of blended learning, blended teaching and associations with approaches to design. *Aust. J. Educ. Technol.* 22. doi: 10.14742/ajet.1289
- Elmer, T., Mepharm, K., and Stadtfeld, C. (2020). Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLoS ONE* 15:e0236337. doi: 10.1371/journal.pone.0236337
- Engzell, P., Frey, A., and Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2022376118. doi: 10.1073/pnas.2022376118
- Eradze, M., Bardone, E., and Dipace, A. (2021). Theorising on covid-19 educational emergency: magnifying glasses for the field of educational technology. *Learn. Media Technol.* 46, 404–419. doi: 10.1080/17439884.2021.1961802
- Ertmer, P. A. (1999). Addressing first-and second-order barriers to change: strategies for technology integration. *Educ. Technol. Res. Dev.* 47, 47–61. doi: 10.1007/BF02299597
- Ferri, F., Grifoni, P., and Guzzo, T. (2020). Online learning and emergency remote teaching: opportunities and challenges in emergency situations. *Societies* 10, 86. doi: 10.3390/soc10040086
- Fidalgo, P., Thormann, J., Kulyk, O., and Lencastre, J. A. (2020). Students' perceptions on distance education: a multinational study. *Int. J. Educ. Technol. Higher Educ.* 17, 1–18. doi: 10.1186/s41239-020-00194-2
- Fosslien, L., and Duffy, M. W. (2020). How to combat zoom fatigue. *Harv. Bus. Rev.* 29, 1–6.
- Garbe, A., Ogurlu, U., Logan, N., and Cook, P. (2020). COVID-19 and remote learning: experiences of parents with children during the pandemic. *Am. J. Qual. Res.* 4, 45–65. doi: 10.29333/ajqr/8471
- García, E., and Weiss, E. (2020). *COVID-19 and Student Performance, Equity, and US Education Policy: Lessons From Pre-Pandemic Research to Inform Relief, Recovery, and Rebuilding*. Washington, DC: Economic Policy Institute.
- Ghateolbahra, A., and Samimi, F. (2021). Classroom management strategies in online environment: a comparative study on novice and experienced teachers. *Turk. J. Comput. Math. Educ. (TURCOMAT)* 12, 510–516. doi: 10.16949/turkbilmat.702540
- Goldstein, D., Popescu, A., and Hannah-Jones, N. (2020). As school moves online, many students stay logged out. *The New York Times*, p. 6.
- González, C. (2009). Conceptions of, and approaches to, teaching online: A study of lecturers teaching postgraduate distance courses. *High Educ.* 57, 299–314. doi: 10.1007/s10734-008-9145-1
- González, C. (2010). What do university teachers think eLearning is good for in their teaching. *Study High. Educ.* 35, 61–78. doi: 10.1080/0307570902874632
- Gonzalez, T., de la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., et al. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PLoS ONE* 15:e0239490. doi: 10.1371/journal.pone.0239490
- Grover, S., Franz, P., Schneider, E., and Pea, R. (2013). "The MOOC as distributed intelligence: Dimensions of a framework & evaluation of MOOCs," in *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSDL 2013 Conference Proceedings Volume 2 - Short Papers, Panels, Posters, Demos & Community Events*, eds N. Rummel, M. Kapur, M. Nathan, and S. Puntambekar (Madison, WI: International Society of the Learning Sciences), 42–45.
- Hargittai, E. (2003). The digital divide and what to do about it. *New Econom. Handbook*. 2003, 821–839. Available online at: <http://www.eszter.com/research/pubs/hargittai-digitaldivide.pdf>
- Harsha, R., and Bai, T. (2020). Covid-19 lockdown: challenges to higher education. *Cape Comorin*. 2, 26–28. doi: 10.13140/RG.2.2.16290.25281
- Henriksen, D., Richardson, C., and Mehta, R. (2017). Design thinking: A creative approach to educational problems of practice. *Think. Skill. Create*. 26, 140–153. doi: 10.1016/j.tsc.2017.10.001
- Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., and Bond, M. A. (2020). *The Difference Between Emergency Remote Teaching and Online Learning*. VTechWorks Home. Available online at: <https://vtechworks.lib.vt.edu/handle/10919/104648>
- Iivari, N., Sharma, S., and Ventä-Olkkonen, L. (2020). Digital transformation of everyday life—how COVID-19 pandemic transformed the basic education of the young generation and why information management research should care? *Int. J. Inf. Manag.* 55, 102183. doi: 10.1016/j.ijinfomgt.2020.102183
- Jindal, M., Bajal, E., Singh, P., Diwakar, M., Arya, C., and Sharma, K. (2021). "Online education in Covid-19: limitations and improvements," in *2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (IEEE)*, 1–5. doi: 10.1109/UPCON52273.2021.9667605
- Joia, L. A., and Lorenzo, M. (2021). Zoom in, zoom out: the impact of the COVID-19 pandemic in the classroom. *Sustainability* 13, 2531. doi: 10.3390/su13052531
- Junus, K., Santoso, H. B., Putra, P. O. H., Gandhi, A., and Siswantinings, T. (2021). Lecturer readiness for online classes during the pandemic: a survey research. *Educ. Sci.* 11, 139. doi: 10.3390/educsci11030139
- Kapasia, N., Paul, P., Roy, A., Saha, J., Zaveri, A., Mallick, R., et al. (2020). Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India. *Child. Youth Serv. Rev.* 116, 105194. doi: 10.1016/j.childyouth.2020.105194
- Kennedy, J. (2020). *Monopoly Myths: Do Internet Platforms Threaten Competition?* Washington, DC: Information Technology and Innovation Foundation.
- Kerr, S. T. (1996). Chapter I: Visions of sugarplums: The future of technology, education, and the schools. *Teach. Coll. Record*. 97, 1–27.
- Khalil, R., Mansour, A. E., Fadda, W. A., Almisnid, K., Aldamegh, M., Al-Nafeesah, A., et al. (2020). The sudden transition to synchronized online learning during the COVID-19 pandemic in Saudi Arabia: a qualitative study exploring medical students' perspectives. *BMC Med. Educ.* 20:285. doi: 10.1186/s12909-020-02208-z
- Kop, R. (2011). The challenges to connectivist learning on open online networks: learning experiences during a massive open online course. *Int. Rev. Res. Open Dis. Learn.* 12, 19–38. doi: 10.19173/irrodl.v12i3.882
- Lai, J., and Widmar, N. O. (2021). Revisiting the digital divide in the COVID-19 era. *Appl. Econ. Perspect. Policy* 43, 458–464. doi: 10.1002/aep.13104
- Lambert, R., Imm, K., Schuck, R., Choi, S., and McNiff, A. (2021). "UDL Is the What, Design Thinking Is the How:" designing for differentiation in mathematics. *Math. Teach. Educ. Dev.* 23, 54–77.
- Lase, D., Zega, T. G. C., and Daeli, D. O. (2021). Parents' perceptions of distance learning during Covid-19 pandemic in rural Indonesia. *J. Educ. Learn. (EduLearn)*. 16, 103–113. doi: 10.11591/edulearn.v16i1.20122
- Lathifah, Z. K., Helmanto, F., and Maryani, N. (2020). The practice of effective classroom management in COVID-19 time. *Int. J. Adv. Sci. Technol.* 29, 3263–3271.
- Lin, C. Y., Huang, C. K., and Chen, C. H. (2014). Barriers to the adoption of ICT in teaching Chinese as a foreign language in US universities. *ReCALL* 26, 100–116. doi: 10.1017/S0958344013000268
- Manca, F., and Meluzzi, F. (2020). *Strengthening Online Learning When Schools are Closed: The Role of Families and Teachers in Supporting Students During the COVID-19 Crisis*. Paris: OECD.
- Muthuprasad, T., Aiswarya, S., Aditya, K. S., and Jha, G. K. (2021). Students' perception and preference for online education in India during COVID-19 pandemic. *Soc. Sci. Human. Open* 3, 100101. doi: 10.1016/j.ssaho.2020.100101
- Nguyen, T. H., Pham, X. L., and, Tu, N. T. T. (2021). The impact of design thinking on problem solving and teamwork mindset in a flipped classroom. *Eurasian J. Educ. Res.* 96, 30–50. doi: 10.14689/ejer.2021.96.3

- Noweski, C., Scheer, A., Büttner, N., von Thienen, J., Erdmann, J., and Meinel, C. (2012). "Towards a paradigm shift in education practice: developing twenty-first century skills with design thinking," in *Design Thinking Research* (Springer, Berlin, Heidelberg), 71–94. doi: 10.1007/978-3-642-31991-4_5
- Olivares, S. L. O., Lopez, M., Martinez, R., Alvarez, J. P. N., and Valdez-García, J. E. (2021). Faculty readiness for a digital education model: A self-assessment from health sciences educators. *Aust. J. Technol.* 37, 116–127. Available online at: <https://ajet.org.au/index.php/AJET/article/view/7105/1832>
- Padilla Rodríguez, B. C., Armellini, A., and Traxler, J. (2021). The forgotten ones: how rural teachers in Mexico are facing the COVID-19 pandemic. *Online Learn.* 25, 253–268. doi: 10.24059/olj.v25i1.2453
- Panke, S. (2019). Design thinking in education: perspectives, opportunities and challenges. *Open Educ. Stud.* 1, 281–306. doi: 10.1515/edu-2019-0022
- Peper, E., Wilson, V., Martin, M., Rosegard, E., and Harvey, R. (2021). Avoid Zoom fatigue, be present and learn. *NeuroRegulation* 8, 47–47. doi: 10.15540/nr.8.1.47
- Pietrocola, M., Rodrigues, E., Bercot, F., and Schnorr, S. (2021). Risk society and science education. *Sci. Educ.* 30, 209–233. doi: 10.1007/s1191-020-00176-w
- Pokhrel, S., and Chhetri, R. (2021). A literature review on impact of COVID-19 pandemic on teaching and learning. *High. Educ. Future* 8, 133–141. doi: 10.1177/2347631120983481
- Prilop, C. N., Weber, K. E., and Kleinknecht, M. (2021). The role of expert feedback in the development of pre-service teachers' professional vision of classroom management in an online blended learning environment. *Teach. Teacher Educ.* 99, 103276. doi: 10.1016/j.tate.2020.103276
- Rahiem, M. D. (2020). Technological barriers and challenges in the use of ICT during the COVID-19 emergency remote learning. *Univers. J. Educ. Res.* 8, 6124–6133. doi: 10.13189/ujer.2020.082248
- Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., and Koole, M. (2020). Online university teaching during and after the Covid-19 crisis: refocusing teacher presence and learning activity. *Postdigit. Sci. Educ.* 2, 923–945. doi: 10.1007/s42438-020-00155-y
- Rasheed, R. A., Kamsin, A., and Abdullah, N. A. (2020). Challenges in the online component of blended learning: a systematic review. *Comput. Educ.* 144, 103701. doi: 10.1016/j.compedu.2019.103701
- Reimers, F. M., and Schleicher, A. (2020). *A Framework to Guide an Education Response to the COVID-19 Pandemic of 2020*. OECD. Retrieved April 14(2020), 2020-04.
- Rogers, E. M. (2001). The digital divide. *Convergence* 7, 96–111.
- Saxena, A. (2017). Issues and impediments faced by Canadian teachers while integrating ICT in pedagogical practice. *Turk. Online J. Educ. Technol.-TOJET* 16, 58–70.
- Scheer, A., Noweski, C., and Meinel, C. (2012). Transforming constructivist learning into action: Design thinking in education. *Design Technol. Educ.: Int. J.* 17, 8–19. Available online at: <https://files.eric.ed.gov/fulltext/EJ996067.pdf>
- Schott, C., and Marshall, S. (2021). Virtual reality for experiential education: a user experience exploration. *Australas. J. Educ. Technol.* 37, 96–110. doi: 10.14742/ajet.5166
- Schweighofer, P., and Ebner, M. (2015). Aspects to be considered when implementing technology-enhanced learning approaches: a literature review. *Future Internet* 7, 26–49. doi: 10.3390/fi7010026
- Scull, J., Phillips, M., Sharma, U., and Garnier, K. (2020). Innovations in teacher education at the time of COVID19: an Australian perspective. *J. Educ. Teach.* 46, 497–506. doi: 10.1080/02607476.2020.1802701
- Serhan, D. (2020). Transitioning from face-to-face to remote learning: students' attitudes and perceptions of using Zoom during COVID-19 pandemic. *Int. J. Technol. Educ. Sci.* 4, 335–342. doi: 10.46328/ijtes.v4i4.148
- Shanahan, L., Steinhoff, A., Bechtiger, L., Murray, A. L., Nivette, A., Hepp, U., et al. (2020). Emotional distress in young adults during the COVID-19 pandemic: evidence of risk and resilience from a longitudinal cohort study. *Psychol. Med.* 52, 824–833. doi: 10.1017/S003329172000241X
- Sherman, K., and Howard, S. K. (2012). "Teachers' beliefs about first-and second-order barriers to ICT integration: preliminary findings from a South African study," in *Society for Information Technology and Teacher Education International Conference* (Association for the Advancement of Computing in Education (AACE)), 2098–2105.
- Singh, V., and Thurman, A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988–2018). *Am. J. Dist. Educ.* 33, 289–306. doi: 10.1080/08923647.2019.1663082
- Singhal, M. K. (2020). Facilitating virtual medicinal chemistry active learning assignments using advanced Zoom features during COVID-19 campus closure. *J. Chem. Educ.* 97, 2711–2714. doi: 10.1021/acs.jchemed.0c00675
- Smith, E. K., and Kaya, E. (2021). Online university teaching at the time of COVID-19 (2020): an Australian perspective. *IAFOR J. Educ.* 9, 183–200. doi: 10.22492/ije.9.2.11
- Spinelli, M., Lionetti, F., Pastore, M., and Fasolo, M. (2020). *Parents and Children Facing the COVID-19 Outbreak in Italy*. Available at SSRN 3582790. doi: 10.2139/ssrn.3582790
- Stelitano, L., Doan, S., Woo, A., Diliberti, M., Kaufman, J. H., and Henry, D. (2020). *The Digital Divide and COVID-19: teachers' Perceptions of Inequities in Students' Internet Access and Participation in Remote Learning. Data Note: insights from the American Educator Panels*. Research Report. RR-A134-3. Santa Monica, CA: Rand Corporation. doi: 10.7249/RR-A134-3
- Svihla, V., Reeve, R., Sagy, O., and Kali, Y. (2015). A fingerprint pattern of supports for teachers' designing of technology-enhanced learning. *Instruct. Sci.* 43, 283–307. doi: 10.1007/s11251-014-9342-5
- Svrcek, N. S., Rath, L., Olmstead, K., and Colantonio-Yurko, K. (2022). "We are still putting out fires": Considering educator intentionality in remote instruction during the COVID-19 pandemic. *Educ. Inform. Technol.* 27, 407–428. doi: 10.1007/s10639-021-10679-w
- Toquero, C. M. (2020). Challenges and opportunities for higher education amid the COVID-19 pandemic: the Philippine context. *Pedagog. Res.* 5, em0063. doi: 10.29333/pr/7947
- Tsai, C. C., and Chai, C. S. (2012). The "third"-order barrier for technology-integration instruction: implications for teacher education. *Australas. J. Educ. Technol.* 28, 1057–1060. doi: 10.14742/ajet.810
- UNESCO (2020). *COVID-19 Educational Disruption and Response*. Paris: UNESCO.
- Vallis, C., and Redmond, P. (2021). Introducing design thinking online to large business education courses for twenty-first century learning. *J. Univ. Teach. Learn. Pract.* 18, 213–234. doi: 10.53761/1.18.6.14
- Van Dijk, J., and Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *Inf. Soc.* 19, 315–326. doi: 10.1080/01972240309487
- Watson, A. D. (2015). Design thinking for life. *Art Educ.* 68, 12–18.
- Wiyono, B. B., Indreswari, H., and Putra, A. P. (2021). "The Utilization of 'Google Meet' and 'Zoom Meetings' to Support the Lecturing Process during the Pandemic of COVID-19," in *2021 International Conference on Computing, Electronics and Communications Engineering (iCCECE)* (IEEE), 25–29. doi: 10.1109/iCCECE52344.2021.9534847
- Zalat, M. M., Hamed, M. S., and Bolbol, S. A. (2021). The experiences, challenges, and acceptance of e-learning as a tool for teaching during the COVID-19 pandemic among university medical staff. *PLoS ONE* 16:e0248758. doi: 10.1371/journal.pone.0248758

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 Yeh and Tsai. 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.



Music Aptitude, Training, and Cognitive Transfer: A Mini-Review

Lu Wang*

Department of Educational Psychology, Ball State University, Muncie, IN, United States

In this mini-review, the genetic basis of music aptitude and the effects of music training are discussed. The review indicates that regardless of levels of innate ability, experience-induced neuroplasticity can occur as a result of music training. When that happens, it can be expressed as functional or structural brain changes. These changes are often accompanied by improvement in performance in tasks involving auditory analysis. Specifically, music training effects can transfer to a closely related cognitive domain such as auditory processing (near transfer). Music training can also affect more distantly related cognitive domains such as spatial and linguistic domains. Lastly, music training can affect general intelligence (“g”) (far transfer). Music training can mold behavioral brain development and confers cognitive benefits beyond music.

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Jaana Oikkonen,
University of Helsinki, Finland
Sabrina Turker,
Max Planck Institute for Human
Cognitive and Brain Sciences,
Germany

*Correspondence:

Lu Wang
hgse.lu.wang@gmail.com

Specialty section:

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

Received: 24 March 2022

Accepted: 13 June 2022

Published: 29 June 2022

Citation:

Wang L (2022) Music Aptitude,
Training, and Cognitive Transfer:
A Mini-Review.
Front. Psychol. 13:903920.
doi: 10.3389/fpsyg.2022.903920

Keywords: music aptitude, training, cognitive transfer, neuroplasticity, nature and nurture

INTRODUCTION

Depending on one's theoretical inclination, music aptitude has been defined in terms of collections of interrelated but distinct skills that involve auditory sensory discrimination around pitch, loudness, rhythm, time, timbre, and tonal memory (Seashore, 1938) or as the collective whole and the dynamic interplays of various auditory discrimination and memory skills. In a nutshell, the former camp endorses an atomistic view of music aptitude that compartmentalizes it into six subsets of skills, whereas the latter camp endorses an omnibus view of music aptitude that emphasizes on unity of the skills required to render a beautiful piece of music on a musical instrument (Gordon, 1969). Both camps agree that music aptitude is at least partially genetically determined (nature). There is a difference in view regarding the extent music aptitude can be improved through training (experience-induced neuroplasticity). This mini-review provides evidence that supports either position and evaluates evidence against the notion of whether effects of music training are transferable to closely related (near transfer), somewhat related (medium-distance transfer), and remotely related domains (far transfer).

The mini-review starts by examining the genetic basis of music aptitude. The effects of music training are then discussed in the context of neuroplasticity research using neuroimaging and behavioral studies. Next, findings from training studies that reveal a transfer of music training to closely related and more distantly related cognitive domains are evaluated. These transfer effects illustrate the extent of experience-induced neuroplasticity through music training (nurture).

Nature and Nurture

There is some evidence that the brain basis of music aptitude is genetically determined (see Richardson, 1990; Freeman, 1999; Gagné, 1999; Strand and Gault, 2014). This genetic evidence highlights nature's contribution to music aptitude. This evidence includes identified genes that are

responsible for auditory discrimination, memory, and the inclination to practice (see Liu et al., 2016; Wesseldijk et al., 2021a). This evidence is reviewed in greater depth in the following section. There is also evidence that highlights nurture's contribution to music aptitude. This evidence reveals the malleability of music aptitude through training (experience-induced neuroplasticity). Schlaug et al. (2005), for instance, showed trends of functional and structural changes in the brain as soon as 1 year after instrumental training where little preexisting differences in brain activation patterns or structural differences were observed prior to or at the beginning of the music training. In a follow-up fMRI study, Hyde et al. (2009) reported structural brain changes in 6-year-olds who participated in 15 months of music training. These structural brain changes were also accompanied by behavioral improvements in fine motor control and auditory skills. For example, areas of the brain that are responsible for melodic and rhythmic processing showed activation and structural changes after music training. Specifically, areas of the brain that are associated with music perception and cognition showed structural enlargement or changes in activation patterns after music training, rather than predisposing certain individuals to participate in music training. These findings delineate nurture's contribution to the development of music aptitude.

Genetic Basis of Music Aptitude

Researchers who emphasize on nature's contribution to music aptitude cited genetic studies to bolster their view. Within this camp, some are convinced that experience-induced neuroplasticity through music training cannot extend inborn music aptitude, whereas others are open to such possibilities. Mosing et al. (2014a) belong to the former camp. These researchers argue that music practice does not causally influence music aptitude. Rather, individual differences in genetics affect both music aptitude and the inclination to practice. In another paper, Mosing et al. (2014b) showed that the genetic basis of music ability included genes that are specifically responsible for auditory functions and genes that also affect general cognitive ability. More recently, Wesseldijk et al. (2021a) identified a partial association between music aptitude and verbal ability after correcting for general intelligence. Shared genetic factors explained 50% of this association and another 35% was explained by non-shared environmental influences. Wesseldijk et al. (2021b) argues that the results from their co-twin-control-analysis suggest music and verbal abilities may share an underlying genetic etiology, rather than music training at a young age causing cognitive transfer in a verbal domain at the age of 18. The research group published another paper (see Wesseldijk et al., 2021a) that discusses the reasons why an early start in music training contribute to musical skills in adulthood. The paper reiterated the genetic contribution of music aptitude, while at the same time placing a greater emphasis on nature and nurture interaction in developing musical skills into adulthood, with an emphasis placed on familial factors. Musically inclined parents not only pass on their musical genes to their offspring but are also more likely to provide family environment that is conducive to developing musical skills and invest in music lessons.

Other recent research on the molecular and evolutionary backgrounds of music aptitude included Liu et al.'s. (2016) work, a case control study that used HploPS, XP-EHH, and FST methods and compared genome-wide genotyping data (641K SNPs) in 148 Finnish individuals, among which there are 74 cases of high music aptitude individuals. HploPS is a method of locating genomic evidence by using the phased haplotypes of multiple samples from a population (Liu et al., 2013). The XP-EHH test is a method of detecting "selective sweeps in which the selected allele has risen to high frequency or fixation in one population, but remains polymorphic in the human population as a whole" (Sabeti et al., 2007). FST as a method reflects on the genetic differentiation among populations and shed light on the evolutionary properties of the genetic structure of the populations being studied (Whitlock and McCauley, 1999). Having assigned three positive selection regions, gene ontology classification revealed that the positive selection regions contained genes that affect inner-ear development that are critical to music perception. Liu et al.'s. (2016) literature review also identified adhesion G protein-coupled receptor (*ADGRV1*), also known as G protein-coupled receptor 98 (*GPR98*) and usherin (*USH2A*), as contributing to auditory perception; glutamate ionotropic receptor NMDA type subunit 2B (*GRIN2B*), interleukin 1 alpha (*IL1A*), interleukin 1 beta (*IL1B*), and Rap guanine nucleotide exchange factor 5 (*RAPGEF5*) as contributing to general cognition and memory; forkhead box P1 (*FOXP1*), regulator of G protein signaling 9 (*RGS9*), *GPR98*, and *GRIN2B* as contributing to song perception. Like Oikkonen et al. (2015), Liu et al. (2016) sought for genetic markers of music aptitude. Oikkonen et al. (2015) identified genes affecting inner ear development and brain functions corresponding to music perception. Alpha-synuclein gene (*SNCA*) was found to be overexpressed when listening to and performing music (Kanduri et al., 2015). The GATA-binding protein 2 (*GATA2*) regulates *SNCA* in dopaminergic neurons. *GATA2* links DNA- and RNA-studies of music aptitude [see Oikkonen and Järvelä (2014) review for related genetic studies of music aptitude; see also Tan et al. (2014) for a comprehensive review on behavioral and molecular genetic studies of music aptitude]. Genomic analyses conducted by Park et al. (2012) research group on 1008 Mongolian individuals identified an intergenic single nucleotide polymorphism (SNP) and a non-synonymous SNP in UDP Glycosyltransferase 8 (*UGT8*) to be strong determinants of music aptitude, as assessed by a pitch-production accuracy test.

Near Transfer, Medium-Distance Transfer, and Far Transfer

A trait's having a genetic basis does not preclude experience-induced neuroplasticity from taking place. In other words, if a child is genetically endowed with high music aptitude, it does not preclude the possibility that through quality instruction and deliberate practice, the child can extend that innate aptitude further. A body of research supports experience-induced neuroplasticity, both at the functional/structural brain level and at the behavioral level. This section discusses findings concerning "near transfer," a transfer of music training to closely related

cognitive domains, “medium-distance transfer,” a transfer of music training to somewhat related cognitive domains, and “far transfer,” a domain-general transfer of music training. In this discussion, music experience is limited to active music experience. This means “the Mozart effect,” which results from music listening, a passive music experience, is beyond the scope of this manuscript.

The Mozart effect refers to a temporal boost of spatial-temporal reasoning after being exposed to a Mozart sonata K448 for 10 min (see Rauscher et al., 1993). This effect spurred a craze about listening to Mozart, classical music, or compositions that show regularities similar to that of Mozart sonata K448. The effect was replicated by a few independent research labs, though even in studies where positive findings were reported, the effect size was small and short-lived (lasts about 12 min). Also, the effect size varies depending on the test that was used to measure spatial ability (Jenkins, 2001). Furthermore, for every successful replication of the Mozart effect, there are about just as many studies that failed to reproduce the Mozart effect. The Mozart effect is considered a passive listening effect. Unlike the Mozart effect, however, the music training effect concerns active participation in music, e.g., receiving music instruction, engaging in deliberate practice.

Near Transfer (Auditory Processing Skills)

Music training, especially when having an early start, was intensive, or lasted for some time, confers cognitive, emotional, and physiological benefits. This manuscript focuses on the effects of music training on cognition. Improvement in auditory processing skills exemplifies near transfer, improvement in verbal and visuospatial abilities illustrates medium-distance transfer, and enhanced executive functions and “g” instantiates the occurrence of far transfer.

Tervaniemi et al. (2011) identified the neural signatures (MMNm) of long-term music training (musical expertise) using magnetoencephalographic (MEG). These neural signatures included the left auditory areas. The researchers also found that musically trained and musically talented groups engage both auditory cortices in an integrative manner. Tervaniemi et al. (2011) study showed that music training and music aptitude have similar neural signatures during chord discrimination. However, their findings cannot answer the question of whether structural and activation differences in the brain is due to postnatal training or innate aptitude.

Direct evidence supporting near transfer is provided by Kuchenbuch et al. (2012) study. Also using magnetoencephalography (MEG), the research group investigated the influence of long-term music training on the processing of partly imagined tone patterns (imagery condition) and identical but perceived patterns (perceptual condition). The magnetic counterpart of mismatch negativity (MMNm) of musicians and non-musicians were recorded as they performed a lab devised task requiring them to detect deviant tone patterns. The results showed that the latency and the laterality of the MMNm in the perceptual condition differed significantly between the groups, with musicians having an earlier MMNm onset, especially in the left hemisphere. At a

behavioral level, the musician group outperformed the non-musician group in detecting deviant tones (perceptual condition) and imaging deviant tones (imagery condition). Kuchenbuch et al. (2012) concluded that processing patterns is faster and more strongly lateralized in musicians. These findings lent direct support for near transfer as a result of long-term music training.

Banai and Ahissar (2013) proposes that the extent music training transfer into other cognitive domains and the degree of generalization may be subjective to a threshold effect. Individuals with poor memory and auditory discrimination skills are likely to experience a broader generalization or cross-domain transfer as a result of music training than their high ability counterpart. Also, they believe that these individuals’ auditory and memory skills are also likely to show a more pronounced positive association with their music-related skills. Due to a correlational design, however, Banai and Ahissar (2013) is unable to definitively conclude a causal effect from music training to auditory processing and, subsequently, verbal skills such as reading.

Medium-Distance Transfer (Verbal and Visuospatial Domains)

In between studies illustrating direct transfer (near transfer) and domain-general transfer (far transfer) are cross-domain transfer (medium-distance transfer). In this section, studies supporting cross-domain transfers from music to verbal and visuospatial domains are discussed. Not only does music and language share an evolutionary history, music also has a special affinity with the phonetic aspect of language. Mackenzie Beck (2003) and Dankovičová et al. (2007), for instance, showed that music aptitude is predictive of phonetic skills. Phonetic skills are a modular aspect of language.

Articulatory rehearsal is another modular aspect of language. Degé and Schwarzer (2017) compared musically trained 10–12-year-old children on two verbal working memory tests, one involving memorizing a word list under a normal condition and the other involving memorizing a similar word list under an articulatory suppression condition. After controlling for gender, socioeconomic status (SES), general intelligence, motivation, music aptitude, and personality, musically trained children still outperformed untrained children under normal but not under articulatory suppression condition, suggesting that music training enhances linguistic ability through improving articulatory rehearsal. Articulatory rehearsal is involved during short-term storage of word lists.

Forgeard et al. (2008) study not only provided evidence for near transfer from music to motor skills and auditory processing, but also cross-domain (medium-distance) transfer. Beside auditory discrimination and motor skills tests, children who participated in the study also completed batteries of non-verbal reasoning (Raven’s Progressive Matrices), a verbal ability test (Vocabulary subtest from the WISC-III), two spatial ability tests (Block Design subtest of the WISC-III and Object Assembly subtest of the WISC-III), and a mathematics achievement test (Keymath-Revised: A Diagnostic Inventory of Essential Mathematics). The results of the analyses showed that musically trained group outperformed their untrained counterpart in verbal and non-verbal reasoning. It should be noted that unlike

Rauscher et al. (1997) study, discussed in greater detail in the ensuing paragraph, Forgeard et al. (2008) study did not find music training to improve spatial ability. Nor was there strong evidence for improvement in mathematics achievement in musically trained children.

Although from Forgeard et al. (2008) study, it appears as though the effects of music training has a limited transferability to other cognitive domains (in this study, it is verbal and non-verbal reasoning, but not spatial and mathematics), and the fact that the study was correlational by design and hence cannot be used as evidence for a causal relationship, Rauscher et al. (1997) study, discussed in greater detail in the ensuing paragraph, is experimental by design and showed a causal relationship between music training and enhanced spatial-temporal reasoning in young children. Similarly, a few other studies are suggestive of a positive association between music training and enhanced math-related skills as well (see Cheek and Smith, 1999; Ribeiro and Santos, 2017; Rodriguez et al., 2019). Cheek and Smith (1999), although correlational by design, showed a positive association between 2 years of private music lessons and enhanced numerical skills in typically developing adolescents. Ribeiro and Santos (2017) studies, although quasi-experimental by design, also showed positive associations between music training and subsets of numerical skills. Specifically, both typically developing (TD) and the group with developmental dyscalculia (DD), a condition marked by impoverished numerical skills, performed significantly better on number comprehension, number production, and calculation tests after a 7-lesson music training program that targeted core skills in music such as rhythms, melodies, and harmonies perception than their baseline performance on the same numerical tests prior to the music training.

The most compelling evidence for the cross-domain (Medium-Distance) transfer come from two studies that used a variant of the experimental design. Using a double-blind prospective case-control study, Rodriguez et al. (2019) studied the effects of eight group sessions of Numeracy Music training (NMT) on 42 school children (age ranged from 8 to 10)'s numerical cognition, working memory, and math anxiety. The results showed a reduction in math anxiety, as well as significant improvements in performance on numerical cognition and working memory tests after the training. Another experimental study that enables a causal inference of transfer of music training to another cognitive domain (spatial cognition) to be drawn comes from Rauscher et al. (1997) study. Rauscher et al. (1997) investigated the impact of music training on spatial recognition and spatial-temporal reasoning in preschool children who enrolled in keyboard lessons. Rauscher et al. (1997) were among the first to provide experimental evidence for a cross-domain transfer from music training to certain spatial skills. The study showed that music training, even of a short duration, improves spatial-temporal reasoning (aka mental rotation, spatial imagery) but not spatial recognition (aka object recognition, visual ability). The magnitude of improvement in spatial-temporal reasoning was greater than one standard deviation. The effect lasted for more than 1 day.

The discrepancy between Rauscher et al. (1997) and Forgeard et al. (2008) findings concerning the effect of music training

on spatial temporal reasoning may be due to musically trained individuals have improved their spatial ability initially. However, at some point, the developmental advantage of music training may taper off. The non-musically trained group catches up. The null finding reported by Forgeard et al. (2008) may also be due to the Block Design subtest not being sensitive enough to pick up the training effects. This is because in other studies that used different instruments to measure spatial-temporal reasoning (see Rauscher et al., 1997; Sarnthein et al., 1997), musically trained individuals showed better spatial-temporal reasoning.

In a causal-comparative (a quasi-experimental design) study, Wang et al. (2018) found a positive association between music training and perspective-taking and spatial imagery abilities in young adults attending 4-year colleges. The Perspective-Taking Ability (PTA) test measures the ability to navigate in real or virtual large-scale space. The Spatial Imagery Ability (SIA) test measures the ability to mentally represent, temporarily store, and manipulate visuospatial information along an allocentric frame of reference.

Far Transfer (Executive Functions and General Intelligence/"g" Factor)

There is also some evidence for far transfer from music to domain-general cognition (Brandler and Rammsayer, 2003; Helmbold et al., 2005). Music training is positively associated with executive functions and "g" (Degé et al., 2011; Okada and Slevc, 2016). Childhood music training predicts better mathematics, language, and spatial skills (see Schellenberg, 2005; Rauscher and Hinton, 2011; Schellenberg and Winner, 2011). However, there are also studies that failed to find positive music training effects on domain-general cognition (see Schellenberg, 2011).

In a selective review of neuroimaging studies that investigated the neural signatures of cognitive training, Putkinen and Saarikivi (2018) found decreases in frontoparietal activities after music training. The researchers concluded that musically trained individuals have more mature executive functions that are characterized by more efficient information processing in frontoparietal areas. This greater efficiency in information processing corresponds to less activation in frontoparietal network, which undergirds executive functions. The researchers then presented their preliminary findings of an fMRI study that explored executive functions development. Consistent with their prediction, musically trained adolescents and young adults showed fewer frontoparietal brain activations than their untrained counterparts when performing challenging executive functions tasks. This neural signature was accompanied by a behavioral advantage in the musically trained group. In the same study, Putkinen and Saarikivi (2018) also measured general intelligence, using a combination of verbal and visuospatial tests, but did not find group differences.

In light of the mixed findings concerning music training and general intelligence, Schellenberg and Moreno (2010) conducted a study using a full-scale Advanced Raven's Progressive Matrices to measure "g." Raven's Progressive Matrices and Advanced Raven's Progressive Matrices are considered a gold standard to capture "g." Their results failed to reveal significant

group differences in “g” between musically trained and untrained individuals.

Hargreaves and Aksentijevic (2011) and Okada and Slevc (2016) discussed factors that could contribute to mixed findings on far transfer. Some of these factors are: (i) the correlations between music training and IQ may vary at different levels of music aptitude; (ii) socio-economic status (SES) differences between musically trained and untrained group, though sometimes statistically controlled for, may still influence the results of the studies; (iii) differences in parental educational background between musically trained and untrained group are often unaccounted for; (iv) the effects of music training on general intelligence may differ across age groups; (v) the effects of music training on general intelligence may vary among individuals with different neurological conditions.

Perhaps most importantly, controversies concerning the effect of music training on general intelligence stems from the age-old “chicken-and-egg” problem. Group differences in general intelligence could be due to preexisting differences in intellectual profiles. Individuals who are drawn to music may have higher IQs to start with (see Swaminathan et al., 2017), rather than music training improving general intelligence (Schellenberg and Moreno, 2010; Schellenberg, 2011). To effectively argue that the direction of effect is from music training to general intelligence, experimental, longitudinal, and neuroimaging studies are needed.

Experience-Induced Neuroplasticity Through Music Training and the Role of a Sensitive Period

In this section, how neuroplasticity occurs as a result of music training and its possible interaction with a sensitive period is explained. Prior research showed that bilingualism, participation in some sports activities, and even experience working as a taxi driver in London for a certain amount of time can induce functional or structural changes in the brain (see Maguire et al., 2000; Habibi et al., 2018). Recent research in the area of language development (see Turker et al., 2021) suggests that a sensitive period exists in language learning. Inter-individual differences in language learning outcomes over the course of development may be explained in terms of the interaction between neuroplasticity and sensitive period. Certain aspects of language (e.g., the phonetic aspect of second language acquisition) may be constrained by a sensitive period. In other words, language acquisition, whether L1 or L2, is more efficient during the sensitive period, although individual differences in linguistic aptitude and motivational factors also play important roles. In Turker et al.’s (2021) words, the interaction between experiential and maturational factors could impact language learning across the lifespan. Late L2 acquisition is constrained on the neural level by “maturational declines in synaptic density, decreased levels of brain metabolism (Bates et al., 1992), and increased axon myelination (Pulvermüller and Schumann, 1994).”

Although music is not singular in inducing neuroplasticity, it is unique in its complexity and the extent it engages the whole brain (Herholz and Zatorre, 2012). For that reason, some suggest that professional musicians’ brains are a good model

to study lifespan neuroplasticity (Münte et al., 2002; Rodrigues et al., 2010; Herholz and Zatorre, 2012). Studies published thus far explored neuroplasticity as a result of music training using a variety of structural and functional neuroimaging techniques including MRI, EEG, fMRI, PET, and DTI (Münte et al., 2002). Some of these studies showed changes in brain activation patterns after just a short period of training (see Rauscher et al., 1997; Rauscher and Zupan, 2000; Norton et al., 2005; Hyde et al., 2009). There is also evidence that intensive long-term music training can induce morphological changes in gray matters from this literature (see Herholz and Zatorre, 2012). What remains to be explored further, however, is the precise time window, the domain-specificity of sensitive period, and the neural instantiations of the interaction between neuroplasticity and sensitive period concerning music training. Given that some aspects of language learning (e.g., phonetics) resembles music learning (e.g., pitch identification or discrimination) and engages the same brain regions, there is a good reason to speculate that the sensitive period for music training at least partially overlaps with that of language acquisition.

In sum, key findings and directions for further research on the topic of experience-induced neuroplasticity and the role of sensitive period in music training are: (i) structural brain changes in adult professional instrumentalists can be predicted by intensity of training, starting age of training, and duration of training (Merrett and Wilson, 2012); (ii) there are large inter-individual variabilities in response to music training; (iii) the effects of music training may vary by types of music training (e.g., instruments, instructional formats, see Reybrouck and Brattico, 2015); (iv) the direction of causal effects between music training and general intelligence remains to be experimentally determined (Schellenberg and Peretz, 2007); (v) due to similarities between some aspects of language acquisition and music learning, music training’s effectiveness will likely be subject to the influence of sensitive period—a time window that demarcates a high level of receptivity and efficiency in learning. The effects of music training will likely to be more pronounced when music training takes place during the sensitive period (White et al., 2013).

CONCLUSION

This mini-review provides a bird eye view of the interdisciplinary field of mind, brain, and education on music aptitude, training, and cognitive transfer. Insights gained from synthesizing these bodies of works can inform educational neuroscience research concerning how gene-environmental interactions contribute to the unfolding of music aptitude across the lifespan, with sensitive period moderating the malleability of music aptitude through training and domain similarity predicting the magnitude of cognitive transfer. Overall, this mini review indicates that regardless of levels of innate aptitude, experience-induced neuroplasticity can occur as a result of music training and may be subject to the influence of sensitive period. Experience-induced neuroplasticity can be expressed as functional or structural brain changes (e.g., Schneider et al., 2002; Schlaug et al., 2005) that are often accompanied by enhanced behavioral performance (e.g.,

Hyde et al., 2009). The right auditory cortex, for example, has been shown to be malleable to music training (see Schneider et al., 2002) and is a good candidate for further investigations and replications.

Music training can directly transfer to closely related cognitive domains such as auditory processing (near transfer). Training effects can also leak into more distantly related cognitive domains such as certain aspects of spatial and linguistic skills *via* medium-distance transfer. More experimental and longitudinal research is needed to establish the neural basis, behavioral manifestations, and causal mechanisms of medium-distance transfer from music to related cognitive domains that include language, mathematics, and spatial reasoning. Finally, music training can also affect general intelligence *via* far transfer.

Neuroplasticity research generally supports the notion that music training, especially when started early, was intensive, and lasted for several years, leave distinct neural signatures in the brain, accompanied by behavioral performance differences between musically trained and untrained groups (Rodrigues et al., 2010). Playing an instrument engages the whole brain that includes auditory processing, motor controls, executive functions, and bi-hemispheric communications. It therefore provides an ideal model to study how nature (music aptitude) and nurture (music training) work in tandem (White et al., 2013) in the making of a musician.

The nature of sensitive period as expressed in music learning and how neuroplasticity moderates the training effects on the

expression of music aptitude (as outlined in the previous section) warrant further research using experimental, longitudinal, and neuroimaging methods. This research will also refine our understanding of issues with broader implications such as mind-brain relationship, the architecture of the mind, and the causal mechanism of the transfer of training effects at the neuronal and behavioral levels. This mini-review suggests that music training molds behavioral brain development, regardless of innate levels of music aptitude, and confers cognitive benefits beyond music. The logical next-step in this research is to pinpoint the time window during which cognitive transfer effects are most robust.

Lastly, it should be noted that not all studies cited in support of some of the arguments made are independent from one another, meaning some studies cited are follow-up from a previous study and was led by a different researcher within the same research group (e.g., studies conducted by Järvelä and colleagues and Ullen and colleagues). Meta-analysts are advised to use cautions when polling studies directly from this mini-review's reference list without discerning the possible presence of dependency in findings reported by different authors within the same research group.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

- Banai, K., and Ahissar, M. (2013). Musical Experience, Auditory Perception and Reading-Related Skills in Children. *PLoS One* 8:e75876. doi: 10.1371/journal.pone.0075876
- Bates, E., Thal, D., Finlay, B., Clancy, B., Origins, P. D., and Int, N. S. F. (1992). "Early language development and its neural correlates," in *Handbook of Neuropsychology 2nd ed*, eds I. Rapin and S. Segalowitz. (Amsterdam: Elsevier), 69–110.
- Brandler, S., and Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychol. Music* 31, 123–138.
- Cheek, J. M., and Smith, L. R. (1999). Music training and mathematics achievement. *Adolescence* 34, 759–759.
- Dankovičová, J., House, J., Crooks, A., and Jones, K. (2007). The Relationship between Musical Skills. *Lang. Speech* 50, 177–225.
- Degé, F., Kubicek, C., and Schwarzer, G. (2011). Music lessons and intelligence: a relation mediated by executive functions. *Music Percept.* 29, 195–201.
- Degé, F., and Schwarzer, G. (2017). Music Lessons and Verbal Memory in 10- to 12-Year-Old Children: Investigating Articulatory Rehearsal as Mechanism Underlying This Association. *Psychomusicology* 27, 256–266. doi: 10.1037/pmu0000201
- Forgeard, M., Winner, E., Norton, A., and Schlaug, G. (2008). Practicing a Musical Instrument in Childhood is Associated with Enhanced Verbal Ability and Non-verbal Reasoning. *PLoS One* 3:e3566. doi: 10.1371/journal.pone.0003566
- Freeman, C. (1999). The crystallizing experience: a study in musical precocity. *Gift. Child Q.* 43, 75–85. doi: 10.1177/001698629904300204
- Gagné, F. (1999). Nature or nurture? A re-examination of Sloboda and Howe's (1991) interview study on talent development in music. *Psychol. Music* 27, 38–51.
- Gordon, E. (1969). Intercorrelations among musical aptitude profile and seashore measures of musical talents subtests. *J. Res. Music Educ.* 17, 263–271.
- Habibi, A., Damasio, A., Ilari, B., Elliott Sachs, M., and Damasio, H. (2018). Music training and child development: a review of recent findings from a longitudinal study. *Ann. N.Y.Acad. Sci.* 1423, 73–81. doi: 10.1111/nyas.13606
- Hargreaves, D. J., and Akstentijevic, A. (2011). Music, IQ, and the executive function. *Br. J. Psychol.* 102, 306–308.
- Helmbold, N., Rammsayer, T., and Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *J. Individ. Differ.* 26, 74–85. doi: 10.1027/1614-0001.26.2.74
- Herholz, S. C., and Zatorre, R. J. (2012). Music training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron* 76, 486–502. doi: 10.1016/j.neuron.2012.10.011
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., et al. (2009). Music training shapes structural brain development. *J. Neurosci.* 29, 3019–3025. doi: 10.1523/JNEUROSCI.5118-08.2009
- Jenkins, J. S. (2001). The Mozart effect. *J. R. Soc. Med.* 94, 170–172. doi: 10.1177/014107680109400404
- Kanduri, C., Kuusi, T., Ahvenainen, M., Philips, A. K., Lähdesmäki, H., and Järvelä, I. (2015). The effect of music performance on the transcriptome of professional musicians. *Sci. Rep.* 5:9506.
- Kuchenbuch, A., Paraskevopoulos, E., Herholz, S. C., and Pantev, C. (2012). Electromagnetic Correlates of Musical Expertise in Processing of Tone Patterns. *PLoS One* 7:E30171. doi: 10.1371/journal.pone.0030171
- Liu, X., Kanduri, C., Oikkonen, J., Karma, K., Raijas, P., Ukkola-Vuoti, L., et al. (2016). Detecting signatures of positive selection associated with music aptitude in the human genome. *Sci. Rep.* 6:21198. doi: 10.1038/srep21198
- Liu, X., Ong, R. T., Pillai, E. N., Elzein, A. M., Small, K. S., and Clark, T. G. (2013). Detecting and characterizing genomic signatures of positive selection in global populations. *Am. J. Hum. Genet.* 92, 866–881. doi: 10.1016/j.ajhg.2013.04.021
- Mackenzie Beck, J. (2003). "Is it possible to predict students' ability to develop skills in practical phonetics?" in *In Proceedings of the 15th International Congress of Phonetic Sciences*, (Bellaterra: Universitat Autnoma de Barcelona).
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4398–4403. doi: 10.1073/pnas.070039597
- Merrett, D. L., and Wilson, S. J. (2012). "Music and Neural plasticity," in *Lifelong Engagement with Music*, eds N. S. Rickard and

- K. McFerran (Hauppauge, NY: Nova Science Publishers, Inc), 123–162.
- Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., and Ullén, F. (2014a). Practice does not make perfect: no causal effect of music practice on music ability. *Psychol. Sci.* 25, 1795–1803. doi: 10.1177/0956797614541990
- Mosing, M. A., Pedersen, N. L., Madison, G., and Ullén, F. (2014b). Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. *PLoS One* 9:e113874. doi: 10.1371/journal.pone.0113874
- Münste, T. F., Altenmüller, E., and Jäncke, L. (2002). Opinion: the musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478. doi: 10.1038/nrn843
- Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., and Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain Cogn.* 59, 124–134. doi: 10.1016/j.bandc.2005.05.009
- Oikkonen, J., Huang, Y., Onkamo, P., Ukkola-Vuoti, L., Raijas, P., Karma, K., et al. (2015). A genome-wide linkage and association study of music aptitude identifies loci containing genes related to inner ear development and neurocognitive functions. *Mol. Psychiatry* 20, 275–282. doi: 10.1038/mp.2014.8
- Oikkonen, J., and Järvelä, I. (2014). Genomics approaches to study music aptitude. *Bioessays* 36, 1102–1108.
- Okada, B. M., and Slevc, L. R. (2016). “Music training: Contributions to executive function,” in *An Integrative Approach to Cognitive and Working Memory Training: Perspectives from Psychology, Neuroscience, and Human Development*, eds M. Bunting, J. Novick, M. Dougherty, and R. W. Engle (New York, NY: Oxford University Press).
- Park, H., Lee, S., Kim, H. J., Ju, Y. S., Shin, J. Y., Hong, D., et al. (2012). Comprehensive genomic analyses associate UGT8 variants with musical ability in a Mongolian population. *J. Med. Genet.* 49, 747–752. doi: 10.1136/jmedgenet-2012-101209
- Pulvermüller, F., and Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Lang. Learn.* 44, 681–734. doi: 10.1111/j.1467-1770.1994.tb00635.x
- Putkinen, V., and Saarikivi, K. (2018). Neural correlates of enhanced executive functions: is less more? *Ann. N.Y.Acad. Sci.* 1423, 117–125. doi: 10.1111/nyas.13645
- Rauscher, F. H., Shaw, G. L., and Ky, C. N. (1993). Music and spatial task performance. *Nature* 365:611.
- Rauscher, F., Shaw, G., Levine, L., Wright, E., Dennis, W., and Newcomb, R. (1997). Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurol. Res.* 19, 2–8. doi: 10.1080/01616412.1997.11740765
- Rauscher, F. H., and Hinton, S. C. (2011). Music instruction and its diverse extra-musical benefits. *Music Percept.* 29, 215–226.
- Rauscher, F. H., and Zupan, M. A. (2000). Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: a field experiment. *Early Child. Res. Q.* 15, 215–228.
- Reybrouck, M., and Brattico, E. (2015). Neuroplasticity beyond Sounds: Neural Adaptations Following Long-Term Musical Aesthetic Experiences. *Brain Sci.* 5, 69–91. doi: 10.3390/brainsci5010069
- Ribeiro, F. S., and Santos, F. H. (2017). Enhancement of numeric cognition in children with low achievement in mathematics after a non-instrumental music training. *Res. Dev. Disabil.* 62, 26–39. doi: 10.1016/j.ridd.2016.11.008
- Richardson, C. P. (1990). Measuring musical giftedness. *Music Educ. J.* 76:40. doi: 10.2307/3401036
- Rodrigues, A. C., Loureiro, M. A., and Caramelli, P. (2010). Music training, neuroplasticity and cognition. *Dement. Neuropsychol.* 4, 277–286. doi: 10.1590/S1980-57642010DN40400005
- Rodriguez, I. A., Do Nascimento, J. M., Voigt, M. F., and Dos Santos, F. H. (2019). Numeracy music training for school children with low achievement in mathematics. *An. de Psicol.* 35, 405–416.
- Sabeti, P., Varilly, P., and Fry, B. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918. doi: 10.1038/nature06250
- Sarnthein, J., VonStein, A., Rappelsberger, P., Petsche, H., Rauscher, F., and Shaw, G. (1997). Persistent patterns of brain activity: an EEG coherence study of the positive effect of music on spatial-temporal reasoning. *Neurol. Res.* 19, 107–116.
- Schellenberg, E. G. (2005). Music and cognitive abilities. *Curr. Dir. Psychol. Sci.* 14, 317–320.
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *Br. J. Psychol.* 102, 283–302.
- Schellenberg, E. G., and Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychol. Music* 38, 209–221. doi: 10.1177/0305735609339473
- Schellenberg, E. G., and Peretz, I. (2007). Music, language and cognition: unresolved issues. *Trends Cogn. Sci.* 12, 45–46. doi: 10.1016/j.tics.2007.11.005
- Schellenberg, E. G., and Winner, E. (2011). Music training and non-musical abilities: introduction. *Music Percept.* 29, 129–132. doi: 10.1525/mp.2011.29.2.129
- Schlaug, G., Norton, A., Overy, K., and Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Ann. N.Y.Acad. Sci.* 1060, 219–230.
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., and Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694. doi: 10.1038/nrn871
- Seashore, C. E. (1938). Musical intelligence. *Music Educ. J.* 24, 32–33.
- Strand, K. D., and Gault, B. M. (2014). “Musical giftedness,” in *Critical Issues and Practices in Gifted Education: What the Research says*, 2nd ed, eds J. A. Plucker and C. M. Callahan (Austin: Prufrock Press), 439–450.
- Swaminathan, S., Schellenberg, E. G., and Khalil, S. (2017). Revisiting the association between music lessons and intelligence: training effects or music aptitude? *Intelligence* 62, 119–124.
- Tan, Y. T., McPherson, G. E., Peretz, I., Berkovic, S. F., and Wilson, S. J. (2014). The genetic basis of music ability. *Front. Psychol.* 5:658. doi: 10.3389/fpsyg.2014.00658
- Tervaniemi, M., Sannemann, C., Noyranen, M., Salonen, J., and Pihko, E. (2011). Importance of the left auditory areas in chord discrimination in music experts as demonstrated by MEG. *Eur. J. Neurosci.* 34, 517–523. doi: 10.1111/j.1460-9568.2011.07765.x
- Turker, S., Seither-Preisler, A., and Reiterer, S. M. (2021). Examining individual differences in language learning: a neurocognitive model of language aptitude. *Neurobiol. Lang.* 2, 389–415.
- Wang, L., Ester, D., and Finch, H. (2018). A Causal Comparative Study of Visuospatial Thinking among Mathematics, Music, and Psychology Students. *Psychol. Educ.* 55, 1–13.
- Wesseldijk, L. W., Gordon, R. L., Mosing, M. A., and Ullén, F. (2021a). Music and verbal ability—a twin study of genetic and environmental associations. *Psychol. Aesthet. Creat. Arts.* doi: 10.1037/aca0000401 [Epub ahead of print].
- Wesseldijk, L. W., Mosing, M. A., and Ullén, F. (2021b). Why is an early start of training related to musical skills in adulthood? A genetically informative study. *Psychol. Sci.* 32, 3–13. doi: 10.1177/0956797620959014
- White, E. J., Hutka, S. A., Williams, L. J., and Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Front. Syst. Neurosci.* 7:90. doi: 10.3389/fnsys.2013.00090
- Whitlock, M. C., and McCauley, D. E. (1999). Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity* 82, 117–125. doi: 10.1038/sj.hdy.6884960

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.

Copyright © 2022 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.



Views of Hospital Nurses and Nursing Students on Nursing Engagement—Bridging the Gap Through Communication Courses

Qing Huang^{1†} and Jack Pun^{2*†}

¹ School of Foreign Languages, Huizhou University, Huizhou, China, ² Department of English, City University of Hong Kong, Kowloon, Hong Kong SAR, China

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Walter Millanzi,
University of Dodoma, Tanzania
Solange Mianda,
University of the Western Cape,
South Africa

*Correspondence:

Jack Pun
jack.pun@cityu.edu.hk

†ORCID:

Qing Huang
orcid.org/0000-0002-2008-9888
Jack Pun
orcid.org/0000-0002-8043-7645

Specialty section:

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

Received: 07 April 2022

Accepted: 30 May 2022

Published: 29 June 2022

Citation:

Huang Q and Pun J (2022) Views of
Hospital Nurses and Nursing Students
on Nursing Engagement—Bridging
the Gap Through Communication
Courses. *Front. Psychol.* 13:915147.
doi: 10.3389/fpsyg.2022.915147

Background: Communicative engagement plays a significant role for effective nurse–patient communication. In the existing college nursing communication training within and outside China, there is a difference between what students are taught and what they can apply in their clinical placements.

Aims: Using needs analysis, this mixed-methods study explored potential gaps between frontline hospital nurses' and college nursing students' perceptions of nurse–patient communicative engagement, and collated a list of effective engagement strategies for pedagogy.

Methods: Surveys and interviews were conducted with key stakeholders, including 16 hospital nurses and 60 nursing students. A new scale named Nursing Engagement with Patients Scale (NEPS) was developed and validated to explore stakeholders' views on nursing engagement.

Results: Differences between the views of nurses and students on engagement were identified. While frontline nurses affirmed the importance of engaging with patients while providing nursing care, nursing students were unsure about the concept and role of engagement, and how to enact it. A list of communication strategies that promote engagement was culled from the interviews with the experienced nurses.

Implications: These can be used to inform nursing communication courses to bridge the gap between what nursing students are currently taught and what they will need in the workplace.

Keywords: needs analysis, communication strategies, engagement, nurses, nursing students

INTRODUCTION

Effective nurse–patient communication can be achieved in a number of ways, one of which is the nursing professional's engagement with patients during their provision of medical care. The engagement has been studied by a number of researchers in the field (Roe et al., 2012; Bright et al., 2018), but the following description distills the various strands that are part of its practice into a concise definition: Nursing engagement is a nurse's active or proactive efforts to interact with a patient through two-way, patient-centered, communication in order to ensure that the nurse

is understood by the patient, and also the nurse can assist the patient in revealing any concerns they may have, and provide empathy and rapport with the patient. Engagement helps achieve “holistic nursing,” characterized by Candlin and Candlin (2014, p. 260) as follows: “*in caring for persons across the lifespan and from all cultural backgrounds—whether sick or healthy—appropriate and ethical care is based on individual need: physical, emotional/psychological, social and spiritual, where all such needs are ineluctably interconnected in sites of nursing engagement.*”

Similar to nurses in other countries, nurses in China pay attention to the needs of their patients and play a major role in health care by engaging with patients on various levels. Yam and Rossiter (2000) found that Chinese nurses valued interpersonal communication that included engagement behaviors, such as encouraging patients to share their personal stories, or eliciting patients’ emotional needs or unexpressed concerns that could help with their overall treatments. Similarly, Dong et al. (2016) found that Chinese nurses emphasized on the importance of meeting each patient’s spiritual needs, one of the aspects of engagement, to help patients cope with stress, and help nurses build empathy with them.

Although nurses recognize the importance of engaging with patients, many often focus more on providing physical comfort than emotional support (Pun et al., 2018). Dong et al. (2016) interviewed Chinese nurses involved in looking after patients with terminal cancer and found that they were focused mainly on their patients’ physical comfort during their nursing communication. Similarly, Jiang et al. (2015) used a nurse–patient interaction caring scale to measure the caring attitudes and behaviors of 260 nurses in mainland China. They found that nurses were more competent at providing operative care, such as drug delivery and injections, than they were at delivering expressive care, such as showing empathy and offering positive rapport, to engage with patients. More specifically, nurses holding higher positions (i.e., senior or supervisor nurses) had higher competency scores for caring and humanistic behaviors (i.e., features of engagement), in comparison to nurses in lower positions (Jiang et al., 2015).

The reasons behind the neglect by nurses of issues, such as emotional, spiritual, or environmental concerns, could lie in the hectic health care environment, where the patients feel vulnerable and the staff is frequently stressed. Meanwhile, meeting patients’ holistic needs, which means all of their needs, including physical, emotional, spiritual, and environmental needs, demands more time and advanced interpersonal engagement skills (i.e., communication skills that promote engagement; Bramhall, 2014). Thus, nurses should have high level engagement skills, and the best way to ensure this is to teach engagement strategies (i.e., communication strategies that promote engagement) to novice nurses and nursing students.

Teaching Engagement in Nursing Communication

The spoken modality is the most important form to enact engagement because it is the speech that is the primary means by which nurses interact with patients in urgent and high-risk

clinical settings (Pangaltsos, 2011). The scope of the current study thus focuses on nurses’ spoken efforts of engagement. To engage with patients in a better way, nurses need to employ a variety of strategies that will help them connect with their patients, and these engagement strategies are best learned through nursing college communication courses, as nurses may not have the opportunity to learn them while on the job from their busy clinical supervisors.

Studies have shown that many existing communication courses outside and within China only teach nursing students a very limited set of engagement-related communication skills, such as techniques for establishing empathy with patients (Shao et al., 2018), understanding patients (Farrell et al., 2007), gathering information (Bosher and Smalkoski, 2002), or counseling and delivering information (Lin et al., 2013). No course was found to include all the aspects of full nursing engagement. For example, in a study by Bosher and Smalkoski (2002), training in information-gathering techniques helped nursing students elicit patients’ concerns but did not equip them with the skills to address patients’ emotional needs. Thus, a dedicated communication course with a concerted focus on spoken engagement skills is needed. Students would learn about specific strategies for each engagement function, and these strategies can be generalized to work across different language and cultural contexts.

In the conventional communication skills training for nursing students worldwide, many trainees have expressed frustration about discrepancies between the communication training in classroom settings and their actual communicative needs in clinical practice. While teachers emphasized general principles of communication (“Be nice, kind, and patient”), nursing students were more concerned with practical communication skills that can help them to perform effectively in the workplace (Corlett, 2000). This mismatch suggests a need to develop better-designed communication courses that can teach practical engagement skills and can be actualized in real-life clinical scenarios. These courses should, thus, ideally be informed by the views and experiences of frontline nurses regarding engagement in clinical settings to give trainee nurses a better picture of what they will encounter in their future workplaces and how they should actively engage with patients. The views of frontline nurses can be identified through a needs analysis, to determine the specific learning goals that need to be addressed in the course design (Woodrow, 2018). For example, needs analysis has been used to elicit the views of stakeholders, such as hospital managers, nurse supervisors, and senior nurses, to develop language advancement courses that train nurses in effective clinical communication skills (Bosher and Smalkoski, 2002). These studies showed that needs analysis is an approach that can be used to determine the target learners’ specific workplace language needs. The goal of the needs analysis in the current study was to answer the following research questions:

- (1) What are the current perceptions of nurses and students in China regarding engagement in nursing communication?
- (2) What specific kinds of communication strategies do the two groups think can lead to better engagement?

TABLE 1 | Demographic characteristics of the participating hospital nurses ($n = 16$).

Variables	Category	<i>n</i>
Age	21–30	10
	31–40	4
	41–50	2
Gender	Female	12
	Male	4
Position	Nurse supervisor	1
	Senior nurse	5
	Junior nurse	10
Years of experience as a registered nurse	5–10	10
	11–15	3
	16–20	1
	21–25	2

TABLE 2 | Demographic characteristics of the participating students ($n = 60$).

Variables	Category	<i>N</i>	%
Age	18	3	5
	19	15	25
	20	41	68
	21	1	2
Gender	Male	6	10
	Female	54	90
Major	Nursing	60	100
L1	Mandarin	60	100

MATERIALS AND METHODS

Using needs analysis, the present study leads to the first step toward a larger project on nursing communication course design focused on engagement. To explore the gaps between students' needs in target communicative situations and their current communicative ability (Woodrow, 2018), this study used the mixed methods approach. The approach “adopts both qualitative and quantitative methods that work in a complementary way” (Paltridge and Phakiti, 2015, p. 28), and thus contributes to comprehensively understanding stakeholders' views on engagement. This approach was done by interviews for the qualitative arm, and surveys for the quantitative arm, in the current study, with 16 hospital nurses and 60 college nursing students, over a period of 3 months, from August 2019 to February 2020 (7 months in total). Their perceptions of engagement were then contrasted, and the gaps between these two groups of participants can be used to design engagement-focused courses in nursing communication [as suggested in Hafner and Miller (2019) and Woodrow (2018)].

Participants

Sixteen nurses (including one nurse supervisor) from seven provincial grade-3 first-class hospitals, which are categorized as the top-ranking provincial hospitals in Guangdong province, and 60 students from a government-funded nursing college in Guangdong, participated in the study voluntarily *via* convenience sampling. The student volunteers were randomly selected from the 2019 cohort ($n = 650$) using a computer software. The demographic characteristics of the participating nurses and students are presented in **Tables 1, 2**, respectively.

Data Collection

Interviews

Each interview was ~25 min and was conducted in the participants' first language (i.e., Mandarin) by the first author, an instructor in health communication from the involved college. All interviews were audio-recorded for further analysis. To avoid

any potential problems from the power imbalance between the interviewer and the students, the participating students were anonymized by being allocated random identification numbers by helpers. Due to the COVID-19 pandemic, the student interviews were conducted online through WeChat group chats. As there was a large number of student participants, the students were interviewed using a focus group format (four to five per group) to discuss how they viewed engagement with patients. The interviews with the hospital nurses, on the other hand, were conducted one-on-one in a quiet office to elicit their opinions fully. To protect the nurses' privacy, all of them were given pseudonyms in the interview transcripts.

The nurses were asked about their views and reflections on engagement practices and the related communication strategies, based on their extensive clinical experience, and also about novice nurses' perceptions and practices of engagement in clinical settings. Since the participating nursing students had had limited clinical experience with patients (e.g., intern-training), they were interviewed with a focus on their understanding of nursing engagement. All the interview questions were adapted from Roe et al., 2012 research on clinicians' perceptions of engagement with patients and are given in **Appendix 1**. The interviews were conducted in Mandarin Chinese to ensure that the interviewees, who were all Mandarin speakers, could express their ideas clearly and fluently in their mother tongue. The interview questions were translated from the original English into Chinese by two independent translators. The two versions were compared, analyzed, and modified until a consensus was reached. To further ensure content validity, five hospital nurses and five nursing students (all not part of the research) were recruited in a pilot study to help improve the clarity of the wording of the questions.

Questionnaires

Given that student participants outnumbered nurse participants, the former completed an anonymous survey, Nursing Engagement with Patients Scale (NEPS), before the interview, for gathering as many students' perceptions of engagement as possible. NEPS included three themes: students' attitudes toward learning engagement skills, perceived importance of engagement, and perceived ability to execute engagement. Three corresponding established scales that had been tested for validity and reliability in previous studies were adapted for this study (Coates, 1997; Langille et al., 2001; Rees et al., 2002). They had

TABLE 3 | Summary of questionnaire items in NEPS.

Items	Type	Description/theme	Example
1, 4, 5, 10, 12, and 16	Opinion survey	Views on the learning of engagement skills	<i>In order to be a good nurse I must have good engagement skills (Item 1).</i>
3, 7, 8, 9, and 14	Opinion survey	Perceived importance of engagement	<i>Checking for patient understanding is generally unnecessary (Item 3).</i>
2, 6, 11, 13, and 15	Self-rating	Perceived ability to execute engagement	<i>I find it difficult to express empathy with patients (Item 2).</i>

been successfully and simultaneously used in previous studies to gain insights into nursing students' communication skills (e.g., Abdrbo, 2017). Since these three established scales were used to explore the perceptions of interpersonal communication, the adapted items for this research narrowed their scope down to one aspect of interpersonal communication: engagement. Thus, the definition and explanation of engagement, as mentioned earlier, were highlighted in the adapted items. **Appendix 2** shows the full version of the questionnaire used in this study. This new scale is named NEPS.

NEPS consists of 16 items, eight of which are written in the form of positive statements about perceptions of engagement while the other eight were negative statements. The inclusion of negatively worded items is an approach commonly employed to lessen the acquiescent response bias in a questionnaire (Qasem and Gul, 2014). These 16 items concern three themes and are deliberately mixed up in the scale so that similar items are not consecutive. Each theme covers roughly one-third of the items. Items 1, 4, 5, 10, 12, and 16 concern views on the learning of engagement skills, Items 3, 7, 8, 9, and 14 deal with the perceived importance of engagement, and Items 2, 6, 11, 13, and 15 involve perceived ability to execute engagement. **Table 3** shows a summary of questionnaire items in NEPS.

Responses to the statements in the questionnaire were on a Likert 6-point scoring scale (*Strongly agree, Agree, Neutral, Disagree, Strongly disagree, and I don't know*). To develop content validity and improve the clarity of the wording, each item on this scale was checked and slightly modified through several rounds of the panel discussion and pilot testing with two hospital nurses, two communication teachers, and two college nursing students (all not part of the research). The pilot test established that the amount of time needed to complete the survey was reasonable. The content validity and reliability of the scale were established using the Content Validity Index (CVI) developed by Waltz and Bausell (1983). The participants in the pilot testing were asked to rate each item based on relevance, clarity, simplicity, and ambiguity on a 4-point scale (see **Table 4**). All the items in NEPS had a CVI score over 0.75, indicating minimal potential threats (e.g., confounding variables, selection bias) to the construct validity of the instrument.

In addition, to confirm the factor structure underpinning the questions for each of the three themes of NEPS, corresponding to the three scales we adapted them from, we conducted factor analyses on our survey data. Our factor analyses confirmed the construct validity of the three themes: each theme had only one underlying factor. The KMO value for each theme was above 0.7, and the cumulative percentage of variance for each theme was above 50%. The three themes were actually already validated

TABLE 4 | Criteria for measuring content validity (Waltz and Bausell, 1983).

1. Relevance

- 1 = Not relevant
- 2 = Item needs some revision
- 3 = Clear but needs minor revision
- 4 = Very relevant

2. Clarity

- 1 = Not clear
- 2 = Item needs some revision
- 3 = Clear but needs minor revision
- 4 = Very clear

3. Simplicity

- 1 = Not simple
- 2 = Item needs some revision
- 3 = Simple but needs minor revision
- 4 = Very simple

4. Ambiguity

- 1 = Ambiguous
- 2 = Item needs some revision
- 3 = Not ambiguous, but needs minor revision
- 4 = Meaning is clear

through factor analysis in the studies we adapted the questions from. Our study provides important confirmation of these earlier factor analyses. We also generated Cronbach's alphas for the three themes to measure their internal consistency (or reliability), since multiple Likert questions were used in each scale/theme. The results were all above the 0.7 thresholds for reliability for the questions on Theme 1, 2, and 3, respectively.

NEPS was distributed to students online through WeChat. The survey was self-administered, with brief written instructions for its completion. Before the survey, the teacher read the definition and the explanation of engagement for the participating students in the WeChat groups in the students' first language (i.e., Mandarin), in order to ensure they understood the concept of engagement. It took 15 min for students to complete the surveys each time through wjx (www.wjx.cn) using their mobile phones. It is an online software for designing questionnaires, administering them, and analyzing the results. Since all the questionnaire items were compulsory, it was set by wjx that the questionnaire could not be submitted unless all the items were answered. Thus, there were no missing data points for the questionnaires in this project. In order to gather more valid information, all the items on the questionnaires were written in the mother language of students (i.e., Chinese). The translation

process was the same as that for the interview questions, as described earlier.

Data Analysis

The questionnaires and interviews were subjected to quantitative and qualitative analyses, respectively. For the quantitative analysis of the questionnaires filled in by nursing students, using the statistical package SPSS, descriptive statistics were generated, such as the percentage of responses for each option for each item. The interviews were transcribed and input into NVivo, version 12, for qualitative analysis. The content of the transcripts was checked for accuracy by the interviewees together with the first author. Following the approach proposed by Braun and Clarke (2006), a thematic analysis of the transcripts was conducted. Themes regarding engagement were identified by re-reading the coded nodes and classifying sub-themes into broader patterns of meaning. A number of communication strategies promoting engagement practices were also identified in the data. To ensure the reliability of the coding framework, the interview data were independently coded by the first author and the second author according to a coding sheet. The two raters established inter-rater reliability of $k > 0.8$ using Cohen's Kappa coefficient.

Ethics

The study was approved by the ethics committee of the involved nursing college before the study began. Consent forms were distributed by the first author to and signed by all participants before data collection that clearly explained the research aims, the confidentiality of the data, how the data would be used, and the participant's right to withdraw their participation for any reason.

RESULTS

The qualitative analyses showed that nurses and students had markedly different views about engagement. The divergences were identified in the following three major areas: (1) the role of engagement in patient care; (2) ways of enacting or realizing engagement; (3) what engagement strategies actually are. The students' perceptions were further supported by the results obtained through statistical analysis of their survey.

The Role of Engagement in Patient Care

The nurse participants stated that engagement helped them improve patients' understanding, elicit patients' concerns and establish empathy and rapport with patients. For example, an experienced nurse named Lily (pseudonyms are used for all participants) shared her experience where her spoken engagement encouraged the patient to express their feelings and concerns and thus feel more relieved:

Once a patient with a tumor asked me, 'I am in so much pain. Am I dying soon?', after his body check report showed he was in poor health. I responded, 'Why do you think so?', to encourage him to vent his feelings. The patient then told me how sad and in pain he was. Then, I said, 'We have many patients who are in pain like you. We will help you. Do you remember when you first came here? Aren't you feeling better now?'... He calmed down after venting and

said, 'I feel better now. Thank you! I have to be brave.' (Lily, female, 5 years of working experience).

Providing patients an opportunity or a space to express their concerns or vent their feelings is therefore one strategy for engaging with patients. Lily first recognized the patient's distress and then empathetically provided him an opportunity to express himself by asking an open-ended question, "*Why do you think so?*" The patient then told her more about his feelings. Thus, the patient's concerns were successfully elicited by the nurse and his emotional state was explored. Further, to communicate her empathy effectively, Lily used "*a lot of patients*" to help the patient feel less isolated, and "*we*" to suggest that there was a community of nurses all working together for the good of patients ("*We have a lot of patients who are in pain like you*"). By communicating empathy, Lily had communicated her understanding of the patient's feelings and expressed her willingness to support the patient, even reassuring him with empathetic, rapport-building responses, such as "*We will help you*" and "*Do you remember when you first came here? Aren't you feeling better now?*" These expressions are manifestations of strategies of engagement that effective nurses use to help patients feel that they are not alone and to allow nurses to establish empathy and connection with patients.

A similar example was given by a nurse supervisor, Mary, to illustrate her belief that engagement can help nurses meet patients' emotional needs and increase patients' cooperation with clinical treatments:

I once gave an IV infusion to a patient and incidentally found that the wound on her right ankle was infected. I asked if she was in pain, and she suddenly cried and said, 'I have been here for a few days. No one asked me!'. I said, 'Don't worry! I will ask the doctor to put another dressing on your wound now'... In the following days, she trusted me a lot and cooperated with me during her treatments (Mary, female, 24 years of working experience).

In this case, the patient was in pain not because of her ankle but because of the loneliness of having no one to talk with about her condition. Because the nurse engaged with the patient by asking if she was in pain, the patient felt cared for and respected, as evidenced by her subsequent cooperation with the nurse. The nurse supervisor further communicated empathetic engagement by offering reassurance ("*Don't worry! I will ask the doctor to put another dressing on your wound now*"). Despite the brevity of this spoken interaction, Mary was able to use engagement strategies to elicit the patient's concerns and establish empathy, rapport, and mutual trust, and thus perhaps better clinical outcomes.

In contrast, some nursing students did not share this view of the importance of engagement. Based on their survey results, some students held negative attitudes toward the importance of engagement (Theme 1 in NEPS) and learning engagement skills (Theme 3 in NEPS). **Tables 5, 6** show the percentages of each response for each item in Theme 1 and Theme 3, respectively. As can be seen from **Table 5**, some students did not see the value of engagement with patients but instead would rather focus on learning procedural knowledge about clinical tasks. For example,

18.33% of the students agreed that acknowledging the patient's experience was not necessary for nurse–patient relationships (Item 14), which is a way to promote engagement. As one participating student said,

Compared with communication skills related to promoting engagement, it is much more important for us to focus on learning procedural knowledge of the different clinical tasks we would have to perform (Student 02).

Another example is related to the concept of small talk, an informal type of conversation. It can be about everyday life unrelated to the medical conditions of the patients (Bosher and Smalkoski, 2002), such as “What are you going to have for dinner?” Some students thought it was unnecessary for them to have small talk with patients during nursing procedures. One student said,

I don't think it is necessary to chat with patients during nursing procedures. It will interfere with my nursing work, and possibly arouse the dissatisfaction of the patients because they may think I am nosy (Student 31).

Unlike the students, many nurse participants believed that having small talk with patients was an important way to establish closer and more trusting nurse–patient relationships, relieve anxiety, assist in the healing process, and thus achieve better clinical outcomes. Frontline health care workers in Hong Kong, in a study by Slade et al. (2015), had a similar view, saying that chatting with patients about aspects of daily life can put patients at ease and reduce the professional distance between nurses and patients.

One salient fact we can derive from the tables is a considerable proportion of students held positive attitudes toward the value of engagement in patient care (Table 5) and positive opinions on learning engagement skills (Table 6). However, among the students who acknowledged the importance of engagement, the attitudes of nearly all of them were “agree” rather than “strongly agree.” For example, only one student (1.67%) strongly agreed that addressing patients' emotions and psychosocial issues was absolutely essential in health-solving today (Item 9). This is possible because they were not clear about what was important about engagement. In the subsequent interviews, it was hard for quite a few students who valued engagement to explain why they thought it was important to engage with the patients during nursing care. As one participating student said,

My first feeling is that engagement is important, because that's what everyone says, but when asked why, I couldn't explain it (Student 01).

Similarly, among the students who held positive attitudes toward learning engagement, the responses of the majority were “agree” while a few were “strongly agree”, suggesting that their degree of positiveness was not strong. For example, only three students (5%) strongly disagreed that learning engagement skills had not helped or would not help them elicit patients' concerns

(Item 16). The reason may be that students did not really know what engagement was and why they had to learn it. They chose positive responses only by intuition or common sense (i.e., what they thought were the expectations of society). As one student said in the interview,

It is important to learn how to engage with the patients is a clichéd view for me. It is not that important to me (Student 42).

Ways of Enacting Engagement

Another difference between nurses and students is that many students lacked confidence in enacting engagement as they did not know how to enact engagement (i.e., what strategies to use, or what to say) even when they were willing to do so. It was supported by statistical results generated from the questionnaire. Table 7 displays detailed data on students' perceived ability to execute engagement (Theme 2 in NEPS), with the percentages of responses for each item. This theme was included to gauge the nursing students' self-perceived ability to execute engagement with patients. Based on Table 7, the majority of students were not confident in their ability to engage communicatively with patients. For example, more than half of the students (51.67%) lacked confidence in their ability to talk to patients from different backgrounds, and only two of the students (3.31%) thought they could do a good job in this respect (Item 6). Many students further expressed their worries in the interviews:

What worries me most is that the patient can't understand or misunderstands what I'm saying (Student 34).

I know it is important to engage with patients, but I don't know how to (Student 03).

I will be most embarrassed to meet patients who don't like to talk. If I ask and then the patient simply answers the questions without adding much information, it is too awkward (Student 05).

When confronted with possible chances of enacting engagement, students and nurses expressed completely different ways of dealing with it. For example, many student interviewees said that they put minimal effort into responding to patients' questions, especially when they did not have an answer. As one of the students noted,

I am afraid that I will give patients the wrong information and mislead them (Student 02).

Unlike the students, nurses believed that cursory responses annoy patients and asserted that it was always important to engage with patients by actively addressing their questions even if there was no immediate explanation or answer to give (e.g., nurses could offer to find someone who had the answer). Frontline nurses believed that engagement strategies could go a long way toward addressing patients' emotional needs in such situations, making patients feel heard and respected.

Another difference is that many students felt that they could rely on basic medical terminology to communicate with patients,

TABLE 5 | Percentages of each response for each item on perceived importance of engagement (Theme 1).

Item	Attitude	<i>n</i>	%
Item 3: Checking for patient understanding is generally unnecessary*	I don't know	0	0%
	Strongly disagree	6	10%
	Disagree	30	50%
	Neutral	17	28%
	Agree	7	12%
	Strongly agree	0	0%
Item 7: Dealing with the emotional problems of patients is the responsibility of psychiatrists, psychologists and social workers, not nurses*	I don't know	1	1.67%
	Strongly disagree	4	6.67%
	Disagree	41	68.33%
	Neutral	10	16.67%
	Agree	4	6.67%
	Strongly agree	0	0%
Item 8: Good nurse–patient communication improves patients' health outcomes	I don't know	0	0%
	Strongly disagree	1	1.67%
	Disagree	1	1.67%
	Neutral	17	28.33%
	Agree	38	63.33%
	Strongly agree	3	5%
Item 9: Addressing patients' emotions and psychosocial issues is absolutely essential in health-solving today	I don't know	0	0%
	Strongly disagree	0	0%
	Disagree	0	0%
	Neutral	18	30%
	Agree	41	68.33%
	Strongly agree	1	1.67%
Item 14: Acknowledging the patient's experience is not necessary in nurse–patient relationships*	I don't know	2	3.33%
	Strongly disagree	5	8.33%
	Disagree	28	46.67%
	Neutral	14	23.33%
	Agree	11	18.33%
	Strongly agree	0	0

*Negative statement.

believing that using medical terms was equivalent to providing patient care. Nurses, by contrast, said that they always aimed to maximize patient comprehension, avoiding using jargon on purpose, as an engagement strategy to ensure the patient's comprehension. More examples of engagement strategies will be presented in the next section.

What Engagement Strategies Actually Are

The nurses believed that they could engage better with patients through the use of various engagement strategies, and had a clear understanding of what engagement strategies were, as illustrated in the examples given earlier. On the contrary, when nursing students were asked about engagement strategies, the majority did not really understand the concept and interpreted the question as being about etiquette, or how they should behave when talking with patients. This was clear from their responses in the interviews: they could only suggest being polite or kind to patients when interacting with them. Unlike experienced nurses, they did not know of any specific communication strategies

they needed to use to achieve the functions of engagement. This is probably because they had not received any training on this issue.

The engagement strategies identified from the interviews with the experienced nurses (listed in **Table 8**) can be divided into three categories corresponding to the three types or functions of engagement included in the definition of engagement: improving patients' understanding, eliciting patients' concerns, and establishing empathy and rapport. These will be discussed in the following subsections:

Improving Patients' Understanding

The nurse interviewees noted that both nurses' understanding of patients and patients' understanding of nurses were essential to achieving engagement. This two-way understanding was recognized in the definition of “*engagement*” given earlier. Strategies to improve patients' understanding mentioned by the

TABLE 6 | Percentages of each response for each item on views on the learning of engagement skills (Theme 3).

Item	Attitude	<i>n</i>	%
Item 1: In order to be a good nurse I must have good engagement skills	I don't know	0	0%
	Strongly disagree	4	6.67%
	Disagree	0	0%
	Neutral	6	10%
	Agree	29	48.33%
	Strongly agree	21	35%
Item 4: Developing my engagement skills is just as important as developing my knowledge of nursing	I don't know	0	0%
	Strongly disagree	0	0%
	Disagree	0	0%
	Neutral	11	18.33%
	Agree	37	61.67%
	Strongly agree	12	20%
Item 5: Learning engagement skills has helped or will help me understand patients	I don't know	0	0%
	Strongly disagree	0	0%
	Disagree	0	0%
	Neutral	5	8.33%
	Agree	47	78.33%
	Strongly agree	8	13.33%
Item 10: Learning engagement skills has improved or will improve my ability to communicate with patients	I don't know	0	0%
	Strongly disagree	1	1.67%
	Disagree	0	0%
	Neutral	6	10%
	Agree	46	76.67%
	Strongly agree	7	11.67%
Item 12: Learning engagement skills is fun	I don't know	2	3.33%
	Strongly disagree	1	1.67%
	Disagree	3	5%
	Neutral	26	43.33%
	Agree	28	46.67%
	Strongly agree	0	0%
Item 16: Learning engagement skills has not helped or will not help me elicit patients' concerns*	I don't know	2	3.33%
	Strongly disagree	3	5%
	Disagree	34	56.67%
	Neutral	9	15%
	Agree	12	20%
	Strongly agree	0	0

*Negative statement.

nurses included offering information that is easily comprehended and avoiding jargon. One nurse shared the following:

Patients always ask, 'What medicine is it?'. If I tell them the exact names of the drugs, they wouldn't understand them because I would be using medical jargon... Patients are likely to feel nervous, as such drug names are sometimes very long and complicated. This may make them feel they are seriously ill. We can just tell them the functions of the medicine with everyday expressions. For example, we can say, 'It is good for your lungs' or 'It can stop vomiting' (Connie, female, 30 years of working experience).

With the communication strategies mentioned above, simplifying medical information and using a layman's language without leaving out any essential information, nurses can better engage with patients and thus can ensure better patient compliance and outcomes.

Eliciting Patients' Concerns

The most common communication strategies nurses mentioned that they had used to elicit patients' concerns were asking open-ended questions (to allow patients free expression), asking focused questions (to limit the range of the response), employing probes (to gather further detailed information on a specific

TABLE 7 | Percentages of each response for each item on perceived ability to execute engagement (Theme 2).

Item	Attitude	n	%
Item 2: I find it difficult to express empathy with patients*	I don't know	0	0%
	Strongly disagree	0	0%
	Disagree	7	11.67%
	Neutral	25	41.67%
	Agree	24	40%
	Strongly agree	4	6.67%
Item 6: I lack confidence in my ability to talk to patients from backgrounds different from my own*	I don't know	3	5%
	Strongly disagree	0	0%
	Disagree	2	3.33%
	Neutral	24	40%
	Agree	27	45%
	Strongly agree	4	6.67%
Item 11: I find it hard to get my point of view across to patients*	I don't know	3	5%
	Strongly disagree	0	0%
	Disagree	2	3.33%
	Neutral	14	23.33%
	Agree	32	53.33%
	Strongly agree	9	15%
Item 13: While communicating with a patient, I am able to verbalize to the patient that I comprehend what is being said	I don't know	7	11.67%
	Strongly disagree	1	1.67%
	Disagree	6	10%
	Neutral	22	36.67%
	Agree	23	38.33%
	Strongly agree	1	1.67%
Item 15: I find it hard to elicit patients' needs and thoughts*	I don't know	6	10%
	Strongly disagree	1	1.67%
	Disagree	3	5%
	Neutral	26	43.33%
	Agree	22	36.67%
	Strongly agree	2	3.33%

*Negative statement.

topic), and exploring personal health-related habits or asking questions about sensitive topics. Nurse Nana said,

Sometimes, patients can't express themselves clearly and just say they are uncomfortable. Then, we need to ask for further details focusing on the specific area where the patient didn't feel well by asking focused questions or using probes (Nana, female, 19 years of working experience).

The nurse supervisor also gave an example of asking open-ended questions:

When I want to find out about patients' sleep quality, I will ask, 'How did you sleep last night?', in order to encourage patients to say more, rather than a yes-no question, 'Did you sleep well last night?' (Mary, female, 24 years of working experience).

Establishing Empathy and Rapport With Patients

The strategies applied by the nurses mentioned for establishing empathy and rapport with patients included: paying attention and responding to patients' cues that could indicate their feelings,

expressing understanding of a patient's feelings or situation, and offering empathetic or supportive comments, especially after learning about a patient's unpleasant or even miserable experiences, praising patients, and offering spoken reassurance. As the nurse supervisor described,

My patient once told me she had three miscarriages, and this was her fifth pregnancy. I said, 'Oh, it's really hard on you.' During the fetal heart monitoring, I asked her to feel baby's heartbeat and said, 'Listen, the baby's heartbeat is so powerful. The baby can feel that you are a great mom!' (Mary, female, 24 years of working experience).

Through the nurse supervisor's empathetic feedback ("Oh, it's really hard on you.") and praise for the patient ("... you are a great mom!"), the patient would have felt understood and supported. Another similar example was provided by nurse Apple, whose patient's tension was relieved after he was told that he was not the only person to have had blood in the urine after an operation:

TABLE 8 | Engagement strategies identified from nurses' interviews.

Types of engagement	Engagement strategies
Category 1: Improving patients' understanding	(1) Ensuring that offered information is easily comprehended (2) Avoiding jargon or spontaneously explaining it well
Category 2: Eliciting patients' concerns	(1) Asking open-ended questions (2) Asking focused questions (3) Exploring personal health-related habits or asking questions about sensitive topics (4) Employing probes
Category 3: Establishing empathy and rapport with patients	(1) Recognizing or creating the opportunity to show empathy when necessary (2) Exploring the patient's emotional state (3) Stating an intention to help and support (4) Expressing understanding of a patient's feelings or situation (5) Showing respect or expressing praise for a patient (6) Paying attention and responding to patient's spoken or non-spoken cues that can manifest their feelings (7) Offering empathetic or supportive comments (8) Offering spoken reassurance through which a sense of hope is created (9) Making small talk with the patients about some unimportant subjects during the procedures (10) Assisting patients to connect with their significant others (11) Responding actively to questions raised by the patients, even if the nurse does not know the answers

One of my patients told me nervously, 'I have blood in my urine,' and I reassured him by saying, 'Don't worry. It is normal because you've just had an operation' (Apple, female, 13 years of working experience).

The above two examples clearly illustrate that paying attention to patients' emotional states is an important part of a nurse's job. Both nurses interpreted medical results positively for the patients, adding empathetic and supportive comments to reassure patients and establish empathy with them.

New Engagement Strategies

Most of the communication strategies mentioned by the nurse participants, such as ensuring that information offered are easily comprehended (Van Zanten et al., 2007; Slade et al., 2015), can be found in the literature. However, several new communication strategies were suggested by the nurse participants in this study. First, they suggested the strategy of assisting patients to connect with their family members and friends. One nurse described how she helped an elderly patient adhere to the clinical treatments with this strategy:

The patient refused to take medicine after the surgery and wanted to go home immediately... because he missed his wife, who was too old to walk here to see him... I encouraged him to have a video chat with his wife... He finally agreed to stay in the hospital (Mary, female, 24 years of working experience).

By facilitating the video chat with his wife, the patient's emotional needs were addressed, and the nurse probably provided him relief from his anxiety, depression, and pain. The second new strategy identified is that of actively responding to questions raised by patients, even if the nurse does not know the answers. While many strategies identified in previous studies are aimed at eliciting patients' responses (Van Zanten et al., 2007;

Slade et al., 2015), this strategy is about ensuring that nurses' responses make patients feel heard and respected. Nurses should offer to find answers to questions raised by patients even if they are not able to immediately answer them, to reassure patients that the matter will be handled. As a nurse said,

Even if I don't know the answers, I will actively respond to the patients. I can say, 'I am not very sure about this. I will ask the doctor and come back to you later' (Lily, female, 5 years of working experience).

Although these communication strategies were not among those reported in the previous literature on nursing communication, our findings suggest that they can lead to positive nursing engagement.

DISCUSSION

The qualitative findings show that the paramount importance of engagement was recognized by all the interviewed hospital nurses. They stated that engagement with patients meant going beyond attending to patients' immediate symptoms or pain during nursing care. This attention to the various dimensions of patients' needs is at the core of engagement, and the nurses' reflections on their behaviors showcased typical engagement behaviors as defined in the previous literature (Roe et al., 2012; Bright et al., 2018). The analysis of frontline nurses' views in this current study established the following three main functions of nursing engagement: (1) to improve patients' understanding of what the nurse is saying, especially in relation to the patients' medical conditions, treatments, and procedures; (2) to elicit patients' concerns, through the use of communication strategies that encourage patients to express themselves more in terms of their different needs (e.g., physical, emotional, spiritual, and

environmental concerns); and (3) to establish empathy and rapport with patients, through communication strategies that convey to patients that they are understood, their fears and concerns heard, and that their value as people is recognized.

In contrast to the views of experienced nurses, the nursing students, based on their qualitative and quantitative findings above, tended to concentrate on learning procedural knowledge of different clinical tasks and did not fully see the value of engagement, nor understand how to enact engagement with patients. With regard to communication training, the divergent views between nurses and students highlight their different priorities. Frontline nurses see communication skills as key to providing comprehensive care to patients because of how effective communication strategies can be at enacting engagement. Nursing students, however, pay more attention to the rote learning of medical terminology, which they believe is how they can be better nurses. The nurse supervisor interviewed for this study said that this mistaken belief was common among students with limited clinical working experience, suggesting a gap in nursing communication education: students are not being taught the full range of communication skills necessary for successful engagement with patients. In traditional nursing education in China, for example, the focus is on teaching students to memorize medical concepts and jargon (Peng and Ran, 2009). Thus, there is a real need for nursing colleges to have focused communication courses that specifically teach nursing engagement skills. While training in medical procedures and terminology is important for nurses' collaborative work with doctors, training in the use of communicative engagement strategies is also essential for the achievement of effective patient care, and the two strands can be combined in redesigned nursing communication courses. The frontline nurses interviewed for this study expressed a wish for nursing students to learn engagement strategies as part of their training at nursing college before they enter the workplace, and strongly suggested that the course materials should involve the study of samples of authentic nurse-patient hospital interactions so that students can learn how to use communication strategies to engage with patients while simultaneously carrying out the necessary medical procedures and other nursing tasks.

The strategies discussed in this current study are not specific to China but are transferrable to other contexts and regions worldwide. As mentioned earlier, the majority of the strategies are consistent with those from previous studies, while the new strategies are also universally relevant. For example, the strategy of assisting patients to connect with their significant others is consistent with the Roy adaptation model (Roy, 2009, 2011), which nurses all around the world use to ensure holistic care, in which the relationships with others who are meaningful to patients play a crucial role.

CONCLUSION

This study is one of the first to explore Chinese nurses' and nursing student's perceptions of engagement in nurse-patient communication. A new scale, NEPS, was thus developed and validated and this scale can also be applied to explore stakeholders' views on engagement in the future. Using surveys

and interviews, the study explored frontline hospital nurses' and college nursing students' divergent views on engagement and suggested that these gaps should be addressed pedagogically in nursing college communication courses. While experienced nurses emphasized the importance of effectively engaging with patients through various communication strategies, the students showed little understanding of how to engage with patients, nor of the important role of engagement in nursing. There is a gap between what nursing students are taught in existing nursing communication training, and what they need to do in their clinical placements. Given proper training and practice in the relevant communication skills, nursing students will be better equipped to engage effectively with patients once they graduate, in spite of the time constraints and stressful conditions that sometimes make being patient-centered quite challenging. Engaging with patients does not necessarily require a lot of time, and if nurses are well-trained, they will be able to use the necessary communication strategies automatically, even under stressful conditions, and thus be more effective as nurses.

The results have pedagogical and clinical implications for nurses and nurse educators everywhere. Nurse participants in this study thought it was important to provide students with authentic, contextualized samples of nurse-patient interactions collected from hospitals—a broad range of interactional nursing scenarios that students can learn effective engagement practices. Such a needs-aligned course will address the current gaps between what students learn and what they can use in clinical settings. It is the job of communication instructors at nursing colleges to teach these workplace-relevant skills, which can meet nursing students' occupational needs. The communication strategies identified in this study, embedded within a well-designed course that focuses on how they function to realize engagement, as exemplified by authentic interactional data, will help novice nurses quickly get up to speed once they are in the workplace and thus provide better nursing care, and, in turn, improved clinical outcomes. In terms of future directions extending from the current research, it is worth exploring whether more statistical methods can provide better insights concerning teaching engagement in nursing 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/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Huizhou Health Sciences Polytechnic and City University of Hong Kong. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

QH and JP designed the study and wrote the manuscript. QH was responsible for data collection, data analysis, and drafting of the

manuscript. JP revised and edited this manuscript. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors wish to thank all the participating nurses and students, and Dr. David Lee, Dr. Lindsay Miller (City University of Hong Kong), Ms. Qianwen Joyce Yu (City University of Hong Kong), Dr. Danyang Zhang (Shenzhen University), Dr.

Junjie Gavin Wu (Shenzhen Technology University) and Dr. Xinyue Xu (Nanyang Technological University) for their valuable suggestions on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.915147/full#supplementary-material>

REFERENCES

- Abdrbo, A. (2017). Assessment of nursing students' communication skills. *Nurs. Educ. Perspect.* 38, 146–151. doi: 10.1097/01.NEP.0000000000000126
- Bosher, S., and Smalkoski, K. (2002). From needs analysis to curriculum development: designing a course in health-care communication for immigrant students in the USA. *Engl. Specif. Purp.* 21, 59–79. doi: 10.1016/S0889-4906(01)00002-3
- Bramhall, E. (2014). Effective communication skills in nursing practice. *Nurs. Stand.* 29, 53–59. doi: 10.7748/ns.29.14.53.e9355
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Bright, F. A. S., Kayes, N. M., McPherson, K. M., and Worrall, L. E. (2018). Engaging people experiencing communication disability in stroke rehabilitation: a qualitative study. *Int. J. Lang. Commun. Disord.* 53, 981–994. doi: 10.1111/1460-6984.12409
- Candlin, S., and Candlin, C. N. (2014). "Presenting in the context of enhancing patient well-being in nursing care," in *The Routledge Handbook of Language and Health Communication*, eds H. E. Hamilton and W. S. Chou (London: Routledge and Taylor and Francis Group).
- Coates, C. J. (1997). The caring efficacy scale: nurses' self-reports of caring in practice settings. *Adv. Pract. Nurs. Q.* 3, 53–59.
- Corlett, J. (2000). The perceptions of nurse teachers, student nurses and preceptors of the theory-practice gap in nurse education. *Nurse Educ. Today* 20, 499–505. doi: 10.1054/nedt.1999.0414
- Dong, F., Zheng, R., Chen, X., Wang, Y., Zhou, H., and Sun, R. (2016). Caring for dying cancer patients in the Chinese cultural context: a qualitative study from the perspectives of physicians and nurses. *Eur. J. Oncol. Nurs.* 21, 189–196. doi: 10.1016/j.ejon.2015.10.003
- Farrell, G. A., Cubit, K. A., Bobrowski, C. L., and Salmon, P. (2007). Using the WWW to teach undergraduate nurses clinical communication. *Nurse Educ. Today* 27, 427–435. doi: 10.1016/j.nedt.2006.07.005
- Hafner, C., and Miller, L. (2019). *English in the Disciplines: A Multidimensional Model for ESP Course Design*. London: Routledge.
- Jiang, L. L., Ruan, H., Xiang, X. J., and Jia, Q. (2015). Investigation and analysis of the caring attitude and behaviour of nurses in Shanghai, China. *Int. J. Nurs. Pract.* 21, 426–432. doi: 10.1111/ijn.12287
- Langille, D. B., Kaufman, D. M., Laidlaw, T. A., Sargeant, J., and MacLeod, H. (2001). Faculty attitudes towards medical communication and their perceptions of students' communication skills training at Dalhousie University. *Med. Educ.* 35, 548–554. doi: 10.1046/j.1365-2923.2001.00921.x
- Lin, E. C., Chen, S. L., Chao, S. Y., and Chen, Y. C. (2013). Using standardized patient with immediate feedback and group discussion to teach interpersonal and communication skills to advanced practice nursing students. *Nurse Educ. Today* 33, 677–683. doi: 10.1016/j.nedt.2012.07.002
- Paltridge, B., and Phakiti, A. (2015). *Research Methods in Applied Linguistics: A Practical Resource*. London: Bloomsbury Academic.
- Pangaltsos, A. (2011). *Medical Thought*. Thessaloniki: M. Barbounaki.
- Peng, L., and Ran, S. J. (2009). Present situation of medical humanities quality education in China. *Northwest Med. Educ.* 5, 929–931. doi: 10.3969/j.issn.1006-2769.2009.05.041
- Pun, J. K. H., Chan, E. A., Wang, S., and Slade, D. (2018). Health professional-patient communication practices in East Asia: an integrative review of an emerging field of research and practice in Hong Kong, South Korea, Japan, Taiwan, and Mainland China. *Patient Educ. Couns.* 101, 1193–1206. doi: 10.1016/j.pec.2018.01.018
- Qasem, M. A. N., and Gul, S. B. A. (2014). Effect of items direction (positive or negative) on the factorial construction and criterion related validity in Likert scale. *Asian J. Res. Soc. Sci. Humanit.* 4, 114–121. doi: 10.5782/2223-2621.2014.17.3.77
- Rees, C., Sheard, C., and Davies, S. (2002). The development of a scale to measure medical students' attitudes towards communication skills learning: the communication skills attitude scale (CSAS). *Med. Educ.* 36, 141–147. doi: 10.1046/j.1365-2923.2002.01072.x
- Roe, Y. L., Zeitz, C. J., and Fredericks, B. (2012). Study protocol: establishing good relationships between patients and health care providers while providing cardiac care. Exploring how patient-clinician engagement contributes to health disparities between indigenous and non-indigenous Australians in South Australia. *BMC Health Serv. Res.* 12:397. doi: 10.1186/1472-6963-12-397
- Roy, C. (2009). *The Roy Adaptation Model*. New Jersey, NJ: Pearson Education.
- Roy, C. (2011). Extending the roy adaptation model to meet changing global needs. *Nurs. Sci. Q.* 24, 345–351. doi: 10.1177/0894318411419210
- Shao, Y. N., Sun, H. M., Huang, J. W., Li, M. L., Huang, R. R., and Li, N. (2018). Simulation-based empathy training improves the communication skills of neonatal nurses. *Clin. Simul. Nurs.* 22, 32–42. doi: 10.1016/j.ecns.2018.07.003
- Slade, D., Chandler, E., Pun, J., Lam, M., Matthiessen, C., Williams, G., et al. (2015). Effective healthcare worker-patient communication in Hong Kong accident and emergency departments. *Hong Kong J. Emerg. Med.* 22, 69–83. doi: 10.1177/102490791502200201
- Van Zanten, M., Boulet, J. R., and McKinley, D. (2007). Using standardized patients to assess the interpersonal skills of physicians: six years' experience with a high-stakes certification examination. *Health Commun.* 22, 195–205. doi: 10.1080/10410230701626562
- Waltz, C. F., and Bausell, R. B. (1983). *Nursing Research: Design, Statistics and Computer Analysis*. Philadelphia, PA: F.A. Davis Company.
- Woodrow, L. (2018). *Introducing Course Design in English for Specific Purposes*. London: Routledge.
- Yam, B. M., and Rossiter, J. C. (2000). Caring in nursing: perceptions of Hong Kong nurses. *J. Clin. Nurs.* 9, 293–302. doi: 10.1046/j.1365-2702.2000.00349.x

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 Huang and Pun. 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.



Learning English as a Foreign Language Writing Skills in Collaborative Settings: A Cognitive Load Perspective

Dayu Jiang^{1*} and Slava Kalyuga²

¹ School of Foreign Languages and Literature, Wuhan University, Wuhan, China, ² School of Education, University of New South Wales, Sydney, NSW, Australia

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Natasha Pourdana,
Islamic Azad University, Karaj Branch,
Iran

Paula Garcia De Freitas,
Federal University of Paraná, Brazil

*Correspondence:

Dayu Jiang
d.jiang@whu.edu.cn

Specialty section:

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

Received: 29 April 2022

Accepted: 03 June 2022

Published: 30 June 2022

Citation:

Jiang D and Kalyuga S (2022)
Learning English as a Foreign
Language Writing Skills
in Collaborative Settings: A Cognitive
Load Perspective.
Front. Psychol. 13:932291.
doi: 10.3389/fpsyg.2022.932291

Learning to write in a foreign language is a complex cognitive process. The process-genre approach is a common instructional practice adopted by language teachers to develop learners' writing abilities. However, the interacting elements of procedural knowledge, linguistic knowledge, and generic knowledge in this approach may exceed the capacity of an individual learner's working memory, thus actually hindering the acquisition of writing skills. According to the collective working memory effect, it was hypothesized that teaching writing skills of English as a foreign language by adopting a process-genre approach in collaborative conditions could lead to better writing performance, lower cognitive load, and higher instructional efficiency. The reported experiment compared learning writing skills of English as a foreign language in individual and collaborative instructional conditions from a cognitive load perspective, a rarely adopted approach in this field. The results indicated that the collaborative instructional condition was more effective and efficient than the individual instructional condition in improving the quality of written products as well as in optimizing the cognitive (working memory) load experienced by the learners. Measures of cognitive load were used to support the cognitive load theory's interpretation of the results, which is the unique contribution of this research study to the field.

Keywords: collective working memory effect, cognitive load theory, collaborative writing, teaching English as a foreign language, process-genre approach

INTRODUCTION

Learning to write in a foreign language is a complex problem-solving process, requiring not only a range of skills from writing English letters to composing complete essays but also the ability to make claims and provide appropriate supporting details (Kirkland and Saunders, 1991; Bruning and Horn, 2010; Howell et al., 2018). Students need to develop the skills of generating, organizing, and refining ideas by being involved in complex activities, such as brainstorming, discussing, outlining, drafting, monitoring, and revising (Raimes, 1992; Hyland, 2003a). Cognitive load theory aims at designing effective instructional materials and procedures to facilitate learners' acquisition of complex knowledge and skills based on the mechanisms of human cognitive architecture (Van Merriënboer and Sweller, 2005; Sweller et al., 2011). According to this theory, learners can build new knowledge about writing processes not only with the help of explicit formal instruction or

through personal reading but also using problem solving via individual or collaborative efforts (through personal introspection or pair/group discussions).

The collective working memory effect in cognitive load theory refers to the working memory space created by communicating and coordinating knowledge by each collaborator (Kirschner et al., 2011, 2018; Sweller et al., 2011). An individual who studies alone processes all the interacting elements of the instructional material in his or her working memory. By contrast, under a collaborative learning condition, all the interactive elements can be distributed among the working memories of group members. This effect allows a better understanding of cognitive processes in collaborative learning environments and the conditions under which such environments provide more efficient instructional options.

However, to our best knowledge, differences between the effectiveness of individual and collaborative instructional approaches in learning writing skills of English as a foreign language from a cognitive load perspective have never been investigated (Kirschner et al., 2011, 2018). Moreover, despite that collaborative writing as a teaching strategy has been actively implemented in foreign language classrooms since the 1990s (McDonough, 2004; Shehadeh, 2011), the issue of how developing writing skills in collaborative settings impact learners' cognitive characteristics has not been investigated extensively. In addition, more empirical research should be done to examine how learners in collaborative learning conditions would perform on individual writing tasks rather than on co-authoring tasks in the post-intervention phase (Storch, 2005; Chen, 2019). Accordingly, the experimental study reported in this paper was conducted in an attempt to fill these gaps.

MODELS AND APPROACHES TO TEACHING WRITING SKILLS

Cognitive Model of Writing Processes

Writing involves a range of cognitive activities. Flower and Hayes (1981) proposed a cognitive model of writing processes, which regarded writing as a decision-making process, consisting of a range of cognitive activities orchestrated in cyclical or recursive rather than linear orders (Racelis and Matsuda, 2013). Flower and Hayes (1981) argued that a writing process “involves three major elements which are reflected in the three units of the model: the task environment, the writer's long-term memory, and the writing process” (p. 369). This cognitive model generally corresponds to the three phases of writing: planning, translating, and revising phases. The three cognitive processes do not necessarily appear in a linear order but can happen at any moment in the writing process (Berninger et al., 1996; Baaijen and Galbraith, 2018, p. 196). Jones (2014) highlighted that the cognitive model of writing processes emphasized the functions of planning (i.e., generating ideas) and translating ideas into texts. Even though Flower and Hayes (1981) stressed that the three types of cognitive activities were recursive, they did not identify the “distinctions involving the temporal dimensions (before, during, or after translation) and spatial dimensions on

which the planning and reviewing/revising processes operate (whole text or a portion of it)” (Berninger et al., 1996, p. 198). The distinctions are of great significance to instructions as an awareness of stages or phases in writing could help learners internalize the phases of writing, which was evidenced in Jones' (2014) study that some of the participants were not fully aware of making distinctions between planning and translating while others were struggled with how to organize ideas in the writing process. It can be assumed that explicit instruction in planning and organizing ideas in the pre-writing stage could improve writing quality. Orchestrating the cognitive activities into stages or phases in this study attempted to actualize these abstract activities for instructional purposes. However, as Bizzell (1982) and Atkinson (2003) noted, this post-cognitivist approach to writing may neglect the genre nature of writings—shared features of texts shaped through social conventions. Therefore, it is of equal significance to teach genre knowledge when adopting the cognitive model of writing processes in teaching writing skills.

Approaches to Teaching Writing Skills

The genre approach and process approach to teaching writing skills have been used extensively to promote learners' abilities to write in English (Hyland, 2003a,b; Muncie, 2009; Keen, 2020). The process-based approach in writing instruction, which was introduced in the 1980s, usually consists of four stages: prewriting, writing, revising, and editing (Tribble, 1996). Participants in Keen's (2020) study adopted a process approach to learning skills: discussing topics in small groups, writing ideas about the topic, writing first drafts, carrying out peer reviews, writing second drafts, and sharing their accounts with the whole class. It was found that the participants developed a sense of ownership and learned how to write more effectively. Even though Keen (2020) used young learners of English as a first language as research subjects, he identified the beneficial role of procedural learning in cultivating students' writing abilities. However, it should be noted that such approaches demonstrate “how some writers write, they do not reveal why they make certain linguistic and rhetorical choices” (Hyland, 2003b, p. 19), as the process-based approach “is seen as predominantly to do with linguistic skills such as planning and drafting, and there is much less emphasis on linguistic knowledge” (Badger and White, 2000, p. 154). In a response, (Hyland, 2003b, 2008) put forward a genre-based approach to teach writing skills, in which genre is conceptualized as “a term for grouping texts together, representing how writers typically use language to respond to recurring situations” (p. 544). The genre-based approach emphasizes explicit instructions for communicative purposes, key language features, and structural patterns.

Graham and Sandmel (2011) advised that “advocates of process writing instruction integrate other effective writing practices into this approach” (p. 405). Researchers (e.g., Flowerdew, 1993; Badger and White, 2000) have endeavored to integrate the process-approach and genre-based approach in teaching writing skills of English as a foreign language as the two approaches could be mutually complementary (Raimes, 1991; Badger and White, 2000; Racelis and Matsuda, 2013; Deng et al., 2014; Huang and Zhang, 2020; Jiang et al., 2021;

Rahimi and Zhang, 2021). For example, Flowerdew (1993) introduced a process consisting of six types of activities to explicitly teach the process of learning specific genres. Badger and White (2000) proposed the process-genre approach to teaching writing skills, which consists of several stages starting from understanding a situation to completing a draft. By process-genre approach, Badger and White (2000) emphasized the significant roles of language skills, situational knowledge, and processes in cultivating writing abilities. Learning to write also means learning the techniques of self-regulating cognitive activities and procedures. Students who learn how to regulate the writing procedures collaboratively could transfer the knowledge when writing independently (Teng, 2020).

Learning English Writing Skills Through Collaboration

Taking a social stance, a process-genre approach to teaching writing skills encourages interactions and collaborations, which involves some kinds of collaborative activities such as “modeling, eliciting, supporting, probing, and suggesting alternatives or extension” to a learner’s initial attempts (Wette, 2017, p. 72). Dillenbourg (1999) and Prince (2004) defined collaborative learning as an instructional method through which students work together in small groups to pursue common learning or writing goals. Although collaborative learning, in general, has a long history of research, learning writing skills through collaboration was not actively implemented in foreign language classrooms until the late 1990s (McDonough, 2004). Learning writing skills through collaboration, with a primary aim of learning curricular content, focuses on both deconstruction and construction processes (Karnes et al., 1997). Granado-Peinado et al. (2019) found that participants who received collaborative practice and explicit instructions about writing synthesis identified more proportions of arguments and higher levels of integration of different sources than those in the collaborative practice conditions without instructions about writing synthesis. However, their research showed that providing collaboration opportunities does not sufficiently warrant effective learning, which also needs not only guides about how to collaborate but also explicit instructions about learning tasks. Accordingly, Teng (2020) investigated the effect of collaboratively modeling text structure and explicitly teaching self-regulated strategies on younger English learners’ abilities to write summarizations and essays. After 1-month intervention, it was found that participants who adopted self-regulated strategies and collaboratively modeled text structures demonstrated better performance than the participants in the control group in terms of the three measurements. It should be noted that the available research studies have reported mixed results about whether learning writing skills through collaborations could effectively improve the quality of written products or not (McDonough, 2004; McDonough and De Vleeschauwer, 2019; Matos, 2021). For example, some studies (e.g., Storch, 2005; Fernández Dobao, 2012; Hsu and Lo, 2018) indicated that texts written by collaborative learners were more grammatically accurate than those by individual ones. However, it has also

been reported that learners in the individual learning conditions produced more syntactically complex text than collaborative learners (McDonough et al., 2018). The divergent findings in the collaborative learning of writing skills can be related to the following three issues: the lack of explicit collaborative tasks in the learning phases, not considering cognitive aspects in the experimental designs, and not evaluating individual writing outcomes. Accordingly, Kirschner et al. (2009) recommended that research in collaborative learning should directly measure learning outcomes in a test condition, focus on one aspect of the learning goals at a time, and investigate the performance of individual learners instead of the group as a whole. They also advocated that research studies need to consider human cognitive architecture to better understand and compare individual and collaborative learning. In addition, (Berninger et al., 1996) noted that “working memory, and not only long-term memory, is involved in writing development” (p. 199), as the cognitive activities in relation to the task environment and writing process should be carried out in working memory.

COGNITIVE LOAD THEORY

Cognitive load theory aims at designing effective instructional materials and procedures to optimize learner cognitive resources in the process of acquiring complex knowledge structures (Sweller, 2010; Sweller et al., 2011). Cognitive load refers to the working memory resources needed for completing a particular learning task. Theoretically, learners may experience two types of cognitive load: intrinsic cognitive load and extraneous cognitive load (Van Merriënboer and Sweller, 2005; Sweller et al., 2011). Intrinsic cognitive load is defined as the working memory resources demanded by the innate complexity of information that a learner must learn (Sweller, 2010). Extraneous cognitive load, conceptualized as the working memory load that is unnecessary and extrinsic to instructional goals, is generated by the presentation manner and structure of the instructional material (Van Merriënboer and Sweller, 2005; Sweller et al., 2011).

The level of cognitive load experienced by the learners is determined by the level of element interactivity which refers to the degree to which information elements or components of a learning task should be processed simultaneously for meaningful learning (Sweller et al., 2011). For example, learning new vocabularies in a list can be considered as low in element interactivity, as individual vocabularies can be acquired without reference to other information in the list. By contrast, most writing tasks have high levels of element interactivity, as the writing process involves a relatively large number of interconnected elements of information, as well as cognitive, metacognitive, and socio-affective activities (Negari, 2011).

The levels of cognitive load that learners experience can be measured by subjective rating scales of effort, a simple and reliable instrument first adopted by Paas (1992). In this type of rating method, learners were asked to recall, reflect, and report the level of mental effort during their previous learning after they completed instructional activities. Even though subjective rating scales were capable of measuring the overall cognitive load,

researchers also needed information about the levels of particular types of cognitive load that learners experience (Paas et al., 2003; DeLeeuw and Mayer, 2008). Leppink et al. (2013) proposed a more recent version of subjective rating scales: three items on intrinsic cognitive load, three items on extraneous cognitive load, and four items on germane cognitive load. However, the results of confirmatory factor analysis in Jiang and Kalyuga's (2020) study showed that the two-factor (intrinsic and extraneous) model was an acceptable fit. Therefore, the cognitive load rating questionnaire in this study, which was developed on the basis of Leppink et al.'s (2013) version, adopted the two-factor model.

Cognitive load ratings are frequently combined with learning performance measures to calculate the relative instructional efficiency for different learning environments. Instructional efficiency in this study was calculated using Paas and van Merriënboer's (1993) formula $E = (P-R)/\sqrt{2}$, in which E stands for efficiency, P for performance z-score, and R for cognitive load rating z-score. In this study, the average of intrinsic cognitive load and extraneous cognitive load ratings were used to calculate the cognitive load z-score. According to this formula, higher values of instructional efficiency are achieved in situations where learning performance is high and cognitive load is low; lower values of instructional efficiency occur under conditions where learning performance is low and cognitive load is high.

Collective Working Memory Effect

Cognitive load theory considers a social interaction situation as a collective working memory system and extends the instructional focus from individual learning to collaborative learning. A collective working memory system can be developed from individual cognitive systems through collaboration, coordination, and communication. The collective working memory effect happens when learners acquire knowledge more effectively and efficiently through collaborating with others than through learning individually (Sweller et al., 2011). The collective working memory space constituted by multiple working memories has a larger capacity and longer duration than any of the constituents in individual working memories. This concept was supported by Dillenbourg (1999) who argued that in the collaborative conditions, "the horizontal division of labor into, for instance, task-level and strategy-level tasks, reduces the amounts of processing performed by each individual" (p. 10). Villarreal and Gil-Sarratea (2019) found that the texts produced by pairs were more accurate and grammatically complex than those by individual learners. They attributed the difference partially to collective scaffolding.

Collective working memory refers to the working memory space created by communicating and coordinating knowledge by each collaborator (Kirschner et al., 2018). An individual who studies alone processes all the interacting elements of the instructional material in his or her working memory. By contrast, under a collaborative learning condition, all the interactive elements can be distributed among the working memories of group members. The multiple working memories constitute a collective working memory space that has a larger capacity and longer duration than individual working memory. As a result, an individual learner in the collaborative instructional condition may experience lower levels of the cognitive load than a learner

who studies alone. The collective working memory effect, a recently developed cognitive load theory effect, occurs when learners learn better through collaborating with other learners than through learning alone (Sweller et al., 2011). This effect assumes that "students working in groups have more processing capacity than students working individually" (Janssen et al., 2010, p. 139). Even though interacting with group members in the collaborative learning condition may generate extraneous cognitive load, the interactive process should be beneficial as elaborating and eliciting could result in forming more advanced knowledge (Dillenbourg, 1999).

Under the individual learning condition, all the interacting elements of the learning task are processed in the individual learner's working memory. By contrast, learners who collaborate with others in their learning distribute all the interactive elements among the working memories of group members. Consequently, a collaborator would experience lower levels of the cognitive load than an individual learner. This assumption was supported by Zhang et al. (2011), who compared the effectiveness of collaborative and individual instructional approaches in learning the complex tasks of designing web pages. They found that the participants in the collaborative learning condition demonstrated better performance and experienced a lower level of the cognitive load than the individual learners.

Task complexity or element interactivity can influence the effectiveness of collaborative learning. For simple learning tasks, individual learning is expected to be more effective and efficient, as the transaction costs associated with sharing knowledge and coordinating communication will nullify the benefits offered by collaborative learning. By contrast, for complex tasks, the benefits offered by the collective working memory could be higher than the transaction costs, thus fostering efficient learning. Kirschner et al. (2009) found that individual learners performed better in remembering biological knowledge (simple tasks) than learners in collaborative conditions, whereas collaborative learners performed better in transferring the skills to solving similar problems (complex tasks) than individual learners. Similar findings were reported by Kirschner et al. (2011) who found that learning low-complexity biological tasks individually was more effective and efficient while learning high-complexity tasks benefited more from the collaborative approach.

EXPERIMENTAL STUDY

Learning writing skills of English as a foreign language has long been regarded as a complex process that usually generates a heavy cognitive load (Vanderberg and Swanson, 2007; Kellogg, 2008). Based on the review of literature on cognitive load theory and writing learning, the study was conducted to examine the following research hypotheses:

- (1) Participants taught through the process-genre approach in the collaborative learning condition would demonstrate better individual writing performances than participants in the individual learning condition.

- (2) Participants taught through the process-genre approach in the collaborative learning condition would experience lower levels of the cognitive load than participants in the individual learning condition.

The reported experiment focused on the effect of collaboration in creating a collective working memory among the members of a group. Previous research studies seldom included controlled randomized experiments and assessed learners' writing products as a means to evaluate the effectiveness of collaborative learning. Therefore, according to the collective working memory effect, the reported experiment was designed to test the hypotheses that learners of English as a foreign language in the collaborative process-genre instructional condition would achieve better individual learning outcomes in terms of writing skills, experience lower levels of cognitive load, and have higher instructional efficiency than learners in the individual process-genre instructional condition.

MATERIALS AND METHODS

Participants

The study adopted a purposive convenience sampling method; 64 undergraduate students (29 females) voluntarily participated in this experiment after reading the recruitment notice. They studied at a technological university in Shandong Province, China. They were also briefed about the aims, the procedures, their rights through the study, and their rights to access the research results. They were requested to return the signed consent form if they determined to participate. These college students were on average 21.5 years old and had spent 11 years learning English as a foreign language at the time of the experiment, so they could be regarded as having an intermediate level of English proficiency. They were randomly allocated into the individual learning condition (IL) ($n = 32$) and the collaborative learning condition (CL) ($n = 32$). The participants in the collaborative learning condition were further randomly allocated into eight groups with four members in each. This arrangement was based on the rationale that groups consisting of no more than six members could maximize participation by all group members (Herner et al., 2002).

The participants were required to write an essay as a pretest. The design of the pretest was based on Task 2 of the writing section in *International English Language Test System (IELTS): General Training*. Two independent raters examined their writings by complying with the *IELTS writing band descriptors*. These raters were proficient IELTS tutors with experience in applying the band descriptors in evaluating IELTS essays. An independent samples *t*-test indicated that the pre-test scores of the IL group ($M = 5.16$, $SD = 0.91$) were not significantly different from the CL group ($M = 5.00$, $SD = 1.02$), $t(58) = 0.61$, $p > 0.05$.

Materials

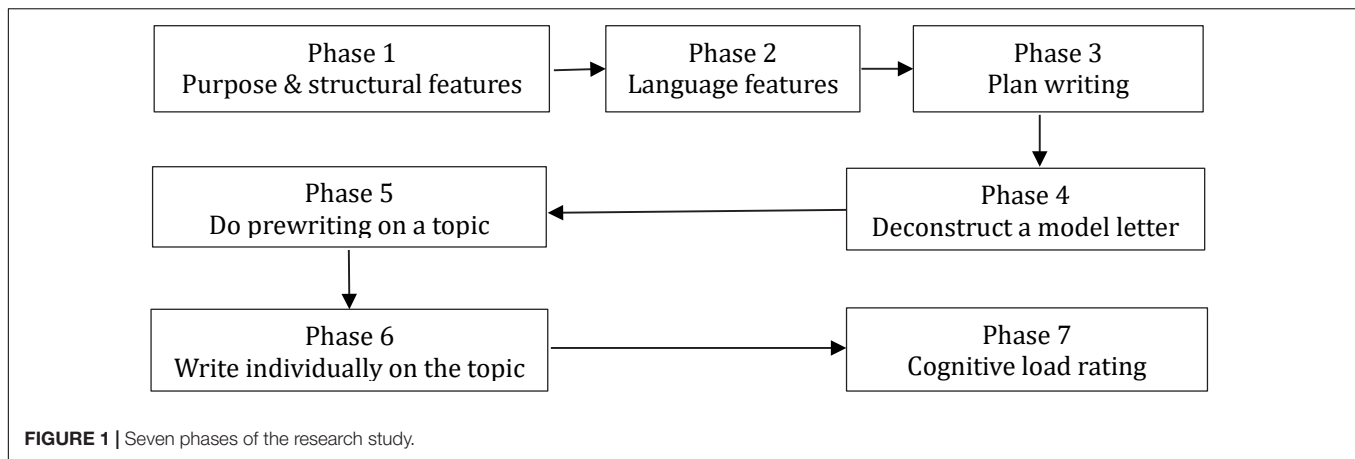
The instructional material was about how to write complaint letters. The development of the teaching material was based on the book *The Official Cambridge Guide to IELTS* authored

by Cullen et al. (2014). The experimental materials included four teaching components (structural features teaching, language features teaching, model essay teaching, and essay planning teaching), one essay planning phase, one testing phase (essay writing), and one subjective cognitive load rating phase (Appendix).

Procedures

The instruction was delivered in seven phases (see Figure 1). The participants in the individual learning condition were allocated to a lecture room. Each participant sat with at least 1-m distance from other participants to prevent collaboration and interference. The 84-member CL instructional groups were put in one lecture room. Each group kept a distance of at least 5 m from other groups to prevent collaboration and interference between groups, if any. The participants in the IL condition were required to complete all the seven phases individually; on the other hand, the participants in the CL condition completed the first five learning phases collaboratively, but the last two phases were completed independently. Associated questions were provided for thinking (for individual learners) and discussing (for collaborative learners) as Proske and Kapp (2013) argued that "learning questions might also be suitable to support the construction of a richly interconnected situation model of a writing topic which in turn may allow writers to produce better text products" (p. 1340). As it was generally believed that cognitive activities involved in writing procedures were recursive and dynamic (Flower and Hayes, 1981), the participants were reminded that they did not necessarily treat the phases as absolute linear orders and had the freedom to revisit the previous phase or skip to next one when they feel necessary.

The first part of the instructional materials (10 min) introduced the purposes and structural features of complaint letters, as well as the functions of each structural component. The structural features of complaint letters covered in this study include the following: the introductory paragraph elicits the purpose of complaint letters; the body paragraphs elaborate on the problems that letters are about and the suggested solutions; and the conclusion paragraph generally states the expectations and closes the letter. The associated questions for thinking (for individual learners) and discussion (for collaborative learners) were: *How do structural features reflect communicative purposes?* and *Are there alternative structures for this genre?* The second part (10 min) elaborated on the common language features of complaint letters, such as phrases and sentence structures for specifying the problem, outlining the consequences, making and justifying a specific claim, and so on, with the questions for thinking and discussion being: *Are there alternative ways to give reasons and solutions?* By using graphic organizers, the third part (10 min) showed the essential steps in planning writing. The question for introspection and discussion in this phase was: *If there exist alternative structures, how can these steps in essay planning be adapted to suit those structures?* The fourth part (15 min) introduced a model letter, in which the participants were required to identify the structural features, explain the functions of each feature, and the language features that were used for achieving the purposes. The associated questions in



this phase were: *What tenses have been used mainly in each paragraph?* and *Why tenses were used in these ways?* The fifth instructional phase (10 min) required the participants to plan a letter on a given topic and scenario. In these five phases, the participants in the collaborative instructional conditions were encouraged to learn the materials through collaboration, share their understandings, ask questions, and provide responses, while the individual learners were encouraged to talk to themselves or engage in an internal conversation. In the sixth phase (15 min), the participants were required to individually write a letter on the topic they discussed in the fifth phase by using the skills learned in the first four phases. The last phase of the experiment (5 min) was a subjective cognitive load rating questionnaire (**Appendix**).

Traditionally, subjective ratings of working memory load have proven to be able to collect reliable and valid estimations of mental load in a non-intrusive way (Jiang and Kalyuga, 2020). The cognitive load rating questionnaire was developed from the questionnaires designed by Leppink et al. (2013), with the first six items on intrinsic cognitive load and the last six items on extraneous cognitive load. The questionnaire was written in Chinese, the research participants' first language. The participants were asked to evaluate the appropriateness of a certain aspect of the instructional design that could orchestrate their mental resources to facilitate learning by choosing a number on a Likert-type scale, ranging from 0 (not at all the case) to 10 (completely the case). In addition, the instructor was available to clarify and explain puzzles and queries, if any.

Scoring

The quality of the letters was assessed according to the *IELTS General Training Writing Task 1: Writing band descriptors* published by the British Council. The band descriptors cover four categories: task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy. Each category has the 9-point scale, ranging from one to nine. Each letter was given one score for each category, and the sum of the scores in the four categories was the rater's score for the letter. The highest mark for a letter was 36. Two independent raters assessed students' letters. The average value of two raters' markings was used as the final score of the letter. The inter-rater reliability was calculated using

a Person intra-class correlation (ICC). The ICC measure of 0.92 indicated a high degree of inter-rater reliability.

RESULTS

Table 1 shows means and standard deviations of the letter scores, the scores of each category, the ratings of intrinsic, extraneous, and overall cognitive load, and the instructional efficiency for the two instructional conditions. The reliability of the subjective cognitive load rating scale as measured by Cronbach's alpha was 0.76.

An analysis of covariance (ANCOVA) was conducted to compare the two instructional groups' letter scores, scores of each subcategory, the ratings of intrinsic cognitive load,

TABLE 1 | Means and standard deviations for essay writing performance scores, individual category score, subjective ratings of cognitive load, and instructional efficiency for two instructional groups.

Group		Individual learning (N = 32)	Collaborative learning (N = 32)
Essay score	M	22.03	24.55
	SD	4.36	4.21
Task achievement	M	5.72	6.19
	SD	1.08	1.05
Coherence and cohesion	M	5.44	6.14
	SD	1.07	1.06
Lexical resource	M	5.50	6.03
	SD	1.15	1.07
Grammatical range and accuracy	M	5.36	6.19
	SD	1.13	1.08
Intrinsic load	M	2.21	1.59
	SD	0.99	0.99
Extraneous load	M	2.40	1.97
	SD	0.68	0.75
Overall load	M	2.30	1.78
	SD	0.61	0.69
Efficiency	M	-0.47	0.46
	SD	1.02	1.08

extraneous cognitive load, and overall cognitive load, as well as the indicators of instructional efficiency. Levene's test was conducted ($p > 0.05$) and the assumptions were satisfied. After controlling for the effect of pretest, it was found that the participants in the CL instructional condition demonstrated significantly better letter writing performance [$F(1, 61) = 27.40$, $p = 0.001$, $\text{partial}\eta^2 = 0.31$] and significantly higher instructional efficiency [$F(1, 61) = 31.97$, $p = 0.001$, $\text{partial}\eta^2 = 0.34$] than those in the IL instructional condition. In terms of category scores, the learners in the CL teaching condition significantly outperformed those learners in the IL teaching condition in all the four subscales: task achievement [$F(1, 61) = 15.72$, $p = 0.001$, $\text{partial}\eta^2 = 0.21$], coherence and cohesion [$F(1, 61) = 30.64$, $p = 0.001$, $\text{partial}\eta^2 = 0.33$], lexical resource [$F(1, 61) = 17.86$, $p = 0.001$, $\text{partial}\eta^2 = 0.23$], as well as grammatical range and accuracy [$F(1, 61) = 41.76$, $p = 0.001$, $d = 0.41$]. The participants in the IL instructional condition experienced significantly higher levels of intrinsic cognitive load [$F(1, 61) = 7.68$, $p = 0.007$, $\text{partial}\eta^2 = 0.11$], significantly higher levels of extraneous cognitive load [$F(1, 61) = 5.83$, $p = 0.020$, $\text{partial}\eta^2 = 0.09$], and significantly higher levels of overall cognitive load [$F(1, 61) = 12.02$, $p = 0.001$, $\text{partial}\eta^2 = 0.17$] than the participants in the CL condition.

The covariate, which is pretest in the study, was significantly related to the letter writing performance, which means that the participants in the CL condition had significantly better performance than the students in the IL condition in terms of the overall scores [$F(1, 61) = 143.44$, $p = 0.001$, $r = 0.84$] as well as the four subscales: task achievement [$F(1, 61) = 127.86$, $p = 0.001$, $r = 0.81$], coherence and cohesion [$F(1, 61) = 128.09$, $p = 0.001$, $r = 0.82$], lexical resource [$F(1, 61) = 125.52$, $p = 0.001$, $r = 0.82$], and grammatical range and accuracy [$F(1, 61) = 146.29$, $p = 0.001$, $r = 0.84$]. In addition, cognitive load ratings and instructional efficiency were related to the covariate, pretest. The correlation to the covariate, pretests, was also observed in intrinsic cognitive load, overall cognitive load, and instructional efficiency. Students in the CL instructional condition had lower cognitive load ratings and higher instructional efficiency than the participants in the IL condition: intrinsic cognitive load [$F(1, 61) = 5.49$, $p = 0.02$, $r = 0.28$], overall cognitive load [$F(1, 61) = 4.58$, $p = 0.036$, $r = 0.07$], and instructional efficiency [$F(1, 61) = 62.88$, $p = 0.001$, $r = 0.51$]. However, it should be noted that the covariate, pretest, was not significantly related to extraneous cognitive load [$F(1, 61) = 0.42$, $p > 0.05$], which indicates that the differences in participants' perception of extraneous cognitive load could be largely attributed to the dependent variable, instructional conditions.

DISCUSSION

The reported experiment was conducted to test the hypotheses generated by cognitive load theory that learners of English as a foreign language in a collaborative instructional condition would show better writing performance, lower levels of cognitive load, and higher instructional efficiency than learners in an individual learning condition. Even though relations to the

covariate, pretest, were observed, the results of the study generally supported the hypotheses. As for the first hypothesis, this randomized experimental study found that the students in the collaborative learning condition demonstrated higher overall post-test letter writing scores and higher subcategory scores (task achievement, coherence and cohesion, lexical resource, as well as grammatical range and accuracy) than the participants in the individual learning condition. The second research hypothesis was also supported as the participants in the collaborative learning condition indicated lower overall cognitive load ratings than the participants in the individual learning condition. It was also found that the collaborative learning condition generated higher instructional efficiency in terms of developing writing skills than the individual learning condition. Moderate and significant negative correlations were found between the ratings of intrinsic cognitive load and the letter-writing performance scores for both instructional conditions. The results demonstrated the collective working memory effect (Kirschner et al., 2009, 2018; Sweller et al., 2011) in the domain of learning writing skills by learners of English as a foreign language. As predicted by cognitive load theory, in the case of complex learning tasks such as writing in a foreign language, the benefits of collective working memory exceeded the possible disadvantages of dealing with transaction costs involved in coordinating individual working memories.

First, the study contributed to the research area of writing in a foreign language by conceptualizing the research and interpreting the findings from the perspective of cognitive load theory. In an attempt to account for the role of specific cognitive mechanisms in improving writing performance, it is possible to assume that the collaborative instructional approach had created an effective pool of knowledge about language and a pool of cognitive resources that beneficially influenced the quality of written products (Storch, 2005; Strobl, 2014). The interactions in collaborative instructional conditions could trigger more learning-relevant cognitive mechanisms, for example, knowledge elaboration and internalization which are essential for meaningful and effective learning. These learning mechanisms could enable learners to organize information into ordered structures and integrate new information with prior knowledge structures (Dillenbourg, 1999; Kalyuga, 2009). In the process of collaborative learning, theme-related knowledge structures would be retrieved from learners' long-term memory and function collectively as distributed cognition including "internal minds, external representations, and interactions among individuals" (Klein and Leacock, 2011, p. 133). The distributed cognition could evolve through members' contributions using stating claims, supporting or challenging others' opinions, providing supporting details, and so on. The mental activities in sharing, understanding, and negotiating meaning involve expressive or introspective elaborations, resulting in conceptual changes in group members (Dillenbourg, 1999). As more sources of information come to the group memory, learners would exercise more knowledge elaborations to establish links between new information and the existing knowledge structures, leading to better performance measures. The multiple learning phases in the collaborative

conditions offered collaborators more opportunities to use the language-related episodes (LRE) and task-related episodes, which were supposed to benefit their writing.

Second, the findings are consistent with the collective working memory effect, in that learning English as a foreign language writing skills in the collaborative instructional condition is more effective and efficient than in the individual learning condition (Kirschner et al., 2009; Retnowati et al., 2018). As learning tasks used for teaching English as a foreign language writing skills are high in element interactivity, and multiple factors (such as linguistic and situational knowledge, understanding of audience and purposes, etc.) affect the learning process, it can be assumed that the participants in each collaborative group would provide collective scaffolding, resulting in learning more sophisticated writing skills in terms of lexical accuracy, grammatical complexity, logic organization, and so on, in the learning phases and consequently in the better performance of these learners in the testing phase than the participants in the individual learning condition.

In addition, this study also indicates that adopting a process-genre approach in a collaborative condition could lead to significantly better writing performance than in an individual learning condition, which is particularly consistent with research studies on developing self-regulation of writing processes and generic knowledge through collaborations (e.g., Graham and Sandmel, 2011; Jones, 2014; Wette, 2017; Villarreal and Gil-Sarratea, 2019; Teng, 2020). According to the genre approach to teaching writing skills, effective instructional practices should “offer writers an explicit understanding of how texts in target genres are structured,” teach “the lexico-grammatical patterns which typically occur in its different stages,” and cultivate writers to command “an awareness of target genres and an explicit grammar of linguistic choices” (Hyland, 2003b, p. 26). However, if all lexical, syntactical, structural, and logical contents were taught without appropriate sequencing and prioritizing, high levels of cognitive load could be generated. Therefore, segmenting a learning task into several phases can ameliorate the complexity of information as the number of interacting elements would be reduced. For example, in a controlled randomized experiment, Klein and Ehrhardt (2015) found that organizing instructional tasks into manageable parts helped learners generate more balanced claims and reduced high-achieving students’ cognitive load in writing persuasion texts as measured by the perceived difficulty of their learning.

Furthermore, the results of the reported study are also consistent with previous research in the field of collaborative learning of writing skills (e.g., Shehadeh, 2011; McDonough et al., 2018), in that the learners in the collaborative instructional condition had better qualities of prewriting/writing performance than the learners in the individual instructional condition. Still, this study contributed to the area of collaborative writing research in two novel ways. First, differently from most of the previous research which required all learners in a collaborative group to write a common single text, this study required every member in a collaborative condition to write a separate text, and the quality of individual texts was assessed to compare the effectiveness of individual and collaborative learning conditions on the same

grounds. This method of measuring learning gains by assessing the quality of individual writing products is more valid and reliable according to Kirschner et al. (2009), as it better fits the learning goals. Second, the use of subjective ratings of participants’ cognitive load in learning and the calculation of instructional efficiency provided additional evidence to support a cognitive load interpretation of the results as the case of the collective working memory effect.

The reported study still has some limitations that require further research. First, this study did not consider the foreign language proficiency of the participants as a variable in collaborative teaching of English as a foreign language writing skill. According to the expertise reversal effect in cognitive load theory, the effectiveness of specific instructional techniques and procedures depends on the levels of the learner’s prior knowledge in the domain (Kalyuga et al., 2003; Kalyuga, 2007; Sweller et al., 2011). This effect has been demonstrated with all other instructional methods developed by cognitive load theory. It is likely that this effect also applies to the collective working memory effect. For example, Storch (2011) claimed that second language proficiency should be taken into consideration in implementing collaborative learning of writing skills. Therefore, future research studies may need to recruit learners at different proficiency (prior knowledge) levels to investigate possible interactions between levels of learner expertise in the area of English as a foreign language writing skills and the effectiveness of individual versus collaborative learning conditions. Second, this study examined the effectiveness of learning approaches (individual or collaborative) by primarily assessing the quality of learning products (i.e., essay). Future studies need to consider and measure other possible contributing factors and performance indicators, such as interactions in the writing processes, the quality of jointly drafted essays, and learners’ perceptions. In addition, as a way to manifest how collaborative learning affects the development of collective memory, future research should record and analyze learners’ interactions during the collaborative learning phases. Furthermore, more research should be done to investigate how learners develop their writing skills in other genres (such as argumentative, informative, and descriptive ones) in individual and collaborative instructional conditions.

In conclusion, the results of the reported experimental study supported the hypothesis generated by cognitive load theory. Learning English as a foreign language writing skills through a process-genre approach in the collaborative instructional condition was more effective and efficient than in the individual instructional condition. Subjective ratings of the cognitive load supported the interpretation of results within a cognitive load framework. The findings have implications for the innovations of teaching approaches, the developments of course materials, and curriculum designs in the field of teaching foreign language writing skills.

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. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Atkinson, D. (2003). L2 writing in the post-process era: introduction. *J. Second Lang. Writ.* 12, 3–15.
- Baaijen, V., and Galbraith, D. (2018). Discovery through writing: relationships with writing processes and text quality. *Cogn. Instruc.* 36, 199–223. doi: 10.1201/9781315152189-2
- Badger, R., and White, G. (2000). A process genre approach to teaching writing. *ELT J.* 54, 153–160.
- Berninger, V., Fuller, F., and Whitaker, D. (1996). A process model of writing development across the life span. *Educ. Psychol. Rev.* 8, 193–218.
- Bizzell, P. (1982). Cognition, convention and certainty: what we need to know about writing. *Pre/Text* 3, 211–245.
- Bruning, R., and Horn, C. (2010). Developing motivation to write. *Educ. Psychol.* 35, 25–37.
- Chen, W. (2019). An exploratory study on the role of L2 collaborative writing on learners' subsequent individually composed texts. *Asia Pacific Educ. Researcher* 28, 563–573. doi: 10.1007/s40299-019-00455-3
- Cullen, P., French, A., and Jakeman, V. (2014). *The Official Cambridge Guide to IELTS for Academic & General Training*. Cambridge: Cambridge University Press.
- DeLeeuw, K., and Mayer, R. (2008). A comparison of three measures of cognitive load: evidence for separable measures of intrinsic, extraneous, and germane load. *J. Educ. Psychol.* 100, 223–234.
- Deng, L., Chen, Q., and Zhang, Y. (2014). *Developing Chinese EFL Learners' Generic Competence: A Genre-Based and Process-Genre Approach*. Berlin: Springer.
- Dillenbourg, P. (1999). "What do you mean by collaborative learning?" in *Collaborative Learning: Cognitive and Computational Approaches*, ed. P. Dillenbourg (Oxford: Elsevier), 1–19.
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: comparing group, pair, and individual work. *J. Second Lang. Writ.* 21, 40–58. doi: 10.1016/j.jslw.2011.12.002
- Flower, L., and Hayes, J. (1981). A cognitive process theory of writing. *College Compos. Commun.* 32, 365–387. doi: 10.2307/356600
- Flowerdew, J. (1993). An educational, or process, approach to the teaching of professional genres. *ELT J.* 47, 305–316.
- Graham, S., and Sandmel, K. (2011). The process writing approach: a meta-analysis. *J. Educ. Res.* 104, 396–407.
- Granado-Peinado, M., Mateos, M., Martín, E., and Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Read. Writ.* 32, 2037–2058.
- Herner, L., Higgins, K., Pierce, T., and Miller, S. (2002). Study groups for active learning opportunities in preservice education. *Professional Educator* 25, 29–40.
- Howell, E., Hunt-Barron, S., Kaminski, R., and Sanders, R. (2018). Teaching argumentative writing to teachers and students: effects of professional development. *Prof. Dev. Educ.* 44, 169–189.
- Hsu, H., and Lo, Y. (2018). Using wiki-mediated collaboration to foster L2 writing performance. *Lang. Learn. Technol.* 22, 103–123.
- Huang, Y., and Zhang, L. (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.
- Hyland, K. (2003a). *Second Language Writing*. Cambridge: Cambridge University Press.
- Hyland, K. (2003b). Genre-based pedagogies: a social response to process. *J. Second Lang. Writ.* 12, 17–29.
- Hyland, K. (2008). Genre and academic writing in the disciplines. *Lang. Teach.* 41, 543–562.
- Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A., and Paas, F. (2010). Making the black box of collaborative learning transparent: combining process-oriented and cognitive load approaches. *Educ. Psychol. Rev.* 22, 139–154.

AUTHOR CONTRIBUTIONS

DJ: conceptualization, methodology, resources, data curation, writing – original draft, and review and editing. SK: supervision and writing – review and editing. Both authors contributed to the article and approved the submitted version.

- Jiang, D., and Kalyuga, S. (2020). Confirmatory factor analysis of cognitive load ratings supports a two-factor model. *Quant. Methods Psychol.* 16, 216–225.
- Jiang, L., Yu, S., Zhou, N., and Xu, Y. (2021). English writing instruction in Chinese students' experience: a survey study. *REL C J.* doi: 10.1177/0033688220980227
- Jones, S. (2014). From ideas in the head to words on the page: young adolescents' reflections on their own writing processes. *Lang. Educ.* 28, 52–67.
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educ. Psychol. Rev.* 19, 509–539. doi: 10.1007/s10648-007-9054-3
- Kalyuga, S. (2009). *Cognitive Load Factors in Instructional Design for Advanced Learners*. New York, NY: Nova Science Publishers, Inc.
- Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (2003). The expertise reversal effect. *Educ. Psychol.* 38, 23–31. doi: 10.1207/S15326985EP3801_4
- Karnes, M., Collins, D., Maheady, L., Harper, G., and Mallette, B. (1997). Using cooperative learning strategies to improve literacy skills in social studies. *Read. Writ. Q.* 13, 37–51.
- Keen, J. (2020). Writing revision: evidence for learning. *Changing English* 27, 121–136.
- Kellogg, R. (2008). Training writing skills: a cognitive-developmental perspective. *J. Writ. Res.* 1, 1–26. doi: 10.17239/jowr-2008.01.01.1
- Kirkland, M., and Saunders, M. (1991). Maximizing student performance in summary writing: managing cognitive load. *TESOL Q.* 25, 105–121. doi: 10.2307/3587030
- Kirschner, F., Paas, F., and Kirschner, P. (2009). Individual and group-based learning from complex cognitive tasks: effects on retention and transfer efficiency. *Comput. Hum. Behav.* 25, 306–314. doi: 10.1016/j.chb.2008.12.008
- Kirschner, F., Paas, F., and Kirschner, P. (2011). Task complexity as a driver for collaborative learning efficiency: the collective working-memory effect. *Appl. Cogn. Psychol.* 25, 615–624. doi: 10.1002/acp.1730
- Kirschner, P., Sweller, J., Kirschner, F., and Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *Int. J. Comput. Support. Collab. Learn.* 13, 213–233. doi: 10.1007/s11412-018-9277-y
- Klein, P., and Ehrhardt, J. (2015). The effects of discussion and persuasion writing goals on reasoning, cognitive load, and learning. *Alberta J. Educ. Res.* 61, 40–64.
- Klein, P., and Leacock, T. (2011). "Distributed cognition as a framework for understanding writing," in *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*, ed. V. Berninger (New York, NY: Psychology Press), 133–153.
- Leppink, J., Paas, F., van der Vleuten, C., van Gog, T., and van Merriënboer, J. (2013). Development of an instrument for measuring different types of cognitive load. *Behav. Res.* 45, 1058–1072. doi: 10.3758/s13428-013-0334-1
- Matos, F. (2021). Collaborative writing as a bridge from peer discourse to individual argumentative writing. *Read. Writ.* 34, 1321–1342.
- McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System* 32, 207–224. doi: 10.1016/j.system.2004.01.003
- McDonough, K., and De Vleeschauwer, J. (2019). Comparing the effect of collaborative and individual prewriting on EFL learners' writing development. *J. Second Lang. Writ.* 44, 123–130. doi: 10.1016/j.jslw.2019.04.003
- McDonough, K., De Vleeschauwer, 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
- Muncie, J. (2009). Finding a place for grammar in EFL composition classes. *ELT J.* 56, 180–186.
- Negari, G. (2011). A study on strategy instruction and EFL learners' writing skill. *Int. J. Engl. Linguist.* 1, 299–307.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *J. Educ. Psychol.* 84, 429–434.

- Paas, F., and van Merriënboer, J. (1993). The efficiency of instructional conditions: an approach to combine mental-effort and performance measures. *Hum. Factors* 35, 737–743. doi: 10.1177/001872089303500412
- Paas, F., Tuovinen, J., Tabbers, H., and van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* 38, 63–71. doi: 10.1186/s12909-015-0452-9
- Prince, M. (2004). Does active learning work? A review of the research. *J. Eng. Educ.* 93, 223–231. doi: 10.1002/j.2168-9830.2004.tb00809.x
- Proske, A., and Kapp, F. (2013). Fostering topic knowledge: essential for academic writing. *Read. Writ.* 26, 1337–1352.
- Racelis, J., and Matsuda, P. (2013). Integrating process and genre into the second language writing classroom: research into practice. *Lang. Teach.* 46, 382–393.
- Rahimi, M., and Zhang, L. (2021). Effects of an engaging process-genre approach on student engagement and writing achievements. *Read. Writ. Q.* doi: 10.1080/10573569.2021.1982431
- Raimes, A. (1991). Out of the woods: emerging traditions in the teaching of writing. *TESOL Q.* 25, 407–430.
- Raimes, A. (1992). *Exploring Through Writing: A Process Approach to ESL Composition*. London: St. Martin's Press.
- Retnowati, E., Ayres, P., and Sweller, J. (2018). Collaborative learning effects when students have complete or incomplete knowledge. *Appl. Cogn. Psychol.* 32, 681–692. doi: 10.1037/1082-989x.2.1.64
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *J. Second Lang. Writ.* 20, 286–305. doi: 10.1016/j.jslw.2011.05.010
- Storch, N. (2005). Collaborative writing: product, process, and students' reflections. *J. Second Lang. Writ.* 14, 153–173. doi: 10.1016/j.jslw.2005.05.002
- Storch, N. (2011). Collaborative writing in L2 contexts: processes, outcomes, and future directions. *Annu. Rev. Appl. Linguist.* 31, 275–288. doi: 10.1017/S0267190511000079
- Strobl, C. (2014). Affordances of web 2.0 technologies for collaborative advanced writing in a foreign language. *CALICO J.* 31, 1–18. doi: 10.11139/cj.31.1.1-18
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educ. Psychol. Rev.* 22, 123–138. doi: 10.1007/s10648-010-9128-5
- Sweller, J., Ayres, P., and Kalyuga, S. (2011). *Cognitive Load Theory*. London: Springer.
- Teng, F. (2020). Young learners' reading and writing performance: exploring collaborative modeling of text structure as an additional component of self-regulated strategy development. *Stud. Educ. Eval.* 65:100870.
- Tribble, C. (1996). *Writing*. Oxford: Oxford University Press.
- Van Merriënboer, J., and Sweller, J. (2005). Cognitive load theory and complex learning: recent development and future directions. *Educ. Psychol. Rev.* 17, 147–177. doi: 10.2307/23363899
- Vanderberg, R., and Swanson, H. (2007). Which components of working memory are important in the writing process? *Read. Writ.* 20, 721–752. doi: 10.1007/s11145-006-9046-6
- Villarreal, I., and Gil-Sarratea, N. (2019). The effect of collaborative writing in an EFL secondary setting. *Lang. Teach. Res.* 24, 1–24. doi: 10.1177/1362168819829017
- Wette, R. (2017). Using mind maps to reveal and develop genre knowledge in a graduate writing course. *J. Second Lang. Writ.* 38, 58–71. doi: 10.1016/j.jslw.2017.09.005
- Zhang, L., Ayres, P., and Chan, K. (2011). Examining different types of collaborative learning in a complex computer-based environment: a cognitive load approach. *Comput. Hum. Behav.* 27, 94–98.

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 Jiang and Kalyuga. 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.

APPENDIX

Subjective Rating of Cognitive Load

All of the following questions refer to the learning activity that you just finished. Please respond to each of the questions on the following scale (0 meaning not at all the case and 10 meaning completely the case).

No.	Question	Scale 0 = not at all the case; 10 = completely the case											
①	The topic/topics covered in the activity was/were very complex.	0	1	2	3	4	5	6	7	8	9	10	
②	The activity covered generic structure that I perceived as very complex.	0	1	2	3	4	5	6	7	8	9	10	
③	The activity covered language features that I perceived as very complex.	0	1	2	3	4	5	6	7	8	9	10	
④	The activity covered writing planning that I perceived as very complex.	0	1	2	3	4	5	6	7	8	9	10	
⑤	The activity covered model essay that I perceived as very complex.	0	1	2	3	4	5	6	7	8	9	10	
⑥	The activity covered brainstorming that I perceived as very complex.	0	1	2	3	4	5	6	7	8	9	10	
⑦	The instructions and/or explanations were, in terms of learning, very ineffective.	0	1	2	3	4	5	6	7	8	9	10	
⑧	The instructions and/or explanations on generic structure during the activity were very unclear.	0	1	2	3	4	5	6	7	8	9	10	
⑨	The instructions and/or explanations on language features during the activity were very unclear.	0	1	2	3	4	5	6	7	8	9	10	
⑩	The instructions and/or explanations on writing planning during the activity were very unclear.	0	1	2	3	4	5	6	7	8	9	10	
⑪	The instructions and/or explanations on model essay planning during the activity were very unclear.	0	1	2	3	4	5	6	7	8	9	10	
⑫	The instructions and/or explanations on brainstorming during the activity were very unclear.	0	1	2	3	4	5	6	7	8	9	10	



The Changing Role of Chinese English-as-Foreign-Language Teachers in the Context of Curriculum Reform: Teachers' Understanding of Their New Role

Man Lei¹ and Jane Medwell^{2*}

¹ Foreign Languages College, Shanghai Normal University, Shanghai, China, ² School of Education, University of Nottingham, Nottingham, United Kingdom

OPEN ACCESS

Edited by:

Claudio Longobardi,
University of Turin, Italy

Reviewed by:

Hongzhi (Veronica) Yang,
The University of Sydney, Australia
A. Moosnagh Auleear Owodally,
University of Mauritius, Mauritius

*Correspondence:

Jane Medwell
Jane.Medwell@nottingham.ac.uk

Specialty section:

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

Received: 25 March 2022

Accepted: 09 May 2022

Published: 06 July 2022

Citation:

Lei M and Medwell J (2022) The
Changing Role of Chinese
English-as-Foreign-Language
Teachers in the Context of Curriculum
Reform: Teachers' Understanding
of Their New Role.
Front. Psychol. 13:904071.
doi: 10.3389/fpsyg.2022.904071

The New Curriculum Standards for teaching English introduced major changes in the culture of teaching and learning English in the Peoples Republic of China (PRC). Changes have been linked to changing goals for English instruction and a revision of Confucian values in schooling. In this article, we argue that this English curriculum proposes a new role, with new demands, for English-as-foreign-language (EFL) teachers in the PRC. In order to implement the curriculum reform successfully, teachers involved in the reform are required to have a shared understanding of its nature, purposes and scope. However, little is known about to what extent EFL teachers understand and engage with their new roles. This study examines teachers' understandings of the new curriculum and of the new, demanding role of teachers implied by the curriculum. This is a mixed methods study involving an analysis of the curriculum document, a survey ($n = 227$) of EFL teachers and semi-structured interviews with a sample of teachers in the cohort ($n = 18$). The findings suggest that many teachers know the content of the curriculum document, but expressed uncertainty about the implications of changes, uncertainty about what a shift to student-centered teaching and learning means and confusion about new professional development demands. The findings of this study have wider implications for EFL teachers, teacher educators, researchers and policy makers in the PRC and similar national contexts. This article highlights that, from an international perspective, introducing new ideas and practices should consider teachers' existing understanding and experiences of the curriculum as well as the way in which they understand the purposes of the changes, and should promote a shared understanding of policy intentions.

Keywords: teacher's role, English-as-foreign language teachers, curriculum reform, China, humanistic values, teacher reflection

INTRODUCTION

Curriculum reform is seen as a dynamic and complex reality in teachers' professional lives (Vähäsantanen, 2015). EFL teachers are required to constantly develop their professionalism to better serve the goals of curriculum reform (Jiang and Zhang, 2021). Professional development of EFL teachers, however, involves not only continuous learning but also the cultivation of teachers'

new roles (Vähäsantanen and Eteläpelto, 2009; Yang, 2015; Tao and Gao, 2017; Jiang, 2022; Lei and Xu, 2022). The New Curriculum Standards (NCS) for teaching English in senior high schools (Ministry of Education [MOE], 2018) in the PRC were built upon 15 years of piloting. The NCS document proposes a new relationship between teachers and students, based on the centrality of Confucian humanistic values, to harness students' emotions and feelings so that they can take a more active role in learning English and engage in more communicative activities. To achieve this, the NCS demands significant changes in teachers' role, teaching practices, and their focus on students. Moreover, the NCS is open to interpretation in ways its predecessors were not, calling for teachers to be creative within the curriculum and to make curriculum choices (Mei, 2019). This raises the question of how well teachers understand and engage with their new roles. In order to implement a curriculum reform successfully, teachers involved are required to have a common understanding of its nature, purposes and scope (Vähäsantanen and Eteläpelto, 2009; Fullan, 2012; Yang, 2015; Lei and Medwell, 2020; Greenier et al., 2021; Lei, 2022). The present article reports the findings of an analysis of the NCS document focusing on two of the key areas of change proposed—the teacher's role and humanistic values. The paper then reports on a study of how teachers understand their new roles and the demands of the NCS, understandings which are at the core of enabling more than one million EFL teachers in China's high school education system to enhance their professionalism and empowerment, and to develop new and demanding approach to teaching.

TEACHERS' UNDERSTANDING OF EDUCATIONAL REFORM

Teachers' understandings of the key concepts of a change and their background training have a great impact on the implementation of an educational reform (Kırkgöz, 2008; Vähäsantanen, 2015; Tao and Gao, 2017). Researchers argue that if teachers are to successfully implement an educational reform, they must have a common understanding of the theoretical principles and classroom applications of the changes proposed by the reform (Vähäsantanen and Eteläpelto, 2009; Fullan, 2012; Yang, 2015; Lei and Medwell, 2020). Of the two, the latter often proves to be the most important, especially when teachers are poorly trained or lack sound knowledge of the educational reform. In the context of curriculum reform, it is very necessary and important to change the way teachers think about key components of an innovation, which is a more complex, inner, and implicit change (Fullan, 2012; Vähäsantanen, 2015). This emphasizes that the change in teachers' understandings and beliefs plays a significant role in the curriculum reform.

Teacher Training for the New Curriculum Standards

The NCS demands that EFL teachers improve their levels of professionalism through continuous learning in a number of areas and demands that they should constantly update their language knowledge and proficiency in order to be good

teachers in a modern society (Ministry of Education [MOE], 2018). Teacher training plays an important role in how far curriculum changes can be successfully implemented (Carless, 1998). Kırkgöz (2008) claims that teachers need guidance and opportunities to learn the new content and methods of communicating with learners, otherwise educational reform cannot be implemented successfully.

Chinese teachers' views of teaching are likely to be influenced by traditional teaching concepts (Jin and Cortazzi, 2006). Teacher training programs, therefore, ought to be capable of updating teachers' knowledge and of bringing about large-scale changes in teachers' existing beliefs to increase their awareness of the changes in the NCS in order to help them adapt to the innovation. However, teacher training in China is delivered mainly through short intensive courses attended by teachers on a selection basis and thus may not support all teachers. Even if short courses have a huge impact on some teachers, these may, without proper guidance, have difficulty understanding the new concepts or fall back on their previous teaching experiences and ignore the innovations (Fullan, 2012).

Moreover, as Ping (2010) pointed out, there is an identified problem for teacher trainers in China, in that their training programs tend to lack interactions, with trainees tending to be unresponsive. Ping (2010) describes such training as involving mainly passive classes with largely silent learners. This has been a frustrating experience for both trainers and trainees, leading to some unexpected results. For example, passive training is generally not stimulating or interesting for trainees; trainees may ignore the value of the things they have learned during the training program because of lack of motivation; the trainer may lack enthusiasm and energy when teaching the class; trainees may not understand the training content well because they tend not to interrupt the trainer with their questions. Such a passive learning approach may be the result of two features. The most important of these is the highly hierarchical nature of Chinese culture, which lowers learners' status and leads them to be passive recipients to receive whatever their higher status teachers transmit to them. The second reason is that questions or challenges from learners may put teachers at the risk of losing face because they may not have the correct answers (Biggs, 1996; Zheng, 2013; Li and Wegerif, 2014; Lei, 2022).

Relationship Between Teachers' Beliefs and Practices

Understanding the beliefs of teachers and their effects on teaching may be a key feature of the success of educational reforms (Fullan, 2012). Brown and Cooney (1982) defined beliefs as the key determining factors of an individual's action which guide their behavior. Teachers' beliefs about English language teaching (including teachers' understandings, attitudes, expectations, values, and theories about teaching and learning) are normally affected by their previous experience as learners at school; their experience as classroom observers; their teaching experience; their prior training experience (Vähäsantanen and Eteläpelto, 2009; Vähäsantanen, 2015; Yang, 2015).

Researchers have varied views on the relationship between teachers' beliefs and their classroom practices. Some studies suggest a consistent relationship between teachers' beliefs and practices, e.g., Pajares (1992) refers to beliefs as *messy constructs* but argues that there is a close relationship between beliefs and knowledge. He points out that teachers' beliefs are far more influential than their knowledge on the way they organize tasks and solve problems, on the kinds of decisions they make, on the way they plan lessons, and on the way they behave in the classroom. Many researchers (e.g., Burns, 1992; Breen et al., 2001; Gu, 2009; Fullan, 2012) have a similar view about the importance of teachers' beliefs for EFL classroom practice, showing that teachers' practices tend to be highly consistent with their beliefs. EFL teachers bring their own beliefs to situations related to English teaching, and their beliefs are normally regarded as important predictors of their general classroom practice. Their concepts of teaching reflect their beliefs about teaching, affecting their understanding and attitudes and also guiding their behavior.

Teaching methods encapsulate the way teachers put their beliefs into classroom practice. Therefore, it is necessary to understand teachers' beliefs in order to design any professional development program that aims to change classroom practices (Medwell et al., 1999; Fullan, 2012). In the case of curriculum reform, Kennedy (1988, p. 329) suggests, "teachers may be required to change the way they think about certain issues, which is a deeper and more complex change." That is to say, it may be necessary and important to change teachers' beliefs in order to implement any educational reform.

However, some studies suggest that changes in teachers' beliefs, understandings, and attitudes are likely to follow changes in their behavior rather than determine it. For example, Huberman's (1981) study of a reading program innovation showed that initial teacher training and ongoing guidance needed to be provided for teachers to help them adapt to an innovation. All the teachers, trainers and administrators in that study suffered a period of high confusion and anxiety because of the introduction of the new program. According to Huberman (1981), after the new program started, the teachers still needed some time to link their behavior with the concepts of the program. Even 6 months later, the teachers still had little sense of why specific behavior patterns could lead to certain results.

Some research also indicates that there may be inconsistencies between teachers' beliefs and their observed practices (e.g., Basturkmen et al., 2004; Farrell and Lim, 2005) and several studies have even found no significant correlation between the two (e.g., Yim, 1993). This may be because there are many other factors that can greatly influence teachers' beliefs during their actual classroom practice (e.g., Farrell and Lim, 2005; Fullan, 2012). For example, there may be inconsistencies between beliefs and practices if the teacher is in the process of coping with changes in his/her beliefs before putting changes into actual practice, when some propositions are incompatible, or when there are multiple belief systems (Graden, 1996). Moreover, it has been suggested that different research methods can affect whether the findings indicate limited consistencies between teachers' beliefs and their practices (Basturkmen, 2012), but sophisticated methods do not necessarily indicate a high degree of correspondence either.

Indeed, the relationship between teachers' beliefs and their practices is complicated (Pajares, 1992); it can be described as *dialectical* rather than *unilateral*. Beliefs and practices can affect each other, i.e., beliefs can guide and shape behavior but reflections on experiences and behavior can influence (and possibly change) beliefs (Breen et al., 2001). The study reported here takes teachers' experiences into consideration when examining teachers' perceptions and needs in relation to the NCS.

AN OVERVIEW OF THE NEW CURRICULUM STANDARDS

The NCS proposes some very significant changes to the role of EFL teachers in resource planning, classroom teaching, choice making, and approach toward students rather than demanding that teachers should strictly comply with textbook-based, top-down curriculum delivery. In contrast to previous attempts to introduce communicative language teaching, which have been criticized as imposing western approaches in China (Hu, 2005), the NCS does put more emphasis on the tenets of Confucianism, of mutual respect and care (Ministry of Education [MOE], 2018; Cheng and Zhang, 2020). The "figure of Confucius has been a key feature of revisionist assessments" of education (Lei, 2020, p. 174), and Confucius still plays an important role as a key cultural philosopher of the modern PRC, influencing education in China and other cultures, particularly the student-teacher relationship (Cheng, 2017). Confucius set a lofty example for all teachers to emulate, and is often referred to as "an exemplary teacher for all ages" and "the greatest sage and teacher" (Rao, 1998, p. 49). From his own teaching practice, Confucius proposed the requirements for being a good teacher in Chinese culture and these values still have a very important influence on understandings about the role of teachers in modern East Asian societies, which are often referred to as *Confucian Heritage Cultures* (Biggs, 1996; Yu, 2011; Li and Wegerif, 2014). Until quite recently Confucius' legacy was interpreted in ways that positioned teachers as holders of all knowledge, the center of the classroom and unquestionable (Hu, 2002; Cottine, 2016). English teaching across the PRC, in the past, was teacher-centered and teacher-dominated (Cortazzi and Jin, 1996; Jin and Cortazzi, 2006; Zheng, 2013; Lei and Medwell, 2020) by teachers who were good models, holders of profound knowledge and in unquestionable command of the class. However, "*good teacher*" is a cultural construct which can change over time and the NCS signals a massive challenge to the traditional Chinese construct of a good teacher. The underpinning for this, Confucianism, has become more popular in the past two decades in the PRC (Deng and Smith, 2018) and has interpreted the role of the teacher rather differently, with the focus placed more on the development of students' "humanistic values" by their teachers. To achieve this, the "good teacher" presented in the NCS has features of the teacher as guide, organizer, reflective teacher, rather than merely the dominant actor in the classroom. The EFL teacher is seen as engaged in a collective enterprise of learning with the students (Ministry of Education [MOE], 2018), rather than being only an authoritative exemplar. This new vision of the teacher and teaching is based,

at least in part, on the promotion of humanistic values in the NCS, a term which has strong Confucian resonances and is used many times in the NCS document. “Humanistic values” is clearly explicated in the NCS as “learning English for not only economic purposes – but also for the cultivation of students’ positive virtues and personality traits, and the development of their world-views, in individuals and the society to which they belong” (Ministry of Education [MOE], 2018, p. 3). This demands EFL teachers consider students’ concerns, feelings, in order to inspire them to form positive attitudes to learning and develop a healthy personality (Mei and Wang, 2018; Ministry of Education [MOE], 2018). Humanistic values also refer to another aspect of the Confucian tradition of valuing “humanity,” which refers to the support and care for one’s students and developing a Confucian concern for the harmony of the students, class and community (Cheng and Zhang, 2020).

As Hu (2002, p. 98) noted, “a fundamental assumption of Confucian tradition is that innate ability does not account for success or failure in education, and there is a strong belief that everyone is educable and capable of attaining perfection.” However, in the past, this has led to an emphasis on the responsibilities of students to study diligently. The NCS shifts responsibility for student progress to the teachers’ actions. EFL teachers are required to motivate all the students, prompt their positive attitudes in English learning process, and help all the students get of continuous progress and development (Ministry of Education [MOE], 2018). The 2003 curriculum document required EFL teachers to provide students with opportunities to facilitate their independent learning and learner autonomy. However, the expected teacher role documented in the NCS goes much further by demanding that teachers guide their students to participate in the learning process more actively through various collaborative work and help students to build their confidence within this collaborative learning process (Ministry of Education [MOE], 2018). The curriculum provides teachers with explicit examples to show them how to guide students in co-operative learning. Furthermore, the NCS suggests teachers should be aware of their responsibility to prompt learners’ willingness to cooperate with peers. This student-centered approach of taking account of students’ feelings, guiding and helping students through planning particular lessons and differentiating content may be a very new and different role to teachers who are accustomed to straightforward correction but it can be traced back to a Confucian emphasis on inculcating humanistic values and harmony in students.

Another area where the NCS seemed likely to challenge teachers’ existing practices and beliefs is in the view of how they evaluate their teaching. The NCS places much greater emphasis on EFL teachers reflecting on their teaching—what went well, not so well and so on—than the earlier approach, by illustrating, in detail, the importance and benefits of teacher reflection and inquiry. Reflection can help teachers to “identify and solve problems in professional work” (Ministry of Education [MOE], 2018, p. 28), so this is beneficial for teachers’ professional growth and development. But most importantly, teachers will undertake reflection not individually, but within peer groups to address the challenges and problems facing them in their

day-to-day professional lives. This emphasis on reflection suggests it can become a mechanism for developing teachers’ knowledge and understanding, and help them to adapt to the changes in their role.

The NCS encourages teachers to create a culture of cooperative learning with other teachers and cooperative inquiry that encourages communication, sharing with colleagues, and cooperative exploration (Ministry of Education [MOE], 2018). This was not included in the 2003 curriculum—the suggestion that teachers should work together as a learning community to prompt each other’s professional development and enhance their understanding and awareness of new role. This integration of teacher communication and reflection to prompt teacher professionalism has been explored in literature (e.g., Lai, 2010; Lei and Medwell, 2020) and underpinned projects like “lesson study-based action education” (Gu and Wang, 2003) to scaffold teacher learning within community. This view of teacher professionalism was a very significant change for teachers to take on.

CHALLENGES POSED BY THE NEW CURRICULUM STANDARDS FOR ENGLISH-AS-FOREIGN-LANGUAGE TEACHERS

The challenges posed by the NCS for EFL teachers can be summarized as three Cs – clarity, complexity, context (Fullan, 2007; Kırkgöz, 2008; Vähäsantanen and Eteläpelto, 2009).

Clarity here means whether teachers are able to identify essential features of the change (Fullan, 2012). Lack of clarity may refer to unclear objectives or unspecified ways of implementation. Even when the need for an innovation has been recognized, pinpointing what teachers need to do differently is always a barrier to the change process. Fullan (2012) pointed out that, in order to achieve clarity, the involved individuals need a sense of purpose that is explicit, shared, flexible, as they are required to adapt to changing circumstances constantly. However, explicit purposes are absent in much curriculum change documentation. Kırkgöz, 2008 found that most teachers in their study were unable to identify the main and key features of the curriculum. Lack of clarity of the curriculum, therefore, represents a major problem during the implementation phase. In our review of the NCS document, however, we found it did clarify its expected goals and gave suggestions and teaching examples to guide EFL teachers. This raised the question of whether the teachers in our study were able to identify the essential features of the NCS.

Complexity refers to the difficulty of and the extent to which practitioners are responsible for, implementation (Fullan, 2012). The complexity of an educational innovation has been discussed by many researchers (Carless, 1998; Vähäsantanen and Eteläpelto, 2009). Generally speaking, the nature of change is multidimensional and takes place in a particular context that includes political, social, economic and moral aspects. The organizations and individuals involved as well as the particular contexts are just a few of the factors in any change effort. More

specifically, concerning the actual components or dimensions of an innovation, the level of complexity mainly depends on the new materials, new teaching strategies and alteration of beliefs (Basturkmen et al., 2004; Fullan, 2012). According to Fullan and Pomfret (1977), changes that provide a certain level of complexity, but not to the extent that the level of adjustment required by teachers becomes overwhelming, are more likely to be implemented effectively in practice. However, changes that are incompatible with teachers' existing beliefs and strategies, and impractical or unpiloted are more likely to pose challenges to implementing the change (Carless, 1998). For example, changes in the use of new materials without any other changes in strategies may be a minor change. If a change includes other aspects, it tends to be rather more complex. The complexity of a change is not only a feature of the change itself but also a feature of teachers in terms of the discrepancy between the teacher's current practice and beliefs, and those assumed by the change. In the case of the NCS reform, the assumed changes are not just to materials but also to aspects such as teaching practices and beliefs. So, the NCS is not a minor change but a more complex one.

In this study, although the educational change is "top-down," it is impossible to say that the NCS is only curriculum change, because it calls for a change in teacher behavior and beliefs. The Ministry of Education published the NCS to change the experience of students and their learning outcomes, to solve the problems of teaching in practice and to cope with the difficulties of teacher development. This revised curriculum includes profound changes both to curriculum content and the role of the teacher (Ministry of Education [MOE], 2018). However, to activate the changes required by the NCS, teachers have to teach differently and need to know different things. Teachers may need to change their attitudes, views and teaching behaviors in order to meet the NCS' requirements. Because of the perceived dominant role of teachers in the classroom in a traditional Chinese teaching context and the influence of teachers' beliefs and experience, it could be difficult and challenging for teachers to adopt the new roles advocated in the NCS which threaten their authority in the classroom.

Analysis of the NCS exposes a fundamentally different approach to the curriculum and the teacher's role in it. The NCS offers a transformative view of what teachers do, implicitly requiring EFL teachers not just to change their teaching practices, but also the ways they learn about teaching. They are expected to adopt a more communicative approach and to develop their own professional knowledge through reflection and communication. Chinese EFL teachers whose training and professional practice comes from past traditions, might find the NCS unfamiliar and very challenging when required to adapt from being "good teachers" in the traditional context to "good teachers" in the modern society (Zhu, 2018), and it is not known what teachers understand about their new roles.

METHODS OF THIS STUDY

The research reported here is a part of a larger research project (between 2018 and 2021) addressing one key research question:

- How do EFL teachers understand their new roles and the demands of teaching the NCS?

The present study involved the administration of a questionnaire to 273 EFL teachers and subsequent semi-structured interviews with a sample of teachers within the cohort. Official approval from the collaborators' universities was sought before carrying out the study involving the teachers.

The questionnaire comprised two sections. The first section of the questionnaire collected the demographic data of the respondents. The second section asked teachers about the extent to which they knew about the key components of the NCS in relation to humanistic values, new roles (as organizer, guide, and reflective teacher), the requirements of the NCS (student-centered teaching and differentiation) as specified in the document. The questions in the questionnaire took the form of Likert Scale type responses to 28 statements about their understanding of their new role and their teaching practices (Tables 1–3). All the respondents answered the questionnaire anonymously and voluntarily. To make sure that the participants could understand the questionnaire, the items and format were piloted and revised with 24 EFL teachers (not included in the main study). The revised questionnaire was then distributed to EFL teachers in a prefecture-level city in Southern China. Cluster sampling was adopted to select the teachers and schools within this geographical location for the questionnaire. 273 questionnaires were distributed generating 227 returns, a very high return rate of 83%. The questionnaire sample had more female teachers (90%) than male teachers (10%). This is a reflection of female dominance in EFL teaching in the PRC (Beijing Normal University, 2013). The teaching experience of the teachers responding to the questionnaire ranged from 2 to 32 years in middle schools (Table 4).

The analysis of the data from the questionnaires began once the questionnaires were gathered and compiled. Participants were asked to choose from 1 "strongly disagree" to 5 "strongly agree" in response to each of 28 statements which represented an aspect of teachers' understandings and beliefs. The analysis of these responses was carried out using the Statistical Package for Social Sciences (SPSS) software. The validity of the questionnaire was achieved through the use of the expert validation method and peer debriefing (Lincoln and Guba, 1985) by asking a British professor and two colleagues to check the face and content validity. Face validity in this case refers not to what the questionnaire actually measured, but to what it superficially appeared to measure. This study was only concerned with how the questionnaire appeared to the subject users—i.e., whether to the ordinary person it looked as if it measured what it was supposed to—not the essential matter of what it really measured. The content validity of the questionnaire was evaluated by getting expert colleagues to assess whether the content of the questions reflected the intended variable or not. The comments from the expert and colleagues went a long way to ensure both the high face and content validity of the questionnaire.

After the questionnaire, in order to capture deeper insights into teachers' perceptions and ideas and to explore teachers' conceptions of teaching (Munn and Drever, 2004), we conducted

TABLE 1 | Teachers' responses to statements about humanistic values.

Statements	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
	n (%)	n (%)	n (%)	n (%)	n (%)
1. EFL teaching is not for not only economic purposes – but also for the cultivation of students' positive virtues and personality traits, and the development of their world-views	4 (2)	6 (3)	50 (22)	133 (58)	34 (15)
2. The NCS defines the nature of the English courses as the combination of humanistic and instrumental values	4 (2)	9 (4)	49 (17)	146 (64)	29 (13)
3. Teaching English to develop students' positive virtues and personality traits, and the development of their world-views is successful in practice	3 (1)	24 (11)	79 (35)	111 (49)	17 (7)
4. Teaching English for the cultivation of students' positive virtues and personality traits, and the development of their world-views is difficult	8 (3)	34 (15)	33 (14)	139 (62)	13 (6)

TABLE 2 | Teachers' responses to statements about teacher role.

Statements	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
	n (%)	n (%)	n (%)	n (%)	n (%)
5. English classes should be student-centered. Teachers should not dominate the class	4 (2)	6 (3)	15 (7)	140 (61)	62 (27)
6. Student-centered teaching is successful in practice	6 (3)	22 (10)	81 (35)	109 (48)	9 (4)
7. Making students as the center of the class is difficult	12 (5)	55 (24)	57 (25)	91 (41)	12 (5)
8. The EFL teacher's role is only to teach knowledge of foreign languages	43 (19)	118 (52)	21 (9)	39 (17)	6 (3)
9. The NCS emphasizes that EFL teachers should be the center of the class	53 (23)	157 (46)	27 (12)	36 (16)	8 (3)
10. EFL teachers should set realistic teaching objectives according to local teaching needs and students' language proficiency	4 (2)	9 (4)	11 (5)	105 (46)	98 (43)
11. The NCS emphasizes that EFL teachers should set realistic teaching objectives according to local teaching needs and students' language proficiency	5 (2)	3 (1)	21 (9)	151 (67)	60 (21)
12. Setting realistic teaching objectives according to local teaching needs and students' language proficiency is successful in practice	4 (2)	19 (8)	59 (26)	131 (58)	14 (6)
13. Setting realistic teaching objectives according to local teaching needs and students' language proficiency is difficult	6 (2)	49 (22)	33 (14)	117 (52)	22 (10)
14. EFL teachers should plan different resources and teaching methods according to students' different situations	4 (2)	8 (4)	18 (8)	85 (53)	75 (33)
15. EFL teachers should plan different resources and teaching methods according to students' different situations	4 (2)	8 (4)	18 (8)	85 (53)	75 (33)
16. The NCS emphasizes that EFL teachers should plan different resources and teaching methods according to students' different situations	4 (2)	1 (0)	30 (13)	158 (70)	33 (15)
17. Planning different resources and teaching methods according to students' different situations is successful in practice	7 (3)	25 (11)	78 (35)	101 (44)	16 (7)
18. Planning different resources and teaching methods to suit different students' situation is difficult	6 (3)	43 (6)	32 (33)	118 (55)	24 (3)
19. EFL teachers should undertake reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives	5 (2)	31 (14)	37 (16)	105 (46)	49 (22)
20. The NCS stresses that EFL teachers should undertake reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives	2 (1)	2 (1)	16 (7)	155 (68)	52 (23)
21. Undertaking reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives is successfully achievable in practice	5 (9)	32 (14)	68 (30)	107 (47)	15 (7)
22. Undertaking reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives is difficult	10 (4)	32 (14)	25 (11)	131 (58)	29 (13)

semi-structured face-to-face interviews in Mandarin Chinese with 18 volunteer teachers. A stratified sampling strategy was used to select the interviewees from 92 volunteers (Table 5).

Conceptions and beliefs may be implicit and tacitly held, and thus, they cannot always be determined by asking direct questions. Therefore, the interview employed a range of

TABLE 3 | Teachers' attitudes toward teacher training for the NCS.

Statements	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
	n (%)	n (%)	n (%)	n (%)	n (%)
23. I think the training program I have already had was lecture-based, spoon-fed the trainees, and lacked interaction	7 (3)	32 (14)	38 (17)	132 (58)	18 (8)
24. I think the training program I have already had was short intensive training	8 (3)	15 (7)	58 (25)	126 (56)	20 (9)
25. I think the training program I have already had was attended by teachers on a selective basis and did not cater for all teachers	4 (2)	25 (11)	32 (14)	109 (48)	57 (25)
26. I think the training program I have already had was closely linked to my actual situation of teaching in practice	7 (3)	36 (16)	61 (27)	100 (44)	22 (10)
27. The training I have already had for the NCS has helped me to develop my subject knowledge	6 (3)	10 (4)	72 (32)	120 (53)	19 (8)
28. The training I have already had for the NCS has helped me to improve my teaching practice	3 (1)	22 (9)	66 (29)	123 (54)	15 (7)

TABLE 4 | The characteristics of the questionnaire respondents ($n = 227$).

Statistics	Location			Gender		Education qualification			English teaching experience (years)					Training experience		
	City	Suburb	Rural	Female	Male	Junior college	Bachelor	Master	<5	6–10	11–15	16–25	>26	Enough	Not enough	No
<i>n</i>	80	70	77	205	22	38	173	16	53	57	87	17	13	36	105	86
%	35	31	34	90	10	17	76	7	23	25	38	8	6	16	46	38

TABLE 5 | The characteristics of the interview respondents ($n = 18$).

Statistics	Location			Gender		Education qualification			English teaching experience (years)					Training experience		
	City	Suburb	Rural	Female	Male	Junior college	Bachelor	Master	<5	6–10	11–15	16–25	>26	Enough	Not enough	No
Total number in the city	327 (33%)	307 (31%)	356 (31%)	881 (89%)	109 (11%)	218 (22%)	704 (71%)	78 (7%)	188 (19%)	277 (28%)	386 (44%)	109 (11%)	30 (4%)	59 (6%)	515 (52%)	416 (42%)
Number in interview	6 (33%)	5 (31%)	7 (36%)	16 (89%)	2 (11%)	4 (22%)	13 (71%)	1 (7%)	3 (19%)	5 (28%)	7 (39%)	2 (10%)	1 (4%)	1 (6%)	9 (52%)	8 (42%)

strategies and tools including scenarios and vignettes to elicit teachers' "inner" views, attitudes, and beliefs from comments and responses on stories depicting scenarios, individuals, and situations (Hughes and Huby, 2012). Although there is not much literature about the use of vignettes, especially within qualitative research or as a complementary method with other data collection techniques, researchers (e.g., Finch, 1987) offer similar descriptions of their use of vignettes in their research. Finch (1987, p. 105) for example, describes vignettes as "short stories about hypothetical characters in specified circumstances, to whose situation the interviewee is invited to respond." In other words, short scenarios in written or pictorial form are used to elicit participants' comments or opinions on examples of people and their behavior. An example, of a vignette involving Mr. Lin is appended (**Appendix**). This is a description of a fairly traditional teacher-centered class. After reading this vignette, participants were asked about their feelings and comments on the vignette and of Mr. Lin's views and actions. The interview questions were pilot tested with three teachers before embarking upon the study, to discover the appropriate probing and prompting techniques.

With the permission of the participants, all the interviews were audio-taped, transcribed, and then analyzed by using thematic coding.

The researchers adopted a descriptive-interpretative analysis approach to analyzing the data collected from interviews, that is, descriptive, recognizing that some interpretation was necessary in the analytical process (Maykut and Morehouse, 2002). The codes used to classify the data were, therefore, generated from the data. The researchers stayed focused on the answers to the research question while reading the transcripts and also made allowances for issues or concepts that emerged progressively. Because the NCS was a new curriculum with limited evidence thus far available, the whole analyzing process was based entirely upon the transcripts rather than preconceptions or pre-established codes (Braun and Clarke, 2006). During the coding process, the themes which had the potential to answer the research questions emerged. When new categories emerged, the researchers revisited the previously coded transcripts to identify any instances of the newly emerging categories that had not been noticed in the initial coding. To ensure the validity of this study, the researchers analyzed the data independently,

followed by discussions, and then constructed the final set of themes, i.e., teachers' knowledge of curriculum content; teachers' understanding of their new roles and the demands of teaching the NCS; and teachers' views of training for the NCS.

FINDINGS

The relevant results from the questionnaires and interviews are presented together, to emphasize the complex relationship between "knowing the curriculum" and "understanding the implications for the teacher's role." Indeed, it was obvious that most of the participants saw this research as an opportunity to have a voice about improving the quality of EFL teaching and learning and to solve the problems they had in their teaching practices.

Teachers' Knowledge of Curriculum Content

In the questionnaire, 98% (223) respondents reported that they were teaching the NCS. All the interviewees had previously taught the 2003 curriculum, and were teaching the NCS when they were interviewed. When all the 18 interviewees were asked, "how do you find the NCS different from the earlier curriculum?" nine teachers said that they had not noticed any differences. The responses were varied. The major differences identified by those who did identify differences were changes in what the curriculum demanded that students learn. Five respondents said that the biggest difference was that the NCS has reduced the learning requirements for students. Two noted the NCS had reduced the number of vocabulary items students were required to learn, and four mentioned that the new textbooks contained less lesson content. These responses recognized change in the textbooks and word lists, but did not recognize the implications, which were that the content of textbooks was expected to play a smaller role in lessons and teachers were intended to supplement the textbook content and vocabulary in ways to best suit the needs of students.

How English-as-Foreign-Language Teachers Understood Their New Roles and the Demands of Teaching the New Curriculum Standards

Understanding Humanistic Values in the New Curriculum Standards

At the heart of the NCS is the renewed emphasis on humanistic values, an idea clearly defined in the curriculum in relation to EFL teaching as teachers considering students' concerns and feelings, in order to inspire students to form positive attitudes to learning and develop a healthy personality (Ministry of Education [MOE], 2018).

Table 1 shows the agreement levels of respondents to items on this topic. The results suggest the vast majority of participating teachers knew about the proposed development of these values in English language teaching (77%), although only 56% believed that their teaching of humanistic development was successfully achievable in practice, and 68% agreed it was difficult. General

speaking, most respondents recognized the requirements of humanistic values in the NCS.

The interview results offer some insights into how teachers understood *humanistic values* as an important and new emphasis in the curriculum. The comments suggested that they linked the idea to a wide range of changes to teaching and learning. Of the 18 respondents, five said that the humanistic value of the English course meant stimulating students' interest in learning English.

Putting more emphasis on humanistic values is definitely good for students' practical use. Under the previous earlier curriculum, we were so concerned with teaching grammar and "teaching to the test" that we had little time to dedicate to "speaking." But now, the NCS encourages learning in happiness and using English in authentic contexts (T4).

However, not all of the respondents showed a positive attitude to the introduction of humanistic values into language teaching and twelve claimed they were unconvinced of the benefits underpinning this approach. Three of them identified humanistic values in the English course as meaning an emphasis on western culture and two said that the humanistic values included promoting humanities exchanges (cultural and educational exchanges). It was notable that many of the comments discussed the changes to textbooks.

Under the earlier curriculum, we did not teach students so much humanistic knowledge because the textbooks include very limited information about this. But now, the new textbooks include more humanistic knowledge, so we put more emphasis on humanistic education (T5).

The interviews suggested the idea of humanistic values was partially understood, remained challenging but that teachers supported the development of this approach.

Understanding the Requirements of the New Curriculum Standards (Student-Centered Teaching and Differentiation)

The questionnaire asked teachers to indicate levels of agreement to a number of statements about student-centered teaching. **Table 2** shows 88% of the teachers agreed that classes should not be dominated by teachers and should be student-centered, but only around half (52%) believing this was successful in their teaching practice, and 46% claimed it was difficult.

Respondents agreed (89%) that they should set teaching objectives to suit the students' learning needs and that the curriculum mandated this (88%), although only 74% believed this was successful in practice and 62% agreed it was difficult.

In terms of planning resources, which might be expected to be challenging for teachers who had relied on textbooks in the past, 86% agreed that teachers should plan resources to suit students' needs, and that the curriculum mandated this, although only 51% agreed this was successful in practice and 58% said it was difficult.

The interview findings revealed more about EFL teachers' views about student-centered teaching and a traditional, class dominating teacher role. One activity in the interview was to comment on a teaching vignette, where Mr. Lin, a "traditional" teacher, dominated the class. Of the 18 teachers, 16 said that they thought that Mr. Lin used a mainly teacher-centered teaching method, the traditional teaching method in the Chinese teaching

context and ten of them said that they used similar teaching methods in their classes.

I like Mr. Lin's teaching method of reviewing the content at the start of the class, and always ask students questions to make them concentrate. This is very similar to my teaching style in class (T3).

However, the picture was far from uniform. Eight teachers said that they did not agree with Mr. Lin's view of his role in the English classroom. They thought good teaching should be student-centered and EFL teachers should be organizers. For example,

I do not agree with Mr. Lin. His teaching is teacher-centered. I think a good teacher is not simply an explainer but also has other roles. The most important role is to organize the contents well and let students learn. Students should be the center of the class (T2).

There were some mixed responses where teachers discussed their own changes in practice. Five teachers said that they did not use the same teaching method as Mr. Lin but were more student-centered and used more communicative activities in class. Seven of the teachers reported that they taught very traditionally but also used some communicative activities because they said this presented good teaching. Four teachers' comments showed that they were strongly affected by the perpetuation of the examination-oriented culture in the Chinese education context. They stated that behavior and attitudes were heavily influenced by their school head teacher's focusing on exam results. They also pointed out that the evaluation of teachers' performance was also based on their students' exam results, which forced teachers to teach for examination purposes rather than adopt a role as organizer.

I understand that students ought to be the center of the class. However, this is demanding and the new approach takes lot of time and energy and cannot help my students to pass examinations. The role as an organizer of differentiated learning may not serve the examination purpose (T7).

The interview and questionnaire results suggested that the emphasis on student-centered teaching and differentiated learning in the NCS was something most teachers knew about, but which they saw as difficult and which they feared may not secure the desired examination outcomes.

Teachers' Views About Undertaking Reflections Within an Inquiry-Led Community

This project asked questions about how teachers learnt about the NCS and about their reflections about their teaching. Four teachers said that they often reflected on and summarized their teaching experiences after their English lessons. Thirteen teachers said that they did this only occasionally. Seven mentioned that they knew that reflection was helpful for improving teaching quality and students' performance, but only did it where a problem, like poor examination results, was identified.

I sometimes do reflection. Our school requires us to always reflect but we do not follow this suggestion. When students get disappointing exam scores, I discuss them my colleagues within the Teaching and Research group (TRG). I do not have enough energy and time to reflect after each lesson (T2).

The comments from teachers indicated that the activities they undertook included a great deal of reflection on their teaching, although they did not consider these to be reflections because they were not in a writing format. It was clear that this was a challenging area of practice for teachers, reinforcing the questionnaire results (Table 2) in which 91% agreed that "the NCS stresses that EFL teachers should undertake reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives" but only half (54%) agreed that "undertaking reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives is successfully achievable in practice."

The majority of the respondents (71%) agreed that "undertaking reflection within peer groups to address the challenges and problems facing them in their day-to-day professional lives is difficult" which indicates that they had some difficulty in implementing reflections within an inquiry-led community.

Teachers' Views of Training for the New Curriculum Standards

Training is not uniform in its effect and different teachers may perceive it differently. The findings from the questionnaires (Table 3) indicated that the training program teachers had received was "lecture-based, spoon fed, and lacks interaction" (66%), "short intensive" (65%) and "attended by teachers on a selective basis" (73%). Most of the teachers interviewed claimed that they did not get adequate training (17). Seven of the interviewees admitted that they had not learnt the NCS well and thus "did not implement the idea of humanistic values or any new roles, new changes in own professional practice" because they had "no sufficient guidance."

However, some of the comments about the training they had received suggested it was not designed to promote teacher autonomy or alert them to the implications of the proposed changes, but, instead, focused on the detailed content of the curriculum document:

Our TRG organizes the training for us. The director asked all the EFL teachers to sit down together and read the NCS content aloud by taking turns. So, we are learning the document together (T18).

This technique of reading the curriculum aloud could be a memorization technique, but it is a very particular view of what it means to understand the curriculum. This suggests that policy makers, trainers, and the heads of TRGs to think about the approach to training if it is to be focused on the particular needs of the teachers or the wider implications of the new teacher roles as proposed by the NCS.

T14 shared his difficulties in getting to grips with the changes. He felt that without training he was "experiencing trouble in understanding the new concepts in the NCS" and thus he "had no choice but fall back on previous teaching experiences." Teachers in this study felt able to confess that they did not understand the NCS but did not recognize many strategies for finding out changes themselves.

All 18 interviewees were asked “what training you would like to have?” They expected more interactive training with real teaching cases (7) and classroom observation with experts’ feedback (9) the most. As T12 mentioned, “by observing experts or successful teachers’ EFL classes, teachers could learn how to solve practical teaching problems and how to implement the NCS in the suggested and effective way.” The interview results suggested that the emphasis in the training “was on the content of the curriculum” (e.g., T9), but had not enabled teachers to understand or address the practical problems they faced.

These teachers are very concerned about their actual teaching practices and wanted more help and support for the substantial issues raised by the NCS – they wanted to relate the training to their own teaching practices.

DISCUSSION

As we argued earlier, teachers’ understandings of a change and their background training have a significant impact on the implementation of an educational reform (Kırkgöz, 2008; Vähäsantanen, 2015; Tao and Gao, 2017). We outlined in our review of relevant literature a theoretical framework to explore the issues of clarity, complexity, and context (Fullan, 2007; Kırkgöz, 2008; Vähäsantanen and Eteläpelto, 2009) that teachers face in getting to grips with a new curriculum document and changing their role as teachers. If teachers are to implement an educational reform effectively, they are required to have a shared understanding of the theoretical principles and classroom applications of the changes proposed by the innovation (Vähäsantanen and Eteläpelto, 2009; Fullan, 2012; Yang, 2015; Lei and Medwell, 2020). Within the context of curriculum reform, it is essential to change the way teachers think about key components of the innovation, which is a more complex and implicit change (Fullan, 2012; Vähäsantanen, 2015). This means that the change in teachers’ understandings and beliefs plays a key role in the curriculum reform.

A “Chinese culture of teaching and learning” has been widely reported in the past (Cortazzi and Jin, 1996; Hu, 2002) as teacher-dominated, characterized by undifferentiated learning tasks and student persistence. The analysis of the curriculum presented in this article argues that the NCS offers a significant challenge to this past characterization and reinterprets Confucian values in ways which place heavy demands on teachers. Analysis of the NCS identified these demands as a wider range of teacher strategies, a greater range of choice and decision making by teachers, new types and uses of assessment and a new approach toward students. We argue that the NCS for teaching English in China seeks to change the culture of English teaching in the PRC and transform the role of teachers.

Analysis of the NCS also reveals that being an EFL teacher now involves teachers in the design of the curriculum in new ways and places on them the expectation of professional development and being proactive in learning about changes. This is undoubtedly a challenge and it is to be expected that it will take many years to change the beliefs and practices of such a huge teaching force, even when the goals are expressed in Confucian terms and not

the language of CLT, which has been so problematic in China in the past (Cortazzi and Jin, 1996; Hu, 2002).

The results of the empirical part of this study of EFL teachers suggest that most teachers now teaching the NCS recognized changes in the content of the curriculum, when these were presented as statements. Moreover, many of the teachers in this study expressed broad agreement with the aims of the curriculum. However, their interviews suggest a much more complex picture. Half of the teachers interviewed did not recognize where the changes had been made or the implications of the changes for their own practices. Even when teachers did recognize changes, they evaluated changes in terms of their existing practices, such as the use and content of textbooks, and discussed assessment of the NCS as focused on goals they recognized from past curricula – performance in tests of reading and writing. These EFL teachers expressed a tension between doing what was best for their students in terms of examination outcomes and changing their practices to make students more independent learners who could use a wider range of resources. This tension between the longer-term goals of the NCS and the short-term assessment mechanisms has been recognized by many scholars (e.g., Kırkgöz, 2008; Yang, 2015; Mei and Wang, 2018; Zhu, 2018; Lei and Medwell, 2020) and extends to the role of the teacher.

The teachers in this study showed cautious enthusiasm for improving students’ experience and enjoyment of learning English, but the inclusion of humanistic values as a key motivator of students and teachers remained only partially understood. The teachers in this study broadly welcomed student-centered teaching, but saw planning activities to suit the needs of all children as impractical and they were suspicious of the practicality and efficacy of such differentiation. Teachers’ responses showed that their choices of teaching methods were determined by their language learning beliefs and the understanding of the NCS. The NCS specifies not only what is to be taught but also how to teach it – a syllabus with teaching guidance for teachers. However, in being specific, it also aims to empower EFL teachers to offer them more choices and options than they may have been used to in the previous, more textbook-based curriculum. In this way, the NCS is both more prescriptive than the old one, in that it specifies teaching methods, but also much less prescriptive because it gives teachers choices. However, many teachers in this study did not appear to recognize the need to develop their own resources, or go beyond the textbook. If this curriculum is to succeed in its goals of producing a “new teacher” who can teach English in student-centered ways, using authentic differentiated resources and seek out the professional development they need, there is some way to go in changing teachers’ views about their roles.

Finally, despite the strong criticism of the delivery and content of the top-down training in the literature (Ping, 2010) and by some teachers in the present study, most respondents had strong faith in the power of this learning approach. The interviews with teachers suggested that they participate in group activities and training and that they would like “more” training. However, the training they described included knowing the NCS, rather than understanding the implications of it.

Moreover, these teachers did not see themselves as responsible for finding out about and understanding new developments in the curriculum—something that may also be a significant change to the culture of education in the PRC. The top-down nature of the reform seemed to have left teachers feeling they were not “owning” but “implementing” the innovations. They did not appear to recognize their opportunities to prompt and influence their own reform practices (Vähäsantanen, 2015). This suggests that teachers, trainers and the heads of TRGs, should review their approach to training and the intentions of that training. There is also a need to explore more informal learning (e.g., reflection and self-study), which leaves more space for teacher agency, professionalism and empowerment (Vähäsantanen and Eteläpelto, 2009; Vähäsantanen, 2015; Yang, 2015; Tao and Gao, 2017).

The findings of this study could help to promote the authority, professionalization and ownership of teachers, and can better serve the goals of educational reform. Although the focus of this study is EFL teachers in the Chinese context, it has broader implications for policy makers, researchers, teacher trainers, heads of TRGs and teachers in similar national contexts. The present research highlights that, from an international perspective, introducing new ideas and practices must consider teachers’ existing understanding and experiences of the curriculum as well as the way in which they understand the purposes of the changes, and should promote a shared understanding of policy intentions. The teachers’ understanding of the roles of cooperative reflection, student-centered, and differentiated learning within EFL teaching in the PRC can lead to similar research with teachers from other subject areas and national backgrounds.

IMPLICATIONS

The challenges of the NCS for teachers were summarized earlier as challenges of clarity, complexity and context (Fullan, 2007; Kırkgöz, 2008; Vähäsantanen and Eteläpelto, 2009) and these challenges are all closely linked. The teachers in this study seemed to have achieved some clarity in their knowledge of the changes made to the NCS. However, there remained serious challenges in terms of their understanding of the complexity of the changes demanded of them. They have sought to interpret these, in some cases, in terms of their existing context, we would argue, the changes proposed to the NCS are a reflection of a much wider cultural change in the expectations of teaching and learning underpinning the curriculum.

Analysis of the NCS showed that it is not simply about changing teaching materials; it is about very profound changes in the goals, methods of English teaching and huge changes to the role of the teacher. The study demonstrated that teachers were engaged this process of reviewing how to implement the NCS, but not necessarily in understanding the underlying implications. The teachers faced challenges with some fundamental issues such as the shift to student-centered teaching and learning and the new role of the teacher. They also identified difficulties with the tension between the unchanged assessment and the new methods

in the NCS. The NCS presents a completely different vision of English teaching, a new, wider definition of the curriculum and a broader role for EFL teachers. It specifies not only what is to be taught but also a broader framework within which it is to be taught and the values behind that teaching. It specifies how students should feel about learning and what teachers’ roles should be – it is a syllabus, but also broader teaching guidance for teachers. This issue of complexity (Fullan, 2001, 2007) was not something all the teachers could engage with.

This study has also identified a lack shared understanding between teachers and the NCS of the word “curriculum” and teachers’ understanding and beliefs about the changes in the NCS. These teachers were finding it difficult to understand the notion of the teacher as professional who is more than a figurehead who simply “knows” and “delivers” the curriculum in the changed new curriculum processes, but the teacher is now regarded as “designing” and “creating” the curriculum for their own students. On this issue, the authors of the curriculum and the teachers do not share clarity of understanding.

This study also exposed clear issues of context which were very important to the teachers and mitigated against the changes which the NCS seems to promote. One issue identified by many of the teachers in this study was the unchanged nature of the assessment system in the context of a curriculum with such a different vision of curriculum and teachers. The emphasis of the NCS is not consistent with what is assessed in the examinations. The tension between the NCS and the examinations shows up most in the areas of speaking and listening, involving students in the curriculum designing process, and more formative assessment. As a result, EFL teachers teach what will be assessed in the examinations (e.g., grammar and vocabulary) rather than teach what is outlined and valued by the NCS. This is a very important context for those teachers, who see success in examinations as highly important. This dissonance between understanding and knowledge could impede teachers’ abilities to implement a change as it was intended (Vähäsantanen and Eteläpelto, 2009; Fullan, 2012).

Underpinning these implications, is the complex nature of Chinese middle school EFL teachers’ beliefs at a time when the NCS innovation meets the Confucian Heritage Culture and their existing teaching practices. This is the cultural context of teaching English in China. The results from the questionnaires and interviews have shown that EFL teachers’ beliefs are complicated. On the one hand, most teachers showed a positive attitude toward many new concepts in the NCS, such as student-centered teaching, and teachers’ new roles. They were supportive of these concepts not only because these are encouraged by the policymakers and school head teachers, but also because these are beneficial to students’ all-round development and communicative ability development. On the other hand, the teachers also hold very traditional beliefs and in practice, tended to teach students in a traditional way, for example, their teaching was teacher-centered, teacher-dominated and textbook-based; their teaching strongly emphasized grammar and vocabulary; teachers focused mainly on examination performance; they tended to show their authority and emphasized the importance of controlling the class. These traditional beliefs have been valued

in the Chinese educational context for a very long time (Cortazzi and Jin, 1996; Rao, 1998; Run-Hua, 2006), but other beliefs are affected and constrained by local teaching conditions, such as inadequate teaching resources, students' low English proficiency, the pressure of knowledge-based assessment (Vähäsantanen, 2015). Holding both new and traditional beliefs simultaneously suggests that some teachers seemed to be flexible, practical and were able to embrace the changes in the NCS and fitted them to their own teaching context so they can choose and create a suitable way to teach English that puts their students in an appropriate place between the local teaching reality and the NCS' requirements. Others may embrace some of the NCS ideals at a conceptual level, but not see them as practicable, while others did not see any changes, apart from the new textbooks, as being either necessary or desirable.

This study has also shown how great the gap between teachers' beliefs and practices can be. Despite teachers' seemingly strong beliefs about the importance of some of the key changes in the NCS (such as student-centered teaching, differentiated learning, reflective teaching), there can be great differences between beliefs and teaching practices of teachers. We have explored the training background of the teachers, but it is unclear how teachers' beliefs were shaped by their training and, given that many of the teachers shared beliefs in this study, it may be that the theoretical content of most training courses about EFL teaching and learning may be somewhat similar. This was outside the scope of this study, as was

the duration, method and intensity of such training, but would be an interesting extension to this study. Curriculum change is complex and it is not simply a matter of changing practices, or even of just changing beliefs.

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.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Shanghai Normal University and University of Nottingham. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

REFERENCES

- Basturkmen, H., Loewen, S., and Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Appl. Linguist.* 25, 243–272. doi: 10.1093/applin/25.2.243
- Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System* 40, 282–295.
- Beijing Normal University (2013). *Female Dominance in English Teaching in China*. Available online at: http://news.xinhuanet.com/english/culture/2013-01/27/c_132131045.html (accessed August 2020).
- Biggs, J. B. (1996). "Western misconceptions of the Confucian-heritage learning culture," in *The Chinese Learner: Cultural, Psychological and Contextual Influences*, eds D. A. Watkins and J. B. Biggs (New Delhi: CERC), 45–67.
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Breen, M. P., Hird, B., Milton, M., Oliver, R., and Thwaita, A. (2001). Making sense of language teaching: teachers' principles and classroom practices. *Appl. Linguist.* 22, 470–501. doi: 10.3389/fpsyg.2021.755043 doi: 10.1093/applin/22.4.470
- Brown, C. A., and Cooney, T. J. (1982). Research on teacher education: a philosophical orientation. *J. Res. Dev. Educ.* 15, 13–18.
- Burns, A. (1992). Teacher beliefs and their influence on classroom practice. *Prospect* 7, 56–66.
- Carless, D. R. (1998). A case study of curriculum implementation in Hong Kong. *System* 26, 353–368. doi: 10.1016/S0346-251X(98)00023-2
- Cheng, B. (2017). A comparative study of the liberal arts tradition and Confucian tradition in education. *Asia Pac. Educ. Rev.* 18, 465–474. doi: 10.1136/medethics-2014-102318 doi: 10.1007/s12564-017-9505-6
- Cheng, B., and Zhang, D. (2020). Cultivating citizens with confucian cosmopolitanism: defining the purpose of liberal arts education in the asian context. *Front. Educ. China* 15, 564–587. doi: 10.1007/s11516-020-0027-3
- Cortazzi, M., and Jin, L. (1996). "Culture of learning: language classrooms in China," in *Society and the Language Classroom*, ed. H. Coleman (Cambridge, MA: CUP), 169–206.
- Cottine, C. (2016). Role modeling in an early confucian context. *J. Value Inq.* 50, 797–819. doi: 10.1007/s10790-016-9576-3
- Deng, J., and Smith, C. A. (2018). The rise of new confucianism and the return of spirituality to politics in mainland China. *China Inform.* 32, 1–21. doi: 10.1177/0920203X18764041
- Farrell, T. S., and Lim, P. C. P. (2005). Conceptions of grammar teaching: a case study of teachers' beliefs and classroom practices. *TESL EJ* 9, 1–13.
- Finch, J. (1987). The vignette technique in survey research. *Sociology* 21, 105–114. doi: 10.1177/0038038587021001008
- Fullan, M. (2001). *Leadership in a Culture of Change*. San Francisco, CA: Jossey-Bass.
- Fullan, M. (2007). *Leading in a Culture of Change*. Hoboken, NJ: John Wiley & Sons.
- Fullan, M. (2012). *Change Forces: Probing the Depths of Educational Reform*. London: Routledge. doi: 10.4324/9780203059005
- Fullan, M., and Pomfret, A. (1977). Research on curriculum and instruction implementation. *Rev. Educ. Res.* 47, 335–397. doi: 10.3102/00346543047002335
- Graden, E. C. (1996). How language teachers' beliefs about reading instruction are mediated by their beliefs about students. *For. Lang. Ann.* 29, 387–395. doi: 10.1111/j.1944-9720.1996.tb01250.x
- Greenier, V., Derakhshan, A., and Fathi, J. (2021). Emotion regulation and psychological well-being in teacher work engagement: a case of British and Iranian English language teachers. *System* 97:102446. doi: 10.1016/j.system.2020.102446
- Gu, L. Y., and Wang, J. (2003). Teachers grow in educational action. *Curric. Teach. Mater. Method* 23, 9–15.
- Gu, M. (2009). College English learners' discursive motivation construction in China. *System* 37, 300–312. doi: 10.1016/j.system.2008.11.006
- Hu, G. (2002). Potential cultural resistance to pedagogical imports: the case of communicative language teaching in China. *Lang. Culture Curric.* 15, 93–105. doi: 10.1080/07908310208666636
- Hu, G. (2005). Contextual influences on instructional practices: a Chinese case for an ecological approach to ELT. *Tesol Q.* 39, 635–660. doi: 10.2307/3588525
- Huberman, M. (1981). *Exemplary Center for Reading Instruction (ECRI), Masepa, North Plains: A Case Study*. Andover, MA: The Network.

- Hughes, R., and Hubby, M. (2012). The construction and interpretation of vignettes in social research. *Soc. Work Soc. Sci. Rev.* 11, 36–51. doi: 10.1921/17466105.11.1.36
- Jiang, A. L. (2022). Identity work as ethical self-formation: the case of two Chinese EFL teachers in the context of curriculum reform. *Front. Psychol.* 12:774759. doi: 10.3389/fpsyg.2021.774759
- Jiang, A. L., and Zhang, L. J. (2021). Teacher learning as identity change: the case of EFL teachers in the context of curriculum reform. *TESOL Q.* 55, 271–284. doi: 10.1002/tesq.3017
- Jin, L., and Cortazzi, M. (2006). Changing practices in Chinese cultures of learning. *Lang. Cult. Curric.* 19, 5–20. doi: 10.1080/07908310608668751
- Kennedy, C. (1988). Evaluation of the management of change in ELT projects. *Appl. Linguist.* 9, 329–342. doi: 10.1093/applin/9.4.329
- Kırkgöz, Y. (2008). A case study of teachers' implementation of curriculum innovation in English language teaching in Turkish primary education. *Teach. Teach. Educ.* 24, 1859–1875. doi: 10.1016/j.tate.2008.02.007
- Lai, M. H. (2010). Teacher development under curriculum reform: a case study of a secondary school in mainland China. *Int. Rev. Educ.* 56, 613–631. doi: 10.1007/s11159-010-9181-9
- Lei, M. (2020). "The changing role of English teachers in China as a result of curriculum change," in *Education and Mobilities*, eds X. Zhu, J. Li, M. Li, Q. Liu, and H. Starkey (Singapore: Springer), 167–186.
- Lei, M. (2022). *Teachers' Perceptions of Curriculum Innovation in English Language Teaching in China*. Wuhan: Huazhong University of Science and Technology Press.
- Lei, M., and Medwell, J. (2020). How do English language teachers understand the idea of professional development in the recent curriculum reforms in China? *Asia Pac. J. Educ.* 40, 401–417. doi: 10.1080/02188791.2020.1717440
- Lei, P., and Xu, J. (2022). Development of college English teachers' psychological empowerment scale: a validation study. *Front. Psychol.* 815:846081. doi: 10.3389/fpsyg.2022.846081
- Li, L., and Wegerif, R. (2014). What does it mean to teach thinking in China? Challenging and developing notions of 'Confucian education'. *Think. Skills Creat.* 11, 22–32. doi: 10.1016/j.tsc.2013.09.003
- Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry*, Vol. 75. Thousand Oaks, CA: Sage.
- Maykut, P., and Morehouse, R. (2002). *Beginning Qualitative Research: A Philosophical and Practical Guide*. Abingdon: Routledge. doi: 10.4324/9780203485781
- Medwell, J., Poulson, L., and Wray, D. (1999). *Effective Teachers of Literacy: A Report of Research Project Commissioned by the TTA*. Exeter: School of Education, University of Exeter.
- Mei, D. M. (2019). *Guidelines for Teachers of English Based on the Curriculum Standards for Teaching English for General Senior High Schools*. Shanghai: Shanghai Educational Publishing House.
- Mei, D. M., and Wang, Q. (2018). *Interpretation of English Curriculum Standards for General Senior High Schools*. Beijing: Higher Education Press.
- Ministry of Education [MOE] (2018). *Putong Gaozhong Yingyu Kechen Biaozhun [Curriculum Standards for teaching English in General Senior High school]*. Beijing: People's Education Press.
- Munn, P., and Drever, E. (2004). *Using Questionnaires in Small Scale Research. Beginners' Guide*. Edinburgh: SCRE.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62, 307–332. doi: 10.3102/00346543062003307
- Ping, W. (2010). A case study of an in-class silent postgraduate Chinese student in London Metropolitan University: a journey of learning. *TESOL J.* 2, 207–214.
- Rao, Z. H. (1998). Confucius's education thought and foreign language teaching in China. *J. Xi'an For. Lang. Univ.* 6, 48–53.
- Run-Hua, X. U. (2006). The teaching theory by confucius. *J. Zhongzhou Univ.* 23:4.
- Tao, J., and Gao, X. (2017). Teacher agency and identity commitment in curricular reform. *Teach. Teach. Educ.* 63, 346–355. doi: 10.1016/j.tate.2017.01.010
- Vähäsantanen, K. (2015). Professional agency in the stream of change: understanding educational change and teachers' professional identities. *Teach. Teach. Educ.* 47, 1–12. doi: 10.1016/j.tate.2014.11.006
- Vähäsantanen, K., and Eteläpelto, A. (2009). Vocational teachers in the face of a major educational reform: individual ways of negotiating professional identities. *J. Educ. Work* 22, 15–33. doi: 10.1080/13639080802709620
- Yang, H. (2015). *Teacher Mediated Agency in Educational Reform in China*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-15925-6
- Yim, L. W. (1993). *Relating Teachers' Perceptions of the Place of Grammar to their Teaching Practices*. Unpublished master's thesis. Singapore: National University of Singapore.
- Yu, D. (2011). *Confucius from the Heart*. Beijing: Zhonghua Publishing House.
- Zheng, H. (2013). Teachers' beliefs and practices: a dynamic and complex relationship. *Asia Pac. J. Teach. Educ.* 41, 331–343. doi: 10.1080/1359866X.2013.809051
- Zhu, Y. (2018). *Language Curriculum Innovation in a Chinese Secondary School*. Singapore: Springer. doi: 10.1007/978-981-10-7239-0

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 Lei and Medwell. 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.

APPENDIX: VIGNETTE

Look at the example case talking about what Mr. Lin does in his daily English teaching, how do you feel about it? Please justify your answer.

The Vignette

Mr. Lin leads his class in an animated way. At the beginning of the class, he asks students some questions that based on the reading task they have done before and his students can give their responses quickly. After the review, he teaches the class new content, he always raises some questions to keep students attentive and listening to what he said.

He sees his role as an initiator, an explainer and a class controller. He thinks students won't learn English unless the teacher goes over the material in a structured way. He believes it is his duty to teach, to explain, and to show students how to learn English and how to do the task.

He says, "it is more practical to set the same teaching objectives for the whole class. Interactive activities such as group work should not take too much time in class because passing the exams is the final teaching goal and the most important thing for English teaching at school."

- 1 What do you think about the teaching approach Mr. Lin adopted? Is there anything else you might do for good teaching?
- 2 What do you think about the role Mr. Lin played in English class? Is there another role you think necessary for good teaching?
- 3 What do you think about Mr. Lin's viewpoint? Why?



English Language Teacher Agency in Response to Curriculum Reform in China: An Ecological Approach

Lian Wang*

School of Foreign Languages, Guizhou University of Finance and Economics, Guiyang, China

This study draws on the ecological perspective of teacher agency to examine the manifestation of English teachers' agency toward the ongoing curriculum reform in China and the factors that impact it. This study surveyed 353 high school English teachers and then collected data from three case study participants through in-depth interviews. The findings showed that the majority of teachers surveyed exhibited positive attitudes and beliefs about implementing the reform and inclinations to change, but the teachers also showed a constrained state of agency in practice. Teacher agency developed as the teachers exerted sustained pedagogical change and reflection on reform-based practices. Through the findings, prior experiences and reform-oriented beliefs were found to mediate teachers' agency, and reform-related experiences were more influential than future goals in shaping agency. The factors of perceived school culture that involved teachers' interaction with students, colleague cooperation, and administrative support also mediated teachers' agency in practice. Implications are proposed for policymakers and school leaders to help teachers coordinate inconsistencies between high-stakes examination preparation and holistic education and make positive sense of professional development in the context of educational changes.

Keywords: ecological approach, teachers' agency, curriculum reform, educational changes, teaching practices

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the Americas,
United States

Reviewed by:

Hongzhi Yang,
The University of Sydney, Australia
Honggang Liu,
Northeast Normal University, China

*Correspondence:

Lian Wang
200301039@mailgufe.edu.cn

Specialty section:

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

Received: 03 May 2022

Accepted: 16 June 2022

Published: 18 July 2022

Citation:

Wang L (2022) English Language
Teacher Agency in Response to
Curriculum Reform in China: An
Ecological Approach.
Front. Psychol. 13:935038.
doi: 10.3389/fpsyg.2022.935038

INTRODUCTION

Teacher agency is imperative for the process of implementing curriculum reforms and educational policies (Lasky, 2005; Pyhältö et al., 2012; Hamid and Nguyen, 2016; Tao and Gao, 2017). Current empirical studies show that teacher agency is a temporal and situated achievement and that teachers exercise their agency to respond to educational change in different manifestations, such as compliance, resistance, and negotiation (Robinson, 2012; Priestley et al., 2015; Yang and Clarke, 2018; Le et al., 2020). These studies constitute an area of inquiry of growing importance, which is known as teacher agency research in policy implementation, and they reveal that personal factors (e.g., teachers' beliefs) and contextual factors (e.g., teaching contexts) contribute significantly to teachers' change and growth during curriculum reform. Nevertheless, notably, teachers' agency development during educational changes is the outcome of the interplay within their professional life experiences and the expectations for future and contextual conditions.

A number of studies reveal that teachers are more likely to resist imposed policy mandates and embrace conventional teaching techniques to avoid taking risks in educational reform if the mandates are incongruous with their personal factors, such as their beliefs, intentions, and prior experiences (Biesta et al., 2015; Bonner et al., 2019). Le et al. (2020),

for instance, emphasize that English teachers are more likely to exercise their agency to struggle and resist educational language policy when there is a conflict between policy mandates and their beliefs, prior knowledge, and expectations. Teachers' agentic actions toward educational change are always mediated by the sociocultural contexts in which they are situated. Sociocultural contexts are reflected not only in classroom teaching, school conditions, and local communities but also in educational policy mandates, the promotion of ideological discourses, and changes in assessment practice that can either enable or hinder teachers' agency enactment (Lasky, 2005; Priestley et al., 2015; Poulton, 2020; Tao and Gao, 2021). For example, Liyanage et al. (2015) identified the struggles and dilemmas experienced by English language teachers in their attempts to exercise agency amid the instructional demands of the exam-oriented community in Inner Mongolia in China. Since teacher agency's interaction with other personal and contextual factors does not operate alone, a focus on teacher agency during curriculum reform should not distract researchers from profoundly examining the integration of their historical experiences, present conditions, and future aspirations that constitute an ecological perspective to explain how teachers make agentic choices and actions during educational changes. In China, the Ministry of Education (MOE) enacted the 2017 Edition of the General High School English Curriculum Standards (课程标准, also known as *xin kebiao*) in 2018, which was initiated to cultivate learners' four key competencies, namely language competence, cultural awareness, thinking capacity, and learning capacity, as the core competencies (核心素养, *hexin suyang*) in basic English as a Foreign Language (EFL) education. Therefore, this study goes inside the classroom teaching in mainland China and investigates how Chinese English teachers' agency in response to the prescribed national curriculum may be interpreted, and it takes an ecological approach with reference to the factors that enable or constrain teacher agency during educational reforms.

English is the foreign language studied by more than 99% of students who participate in formal education in China (Wang, 2007). Although English language teaching (ELT) and assessment in China has experienced a considerable transition in recent decades, ELT has continued to rely on a tradition of centrally administered exams to determine achievement and opportunity, and constraints continue to operate as major structural elements of the English teaching environment in China (Liyanage et al., 2015). In combatting the chronic malady of "only scores" and "only further education," the advent of *xin kebiao* is considered a driving force for a new round of English curriculum reform in China, and the new syllabus gives teachers guiding principles and pedagogical approaches, such as the activity teaching method, to teaching in action, which challenges teachers to take more responsibilities in innovating professional skills and nurturing quality education. Meanwhile, notably, the implementation of educational policies in China follows a traditional centralized and top-down manner in which prescriptive policies are initiated and imposed at the macrolevel (i.e., the MOE). Although a few studies have explored Chinese university teachers' agency and their professional development in educational reforms (Tao and Gao, 2017; Yang and Clarke, 2018; Tao et al., 2020), more research

is needed to examine how English teachers from different high schools teach in accordance with the core competencies (*hexin suyang*) in response to the prescribed national curriculum and the extent to which they might exert positive agentic changes in the context of China.

TEACHER AGENCY IN CURRICULUM REFORM

Research has noted that there is often a tension in educational policies about the amount of control that teachers are allowed to exert over the curriculum and what and how they teach (Hamid and Nguyen, 2016; Poulton, 2020). That is, implementing innovation with an educational reform is not a matter of straightforwardly executing policies; rather, it involves a process of sense-making through which teachers make meaning from their work environment (Vähäsantanen, 2015). Regarding English language teachers, a few studies have explored teacher agency during educational changes (Hamid and Nguyen, 2016; Tao and Gao, 2017; Le et al., 2020). Among them, Nguyen and Bui (2016) emphasize that teachers do not passively follow a set of norms mandated by policymakers but act as agents in shaping and reshaping a language policy through their pedagogical practices. It has been noted that a positive association was found between teachers' agency and their engagement with research and reflection on teaching in the context of the Chinese national college English reform (Yang and Clarke, 2018). The studies of Li et al. (2020) showed that English primary teachers in Vietnam attempted to adapt the new language policy mandates according to their interpretation, preferences, choices, and current teaching conditions. One noticeable contribution of these empirical studies is that English teachers can act as agents for curriculum reform, in which the dynamic development of teachers' professional practices can be partly explored through their different agency commitments. As agency in educational changes is practiced as a mediational tool that is crucial for teachers to resolve structural conflicts among personal beliefs, language policies, and hierarchical power relationships (Liyanage et al., 2015; Yang and Clarke, 2018), the research on English language teacher agency is, despite the growing interest in recent years, still modest in the current language policy literature.

DEFINING TEACHER AGENCY

The notion of agency has long been researched from quite different theoretical perspectives. There is a lack of consensus on the conceptualization of agency. Specifically, agency has been viewed as individuals' intentional acts to make things happen and participate in their development, adaptation, and self-renewal with changing times (Bandura, 2001). This concept emphasizes the psychological mechanism of the self-system in one's agency formation, while some scholars have argued that agency is a socioculturally mediated capacities to act (Ahearn, 2001; Lasky, 2005; Kayi-Aydar, 2019). Goller and Harteis (2017) derived a definition of agency on both a psychological and practical basis suggesting that human agency is the capacity and tendency

to make intentional choices, initiate actions based on these choices, and exercise control over the self and the environment. More recent attempts have been made to define agency in an ecological term as “a matter of personal capacity to act, combined with the contingencies of the environment within which such action occurs” (Priestley et al., 2012). The ecological perspective of agency argues the notions that agency is the capacity and tendency to make intentional choices, initiate actions based on these choices, and exercise control over the self and the environment (Eteläpelto et al., 2013). These notions suggest that even if actors have some type of capacity, whether they can achieve agency depends on the interaction of the capacities and the ecological conditions. Actors always act through an environment rather than simply in an environment (Biesta and Tedder, 2007; Priestley et al., 2015).

Agency has also been researched in the field of teacher education. For teachers, it is largely about repertoires for maneuvering or the possibilities for different forms of action available to them at particular points in time (Priestley et al., 2012). Therefore, Priestley et al. (2015) proposed an ecological perspective to acknowledge that teacher agency, as a temporal situated achievement, contains the three dimensions: iterational, practical-evaluative, and projective (see **Figure 1**). Language teachers, as active agents in social and educational contexts, directly and indirectly, experience their professional qualities as being naturally shaped by the environment in which they live (Chu et al., 2021). In this regard, teacher agency in response to curriculum reform may be afforded or constrained as the outcome of the interplay among an individual's past experiences, present conditions, and future goals.

The ecological perspective has been considered an effective approach for understanding the phenomenon of one's agency not only in context but also in one's life history (Tao and Gao, 2021). This model offers a more explicit ecological approach to teacher agency that constitutes both a methodological and a theoretical framework for empirical inquiry that relates to the approaches by which teachers achieve agency in practice. Agency is not simply concerned with the ways in which we engage with our contexts-for-action but rather with the capacity to shape our responsiveness to the situations that we encounter in our lives (Biesta and Tedder, 2007). The ecological approach provides useful insight into how teacher agency can be located through a consideration of their histories and beliefs, their abilities to visualize the alternative future, and their interplay with the sociostructural and material conditions in which individuals act (Priestley et al., 2015). Since the ecological perspective allows a historical perspective, Tao and Gao (2017) examined language teachers' professional trajectories in which their agency was situated by life-history exploration. Framed within an ecological conceptualization of teachers' agency, Poulton (2020) explored Australian primary teachers' reported experiences of agency and identified potential enablers and constraints to teachers' agency in curriculum planning and teaching; they found that strong beliefs, teacher knowledge, and skill and aspirations for school-based assessment helped teachers report greater experiences of agency. Liu's et al. (2022) study highlighted that English teachers, as active agents situated in the educational context, could positively

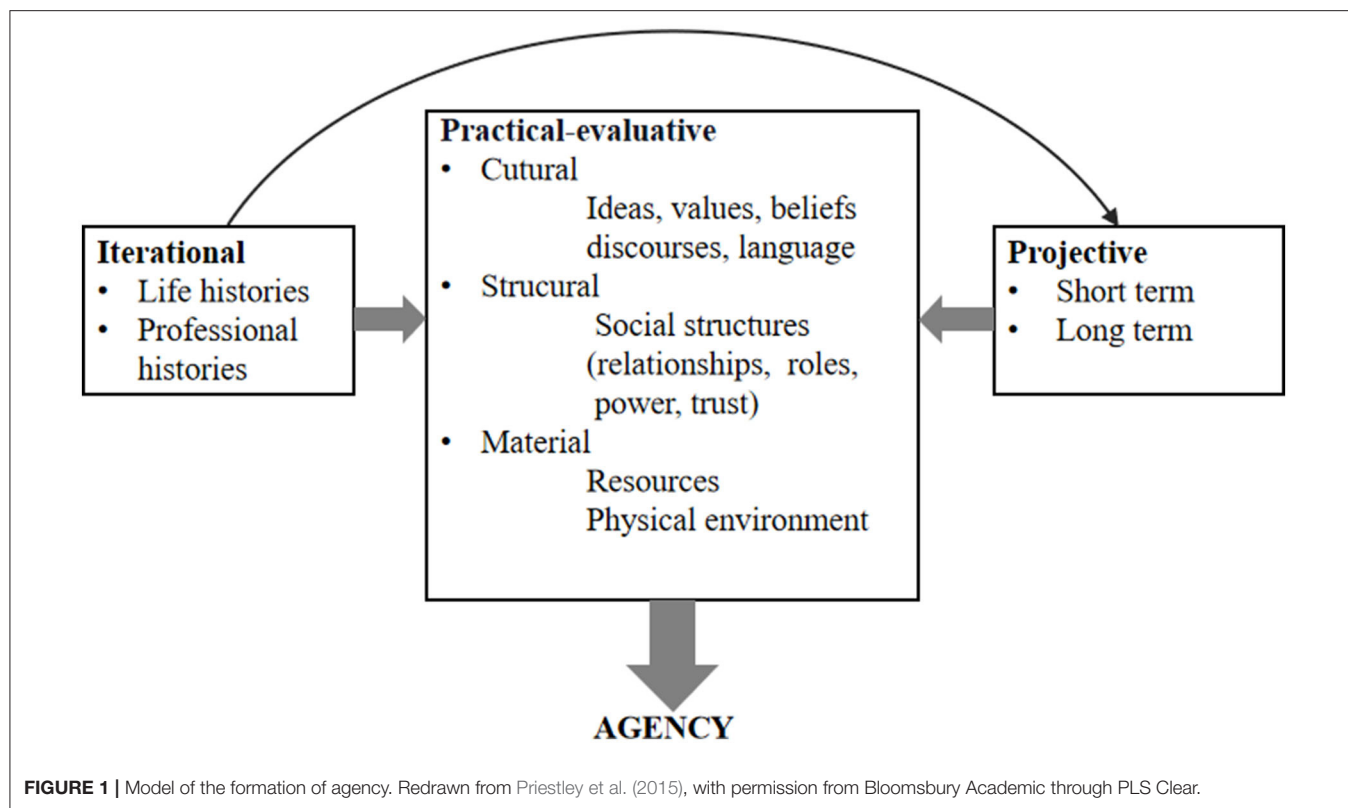
negotiate with their surroundings, and exert their agency to cope with professional and personal pressures. To conclude, the exploration of teacher agency has gained much empirical support and demonstrated diverse theme, while EFL teacher agency in response to curriculum reform has not been thoroughly examined from both psychological and ecological perspectives. In this context, mainly drawn on the ecological model of teacher agency proposed by Priestley et al. (2015), this study aims to adopt a mixed-method design to enrich the understanding of EFL teacher agency in response to curriculum reform in China by addressing the following questions:

- (1) In what way do teachers exercise agency in response to top-down curriculum reform?
- (2) What are the factors that enable and constrain teacher agency during curriculum reform?
- (3) How do these factors interact with teacher agency to facilitate the sustained implementation of the reform?

METHODS

The study employed a two-phased, sequential mixed methods explanatory design framework (Cresswell and Plano Clark, 2007). Precisely, the method used for the initial phase of the study was a survey of Chinese high school English teachers' agentic performance during the reform. For this purpose, a questionnaire has been constructed based on Goller and Harteis (2017)'s perceptions of agency constitutions. The questionnaire formulated three dimensions of teacher agency: intention level, action level, and regulation level. Intention level refers to teachers' action plans and strategies for participating in the reform. Action level refers to teachers' self-initiated and goal-directed behaviors that aim to take control over themselves and the environment. Regulation level refers to teachers' self-motivation and self-reflectiveness in light of educational change. The questionnaire consists of questions, such as “Do you agree that the core competencies should be integrated into actual teaching practice?”, “Do you often consciously nurture students' thinking capacity in class activities?”, “Do you often exchange teaching experience with colleagues about the core competencies?”, “Do you often reflect on your teaching process and adjust actions in line with *xin kebiao*?”, etc. To investigate the factors that had an impact on teachers' agency, this study also developed an instrument to examine the influential factors on teacher agency from the three established dimensions of participants' demographic characteristics, their reform-oriented beliefs, and their perceived school culture. The questionnaire uses a Likert scale from 1 (totally disagree or have never done) to 5 (absolutely agree or have always done).

Overall, the established questionnaire has four parts that consist of participants' demographic variables (teaching age, gender, education level, professional title, and school type), items concerning teachers' agency ($n = 25$), items concerning teachers' reform-oriented beliefs ($n = 7$), and items concerning teachers' perceived school culture ($n = 13$). The survey items and instructions were revised and validated in a pilot test before delivering the final questionnaire by using SPSS 21.0 and AMOS



7.0 in the aspects of item analysis, validity, and reliability analysis. Item analysis was first performed to detect the discriminant validity of each item. After a series of adjustments, five items were deleted from the scale of teacher agency, and three items were deleted from the scale of teacher-perceived school culture. Since the factor structure drew on previous conceptualizations of agency (Goller and Harteis, 2017), a confirmatory factor analysis (CFA) was adopted to examine the goodness of fit of the items from the teachers' agency scale. The CFA showed that the load values of 5 items included in the intension variable ranged from 0.65 to 0.76 on this factor, the load values of 7 items in the action variable ranged from 0.61 to 0.73 on this factor, and the load values of 8 items in the regulation variable ranged from 0.61 to 0.80 on this factor. The reported degree of the fitting indices of the three variables was acceptable as $\chi^2/df = 2.488$, CFI = 0.927, PCFI = 0.815, RMSEA = 0.065. These results indicated that 20 items from the revised scale of teacher agency were well-supported. Regarding teacher-perceived school culture, both an exploratory factor analysis (EFA) and CFA were used to examine the factor structure, and a two-factor structure composed of interpersonal conditions and the school system was obtained to identify it. The KMO measure and Bartlett's test of sphericity were conducted to ensure that the data were eligible for factor analysis. The KMO index was 0.856, and Bartlett's test of sphericity (Approx. Chi-Square = 572.725, $df = 45$) was significant at a level of 0.000. The varimax method showed a high correlation between the items and the common factors, with all factor loadings >0.5 and commonality >0.4 . Moreover, a CFA

for each factor was conducted, and one item was removed from the teacher-perceived school culture dimension with a loading lower than 0.40. The CFA results indicated that 9 items of the adjusted scale of teacher-perceived school culture were much better than the initial scale ($\chi^2/df = 1.847$, CFI = 0.960, PCFI = 0.818, RMSEA = 0.049). After the CFAs, an internal consistency reliability analysis was used to assess each of the subscales. The alphas were 0.890 for teachers' agency, 0.889 for reform-oriented beliefs, and 0.892 for perceived school culture. The pilot results helped to revise the total items from 45 to 36 and established the final questionnaire for the participants ($n = 387$).

Based on convenience sampling strategies, the questionnaire was written in Chinese and distributed to teachers from different high schools in five provinces of China (Jiangsu, Hebei, Guizhou, Guangdong, and the Guangxi Zhuang Autonomous Region). A total of 353 questionnaires were collected, and the proportion of cleared samples was 91.2%, with 84 from Jiangsu, 77 from Hebei, 57 from Guizhou, 36 from Guangdong, and 99 from the Guangxi Zhuang Autonomous Region. In addition, the gender distribution of the convenience sample (male = 73, female = 280) was in accordance with the fact that male English teachers are largely outnumbered by female English teachers in China.

Within the second phase of the study, a multiple case study with three participants was adopted to explore how these potentially influential factors interact with teacher agency to facilitate the sustained implementation of this reform, which allowed an in-depth understanding of the participants' perspectives, beliefs, and action regarding educational changes.

TABLE 1 | Profile of the case study participants.

Characteristic	Participants		
	Zhang	Xu	Wang
Gender	Female	Male	Female
Age	48	38	29
Teaching age	26 years	16 years	7 years
Grades	Grade 2	Grade 2	Grade 3
Qualifications	Master	Bachelor	Bachelor
Professional titles	Senior	Senior	Intermediate

Three participants who were born in the 1970's, 1980's, and 1990's were all involved in the initial survey and voluntarily agreed to participate in the second phase of the study. They were from different schools, and their EFL teaching experience varied from 7 to 26 years, which demonstrated that their educational background, teaching proficiency, and professional development conditions were very different. **Table 1** describes the demographic characteristics of the three participants.

DATA ANALYSIS

To answer RQ1, the questionnaire data were analyzed quantitatively with SPSS 21.0 for descriptive analysis and variance analysis. To answer RQ2, a one-way analysis of variance (ANOVA) and regression were used to examine the correlations between teachers' agency and their background information, beliefs toward the reform, and perceived school culture. To answer RQ3, qualitative data were collected that involved the participants' reform-oriented beliefs, prior experiences, future goals, in-class teaching decisions, and actions related to curriculum reforms through semi-structured interviews. Interviews consisted of questions, such as "Have you ever taken a part in a national curriculum reform in the past, and can you describe its impacts on your teaching practice?", "Can you describe the teaching objectives in your classes at present?", etc. All these face-to-face interviews were conducted with each participant in their actual daily workplaces in an ongoing and cyclical process over 6 months. Drawn on the across-case analysis of qualitative study, categories and patterns were inductively generated (Miles and Huberman, 1994). Drawing on the model of the ecological approach to teacher agency, influential factors of the iterational, practical-evaluative and projective dimensions on their enactment of agency during this reform were identified in the initial reading and coding of the data. In particular, the three integrated elements of teachers' agency mentioned above guided the whole encoding work regarding the choices and actions of the pedagogical changes made by the participants in practice.

RESULTS

EFL Teachers' Agency in Curriculum Reform

The descriptive statistics for the teachers' agency scores based on the scale showed that the mean value of teachers' agency to

implement the reform is 4.063 in total. The results appear to suggest that most English teachers advocate for the mandates proposed in *xin kebiao*. Although this finding reveals the big picture of teachers' supportive perspectives of the reform, in discussing teacher agency, it is a question of not only how teachers position themselves in relation to reforms but also how they act in engaging with the reforms (Vähäsantanen, 2015). To have a close examination of teachers' agentic changes in accordance with policy-related demands, this part conducts a contrastive analysis of Wang's, Xu's and Zhang's agency.

Wang: Constrained Agency

In the interviews, Wang mentioned that her motivation to take on new pedagogical practice was not strong. She stated that compared with her colleagues in the English department, she was less capable of interpreting educational theories of the new syllabus into effective teaching practices. As she claimed, she did try to enact some of the innovative ideas and new teaching methods recommended in *xin kebiao*, but the interactions between the students and her turned out to be unsuccessful. Thus, she rested on traditional pedagogies by promoting word rehearsal, sentence translation, and grammatical drilling in class, which she believed to be helpful for the students in preparation for *GaoKao*, a high-stakes exam used to determine the selection to study at prestigious key universities, ordinary universities, colleges, or other higher education institutes (Fang and Warschauer, 2004). In Wang's case, she demonstrated a constrained state of agency for engagement in the reform. On the one hand, she struggled considerably with a sense of powerlessness for the frustrated adoption of new teaching methods in class; on the other hand, she made great efforts to fulfill her accountability for students' test-based performance to such an extent that her teaching practices were identical to her testing practice (Liyange et al., 2015), which is contrary to the recommended teaching methodology in policy. In addition, Wang said she often "searched for professional assistance from experienced colleagues in deeply interpreting the connotation of some educational theories." It may be inferred that Wang's minimal efforts to perform changes in practice could be explained by her resistance to taking risks at the cost of exam scores.

Extract 1

Wang: I designed some teaching activities guiding students to guess the meanings of words in contexts, but they couldn't answer my questions and were not interested in these activities, let alone fostering their thinking capacity. The educational ideas and goals recommended in *xin kebiao* are so good for students, but currently, the main goal in my classroom is to practice and memorize words and collocations. Thus, the meaning of texts can be understood.

Xu: Transformative Agency

Xu mentioned that given the competitive exam-based assessment system in the context of China, it was difficult to fully cultivate key competencies among students. Simultaneously, he felt obliged to improve his teaching practice and competence by changing beliefs, pedagogies, and practices in the classroom

and maintaining a balance between test goals and holistic educational goals. Xu pointed out in interviews that the new syllabus focused on language skills in *listening*, *speaking*, *reading*, *writing*, and *watching*. For *watching*, he actively brought varying “Thinking maps” to innovate pedagogical practices to thus shift class practices from a grammar-translation approach to a more interactive approach. This teaching method is consistent with the ideas advocated in the reform. It seemed that Xu demonstrated a transformative state of agency in response to the reform. Faced with conflicts and dilemmas between public exam-oriented mindsets and holistic education required by the reform, Xu enacted an ongoing transformation of teaching expertise exempt from routine work to achieve innovations. These reform-based experiences that accompanied reflection constructed his practical knowledge and laid a solid foundation for his class practices, teaching contests, and pedagogical research. Xu talked about his impact on the students: “I always imparted new ideas reflected in examination to the students; they were good at problem-solving and critical thinking skills while aligning with academic standards.” Notably, a well-interactive relationship with students helped Xu make sense of his agency to enact pedagogies that he believed to be valuable for student learning.

Extract 2

Xu: I acknowledge that scores are always of the primary importance for students, most of their time and energy are put into the preparation for exams due to the present examination-success-oriented context, and very limited space of education has been left for students’ all-around development. However, I always possess the capability to make decisions on my own, so I keep reflecting on how to infiltrate the core competencies recommended in *xin kebiao* into my class and concentrate more on developing students’ cultural awareness and critical thinking capacity.

Zhang: Progressive Agency

As the head of the English department, the master teacher of beginning teachers, and the leader of the teaching workshop at school, Zhang has always kept learning on advanced curricular theories and embracing educational changes to enhance her professional expertise. She indicated that students’ English learning and academic achievements could be greatly enhanced if innovative, flexible, and appropriate pedagogies were adopted in practice. She put considerable energy into learning and thinking critically about *xin kebiao* since the government enacted the policy. Specifically, she kept trying varied teaching approaches to explore effective practices of educational theories contained in *xin kebiao* with self-confidence. She spoke with evident enthusiasm about conducting reform-based research associated with pedagogical changes that innovated her practical teaching and achieved well in meeting her students’ needs in creative ways, through which her conviction and responsibilities as curriculum developer strengthened greatly and motivated her to become one of the first adopters of the reform. To shift her role as a facilitator instead of being an authoritative instructor in class, Zhang cultivated students’ capability of active learning by engaging them in self-assessment and peer assessment. Moreover, Zhang showed

progressive agency with leadership in her school community: “I felt obliged to facilitate young unexperienced teachers in school to develop.” Zhang enacted her agency to negotiate her identity in her classroom with her students and within the school with colleagues; then, she took control of what was happening around her in the reform.

Extract 3

Zhang: I was once sent to participate in a national reform-based teaching competition on behalf of Guangxi Zhuang Autonomous Region. At the beginning, I felt stressed because the new syllabus only sets very general goals for English learning and teaching. It’s my work to determine which teaching approaches, such as “task-based instruction,” “communicative language teaching,” “activity teaching method,” and “san yi san xiang method,” will be contextually appropriate for accomplishing my teaching objectives. Although preparing for the competition was laborious and painstaking, it was worth trying. I gained a lot during the process.

The manifestations of Wang’s, Xu’s, and Zhang’s agency toward this reform at the individual scale can symbolize certain typical scenes and cases during educational changes in China. To summarize, the three teachers all showed positive perceptions of *xin kebiao* and its ideas, but they also exhibited different agentic changes in practice. In contrast with Wang’s weak engagement with the reform, Xu and Zhang demonstrated critical shifts in response to the reform by constructing mutual support with students and sharing valuable resources with peer teachers.

Demographic Characteristics and Reform-Oriented Beliefs

A series of one-way ANOVA tests were performed to establish whether there were significant differences in the respondents’ demographic characteristics. The statistical analysis reported that no significant difference was found in divided age groups ($p = 0.273 > 0.05$) and layered educational groups ($p = 0.261 > 0.05$) on their agency values. However, the one-way ANOVA indicated that all groups of varied professional titles manifested significant differences ($F = 4.41$, $p = 0.005$). As seen in **Table 2**, the median of senior teachers’ agency ranked the highest ($m = 4.72$), followed by associate senior teachers ($m = 4.10$), intermediate teachers ($m = 4.09$), and primary teachers ($m = 3.99$). This finding suggests that teachers with higher professional titles exercised more agency to implement advocated pedagogies and pioneer changes in class. This difference might be related to their variable identity commitments. Most experienced teachers with high professional titles tend to incorporate diverse identities as teachers, researchers, administrators, and policy practitioners who are most likely to pioneer educational changes at school or on indigenous levels.

Furthermore, a correlation analysis and a regression analysis were conducted to examine how well the scores of quantitative participants’ agency were correlated with their reform-oriented beliefs. As shown in **Table 3**, there is a significant positive relationship between teachers’ reform-oriented beliefs and agency. The findings indicate that teachers’ beliefs significantly predict teachers’ agentic performance, with teacher agency

TABLE 2 | Teachers' agency differences among professional titles.

	Mean				F	Sig.
	Primary (n = 120)	Intermediate (n = 150)	Associate senior (n = 64)	Senior (n = 4)		
Teacher agency	3.99	4.09	4.10	4.72	4.410	0.005**
Intention	3.92	4.06	4.11	4.70	4.589	0.004**
Action	3.96	4.04	4.03	4.71	3.139	0.026*
Regulation	4.04	4.13	4.15	4.75	3.645	0.013**

The * and ** symbol indicates the values of $p < 0.05$ and $p < 0.01$ respectively.

TABLE 3 | Correlations of the variables.

Variables	Beta	t	p	R ²	F
Teacher agency	-	-	0.000**	0.412	244.949
Teachers' beliefs	0.642	15.651	0.000**		

** $p < 0.01$.

serving as the dependent variable and teachers' beliefs serving as the predictor, $R^2 = 0.412$, $p < 0.01$.

The interview data with the three case study participants further confirmed the close correlation among teachers' professional qualifications, reform-oriented beliefs, and agency presented in the survey data. That is, the qualitative data confirmed that the three teachers' past experiences and reform-oriented beliefs were closely linked to their agentic choices and actions, which was especially relevant to their sensitive perceptions of the current reform. In particular, the past experiences that the three teachers drew on when making a choice in accordance with policy mandates were often related to their perceptions, responsibilities, and participation in prior curriculum reforms. As the youngest teacher among them, Wang had obtained her Bachelor's degree as an English major at a prestigious foreign language university in China. She had not received teacher education prior to entering high school and had not experienced any previous national curriculum reforms. As shown in Extract 4, she has been struggling with the contradictions of students' passive learning, has clung to traditional teaching methods, and has been expected to adopt new approaches of pedagogy in the changing environment, which resulted in the incongruence between her reform-oriented beliefs and her actual behaviors in class. It may be inferred that superficial beliefs in curriculum reform and limited knowledge of educational theories are not sufficiently solid for teachers to make a professional shift beyond their immediate classroom experiential strategies.

Extract 4

Wang: Theoretical knowledge is certainly helpful for the quality of teaching, and the core competences were often heard when I attended academic lectures and peer colleagues' open classes. For me, however, I had little reference to the theories in my

daily work and drew on only some advanced theories when participating in teaching contests. One reason is that the students' test scores in my class are the worst compared with other classes' students. The reform requirements are too difficult for them to meet.

Compared with Wang, Xu and Zhang are experienced teachers and the backbones of their schools. They were involved in previous curriculum reforms and acquired a wide repertoire of experience to cope with puzzles raised by educational changes, which might account for their adaptation to the current reform and confidence in their professional knowledge and pedagogic skills. During interviews, Xu and Zhang both stated that they have already published some reform-based research papers; meanwhile, they expressed strong desires to conduct more research on the basis of their pedagogical beliefs, practical knowledge, and self-reflection. The interviews revealed that deep insights into curriculum reform and critical reflection on creative pedagogies underpinned teachers' stronger sense of agency to engage in changes.

Extract 5

Xu: Since I had learned some new ideas, for example, the new term "cultural awareness" would replace the old term "cultural character" in the upcoming curriculum reform when communicating with some policy-makers and administrators of local educational office, I was prepared for the current reform. In the past, English teachers did pay enough attention to foster learners' Chinese cultural confidence, which had been modified in *xin kebiao*. Respect and appreciation for excellent cultures have been advocated to strengthen learners' patriotism and pride in Chinese historical culture.

Extract 6

Zhang: In the past 10 years, I have been committing to a series of training projects, namely, "Excellent Teachers," organized by the Municipal Education Department, where I kept learning and researching on advanced pedagogies; thus, my professional expertise has been greatly enhanced. This reform exerts pressure on me, but its policies are great. It centers on fostering students' integral competences rather than only judging them by academic performance; teachers are not simply imparters of knowledge but also helpers to scaffold students to establish correct values.

TABLE 4 | Correlations and correlation of the variables.

	Perceived school culture	Interpersonal condition	School system		
Teacher agency	0.563**	0.528**	0.422**		
Variables	Beta	<i>t</i>	<i>p</i>	<i>R</i> ²	<i>F</i>
Teacher agency	-	-	0.000**	0.317	162.181
Perceived school culture	0.642	12.735	0.000**		

** $p < 0.01$.

The results above suggest that Xu had focused on developing learners' cultural awareness and had reflected considerably on this term's connotation beforehand. Teachers are most likely to act as reform agents when their prior beliefs are in congruence with reform principles (Bonner et al., 2019). Xu acknowledged that it was impossible to reflect all the core competencies advocated by *xin kebiao* in assessment, while teachers should adhere to holistic educational goals in the long term and take into account students' academic excellence simultaneously, which inspired him to integrate valuable ideas of *xin kebiao* and high-stakes testing goals into his class. The interview data with Zhang confirmed that she exerted great efforts to foster learners' critical thinking by promoting academic knowledge and social justice to ensure that learners were equipped with a lifelong ability to learn and a sense of social responsibility. In general, the analysis of the three teachers' past experiences and beliefs revealed that teachers' experiences in career trajectories and reform-oriented beliefs are important prerequisites for alternative evaluation, decision making, and acting in a particular professional situation.

Interaction With the Structural Environment

In this section, correlation analysis and regression analysis were conducted to examine how well the quantitative participants' perceived school culture correlated with teachers' agentic performance. A significant positive relationship was reported between perceived school culture and teacher agency, whose items of interpersonal condition were highly correlated with teachers' agentic performance. The findings shown in **Table 4** indicate that perceived school culture significantly predicts teachers' agency level, $R^2 = 0.317$, $p < 0.01$. These results might reveal that harmonious interactions with students and strong professional bonds with coworkers guarantee sustained agentic development for individual teachers.

Teachers' perceived school culture is another important factor in mediating teachers' agency during the reform. The qualitative analysis showed that the schools in which the three teachers worked had recontextualized policy mandates in different forms that had socially influenced their perceptions of their roles in the reform. Wang's data suggested that the school that she worked

in concentrated more on teachers' examination accountability for evaluating students' learning and teachers' teaching, mainly based on test scores. Extract 7 also revealed that little intellectual support and interactions at the school level reduced Wang's possibilities of professional cooperation with colleagues. This evidence might partly explain Wang's great emphasis on students' concrete language skills in exam preparation.

Extract 7

Wang: Since English teachers at my school are overburdened with loads of classes, we have no stable time and space for regular teaching and researching sessions and simply gather together when there is a research activity held by the local educational office. Most of the time, we communicate *via* QQ group chat (a Chinese social medium) and share some teaching materials and resources that we believe to be useful. I truly hope that the school administrators could offer us more opportunities to learn advanced knowledge and skills in diverse ways.

Xu described that lately he had engaged in interschool English teachers' professional development training in association with curriculum reform. The teacher community mentioned in Extract 8 was organized to empower teachers to acquire a profound understanding of *xin kebiao*, which helped Xu to self-perceive as an active subject in educational changes. The data revealed that the creation of a learning environment for teachers who are enabled access to new educational theories and pedagogies in professional communities may enhance their sense of agency, confidence, and optimism regarding the reform.

Extract 8

Xu: My school has established a training program for teachers' professional development hand-in-hand with the Basic Education Institute of Jiangsu Academy of Educational Sciences, in which some chosen English teachers across the province are regularly organized to receive trainings for a week or two. During the session, I had chances to ask questions, discuss doubts and seek advice from experts and instructors. At the beginning, we all had no idea about the new curriculum reform, but we harvested many new theories when a series of trainings were completed.

Zhang mentioned in interviews that she and other faculty members in the English department met on the same workday each week to work on teaching plans and learn official documents together. Extract 9 demonstrates Zhang's enthusiasm for writing extended school-level textbooks and enriching teaching materials in a subgroup. This may indicate that intellectual exchanges occur through professional cooperation in different social activities, which create an opportunity for individual teachers to actively engage in the reform. It seems that Zhang not only perceived herself as an active performer of the reform but also performed in a larger picture of facilitating inexperienced teachers to take on new pedagogical innovations.

Extract 9

Zhang: School authorities require all English staff to attend regular weekly meetings to solve routine issues of daily work. We usually share some valuable resources, such as teaching plans, teaching design and PPTs, in giving and generous attitudes.

To enrich students' cultural awareness, I set up a team in the department to build a school-level curriculum materials resource base, that aimed to work with other colleagues to intrigue learners' interests by containing more local cultures that they are familiar with in life.

Aspirations for Professional Development

The projective dimension of teacher agency concerns both their long- and short-term work aspirations (Priestley et al., 2015). The interview data with the three teachers demonstrated that their future professional goals were closely linked to their agentic choices and actions during the reform. Wang said that it was absolutely necessary to “enact some advanced teaching theories and skills in class. However, for the students, entrance examination preparation to college upon which their future would be determined are most important, I don't dare to implement innovative pedagogies when *Gaokao* is approaching.” Additionally, she has undergone a professional dilemma and aspired to promote her professional learning by “participating in a Master program in the future and learning more knowledge of EFL educational theories.” It seems that her aspiration for a Master's degree regarding long-term professional goals articulated her hope of controlling her professional development trajectory with autonomy, which was in conflict with her short-term goal of pursuing higher scores and meeting students' academic standards set by school authorities. Comparatively, Xu's long-term goals in work were to “facilitate his students to gain more holistic knowledge and improve their language proficiency,” and his short-term goals were to “expand practical knowledge and conduct more reform-based research.” Zhang's long-term professional goals were to “explore deeply in educational theories and reinforce her students' social responsibilities, including how they could gain capability of life-long learning and how they could gain independence of thought and moral sense of judgment.” Her short-term professional goals were to “enhance professional expertise and academic competence through self-directed learning, cooperative learning and researching.”

The interview data and the teachers' published papers revealed that they had different expectations to conduct research. For example, Xu utilized mind maps in practical teaching and published academic papers based on his practical knowledge. Thus, he gave a future orientation for conducting more action research on relevant themes. Similarly, Zhang brought up her research interests as incentives to take on pedagogical changes. Extract 10 describes her research experience and agentic actions in focusing on activity teaching methods recommended in *xin kebiao*.

Extract 10

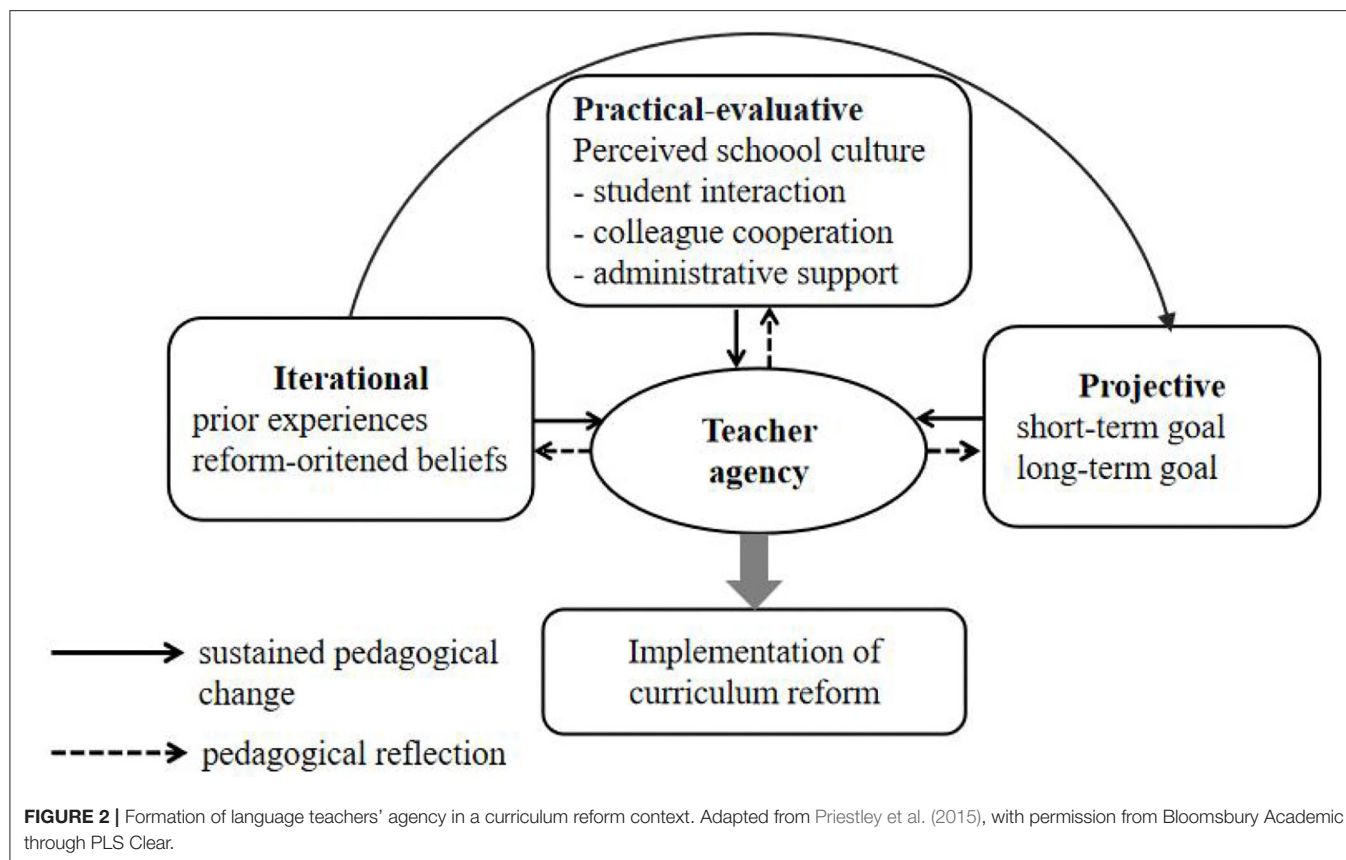
Zhang: Last summer, I attended a lecture given by an expert of curriculum reform on activity teaching methods, which raised my interests and inspired me to learn more about it. In the process of teaching practice guided by the method, I wrote an article and eventually published it, which gave me a great sense of achievement. I believe theoretical knowledge can instruct a teacher to supervise and evaluate his class, so I will continue to do research.

In all, Xu and Zhang were highly self-conscious that the class was inherently an “experimental field” where they enacted instrumental reform to thus promote their reflections on existing teaching practices. This evidence may indicate that teachers are most likely to act as reform agents when they can operate closely with positive interactions among their pedagogical beliefs, research interests, professional goals, and the reform policy.

DISCUSSION AND IMPLICATIONS

In this study, the researcher endeavored to examine the agency of English teachers and analyze the influencing factors from an ecological perspective. Accordingly, the majority of teachers surveyed exhibited positive attitudes and beliefs about implementing the reform and noticeable inclinations toward change. Regarding the demographic characteristics, the findings confirmed that the scores of senior teachers' agency were much higher than those of other cohorts with lower ranks of professional titles. These results reveal that the imposition of national top-down educational policy from the macrolevel (i.e., the MOE) has strong impact on teachers' potential changes in their beliefs, behaviors, and competence in classroom practice. The teachers' abundant professional experience and resources may allow them to actively engage in policy implementation as motivational forces in its infancy.

Although the agency of English teachers demonstrated a relatively high implementation of policy mandates, their agency to act on changes is at variance with and is mediated by the personal qualities that they bring to their work. When the prescriptions and policy mandates are enacted at the microlevel (the classroom context), the qualitative study reveals that English teachers' high agency in the reform manifested in a series of sustained pedagogical change behaviors and reform-based reflections. Although the three case study teachers' ideas and beliefs are in congruence with those of the curriculum reform, they exhibited different agentic actions in the teaching process and pedagogical shift. Three forms of agency are found in this study: constrained agency, transformative agency, and progressive agency. More specifically, Wang demonstrated the constrained state of agency with notable inconsistency between positive reform-oriented beliefs and reversed actions in classroom practice. The problems in Wang's powerlessness in change and weak participation in the reform are predictable on the basis of her inexperienced practical knowledge and unsupportive school environment. Contrary to Wang, Xu and Zhang displayed considerable reflections on educational change and thus demonstrated themselves as active agents in consciousness and actions. The example of Xu supports the statement of being transformative by showing that his realistic intention of preparing the students for examinations and his ideal intention of promoting holistic education are integrated into his efforts to overcome the contradiction to then enhance teaching quality. In the same way, Zhang's progressive agency in the reform was manifested by actively constructing the expected identity imposed by educational changes and facilitating colleagues' growth in teaching.



These findings draw attention to how particular ecological conditions interconnect and interact with individuals, as captured in **Figure 2**.

Figure 2 presents a complex model of understanding the formation of language teachers' agency as a continuum that contains the three dimensions of iterational, practical-evaluative, and projective. Regarding the iterational dimension, as shown in **Figure 2**, prior experiences and reform-oriented beliefs were found to mediate teacher agency, which allows one to sustain previous practice or change on the microlevel. The implications of this model on the interplay between teacher agency and the iterational dimension that impacts it are as follows. First, this study supports the statement that particular experiences in a teacher's biographical dimension provide a strong 'drive' toward the future and make a clear difference in the here and now (Priestley et al., 2015). This study extends this perspective by finding that teachers' agency to change may originate from the recognition of one's agentic capacity that relies on transferring prior reform-related experiences into the present reform. A prior successful experience is likely to help teachers maintain confidence in their teaching ability (Chu et al., 2021; Liu et al., 2021). These prior successful experiences seem to enable teachers to have a clear awareness of implicit developmental opportunities, possible resources, and behavioral principles accompanied by educational changes and to thus take initiative to make reform-oriented choices and practical actions to transform in line with the reform. These findings also emphasize that the accumulated experiences and thoughts

from the past are always intertwined with practical activities and are more significant than future professional goals in shaping teacher agency in curriculum reform in teachers' developmental trajectories. Second, reform-oriented beliefs are considered to be important factors to achieve agency in the iterational dimension. This is incongruous with the proposed pattern, which includes beliefs in the practical-evaluative dimension (Priestley et al., 2015) (see **Figure 1**), as the study found that teachers' reform-oriented beliefs that directed their actions were rooted largely in prior experiences and were more difficult to change than knowledge or practice during reform enactment. Although teacher beliefs always act in the present, belief change in teacher practice is difficult, slow, and often transient (Bonner et al., 2019). These findings also confirm that the agency to change is not activated by beliefs alone; a partial adoption of reform beliefs can also coexist with the agency of resistance toward the reform, but teaching and learning achievement gains for implementing reforms appear to be precursors of deep change in beliefs (Biesta et al., 2015; Naraian and Schlessinger, 2018; Bonner et al., 2019; Bao et al., 2020). For instance, the tension between Wang's stated beliefs and inconsistent behaviors in class was largely motivated by her perception of students' reactions and the safety of using innovative teaching methods in her classrooms. Since an insufficient experience with effective pedagogical strategies and practical knowledge is the primary cause of teachers' "willing spirit but weak power" phenomenon, it seems that policies prescribed in *xin kebiao* should be interpreted and implemented in more flexible and diverse senses rather than in narrow and

rigid ways, which might lead to inappropriateness to the realities of the students. Thus, the critical question becomes how reform-related teacher educational programs could enable teachers to generate agency in challenging changes and gaining autonomy in contextual conditions. Third, this study implies that systematic reflections on prior experiences, future goals, and actions are seen as indispensable parts of teacher agency development. The three teachers interviewed reported their experiences of transferring familiar valuable acting principles into the reform context, in which reflective thinking was intertwined with their decision-making and action-taking in professional practices. Thus, reflection on action is considered to be the dominant activity for expanding practical knowledge and supporting teacher learning in general (Le et al., 2020), especially proactive reflection in evaluating aim accomplishment after experiences lay the groundwork for teacher learning in educational change.

For the projective dimension, future professional short-term goals and long-term goals on the microlevel were found to have been applied when evaluating alternatives, making decisions, and acting in a particular situation (Leijen et al., 2020). The results from the study concur with early research (Priestley et al., 2015), which suggested that teachers' short-term goals with respect to their day-to-day decision-making and teaching action are shaped by a perceived need to deliver enjoyable lessons and keep students engaged. For instance, examination preparation was prioritized when Wang exerted pedagogical practices in routine school work. Regarding long-term professional goals, the three involved teachers set purposes for improving teaching competence in different ways, which demonstrates that teachers are conscious of the imperative shifting of educational context in China and the need to transform their professional identity. Wang expressed aspirations to make progress in career development and be promoted as a senior teacher. Xu and Zhang mentioned digging deeper into the domain of the core competencies and developing wider social purposes of education. It is evident in these findings that when teachers' long-term professional goals are more congruent with ideas proposed by the policy, they will be more engaged in the reform. This suggests that educational innovation aligned with teachers' professional interests and expectations for the future can sustain and inspire their dedication to the whole process of curriculum reform. Therefore, in the early phase of curriculum reform, opportunities, social dialog, and time should be offered to teachers to become aware of their transformed roles and make sense of their views on the reforms (Vähäsantanen, 2015). It is also essential for policymakers and school administrators to create more opportunities for teachers to make a positive sense of professional development in the context of educational change.

Concerning the practical-evaluative dimension, a closer examination of how the variables of perceived school culture mediated the exercise of teacher agency on the meso level yielded additional evidence with implications for practice. The challenge of localizing the national curriculum should not be underestimated (Kennedy, 2013; Le et al., 2020). The results indicate that teachers' relationships with students, colleague cooperation, and administrative support strongly mediated teachers' practices during the whole reform period. Wang's previous intention to change was discouraged by students'

inactive responses in class and insufficient support from the external environment. In contrast, Xu enhanced his sense of agency by participating in teacher communities organized by the school, and Zhang had access to more socially supportive interactions with the students and colleagues to develop relevant educational expertise. Therefore, the study confirms that whether school authorities attach importance to the reform by organizing training programs at the school level and interschool level and providing pedagogical resources for teachers affects their agency to a different degree. Given the deep-rooted tradition of examination-driven education in the Chinese context, localized school institutions always implement macrolevel curriculum reform as policy translators by changing the existing curriculum to various degrees based on their diverse financial conditions, deep-rooted culture, and hierarchical context. Notably, schools at the meso level interact with national educational policy, social ideologies, students' anticipations, and teachers' agency in a very complex way. Highly hierarchical school systems usually have close ties with less individual autonomy, which might be seen as a key competency of agency (Tao and Gao, 2021). Wang's constrained state of agency was largely accountable to her school context, where the regime of examination accountability prevailed and was dominant, while inadequate learning opportunities and professional help were offered. The finding echoes other published empirical results (Liu et al., 2016; Yang and Clarke, 2018; Poulton, 2020). Change requires both reculturing and restructuring of schooling (Biesta et al., 2015). However, the influence of schooling does not operate in a linear way on teachers because boosted agency may gradually drive the decomposition of traditional school culture and rebuild new culture, and vice versa. For example, Zhang constructed a teacher-learning community and created mutual support among colleagues through cooperation, which empowered her as a leading practitioner in the reform. Thus, suggestions for school institutions to relieve teachers' anxiety about educational change and enhance their agentic change include nurturing the role of reform leaders among teachers and providing them good opportunities for school-guided collaborative interaction, learning, and research. This study also calls for teacher education programs that concentrate on teachers' assessment literacy and help them find approaches in more profound ways to coordinate the inconsistencies between high-stakes examination preparation and holistic education.

CONCLUSION

This study contributes to future research, as it demonstrates that English teachers' agency in response to the reform manifested in constrained, transformative, and progressive forms at the early phase of curriculum reform. In particular, the study found that teachers developed deep agentic changes through sustained pedagogical change and reflection on reform-based practices. Another important contribution of this study shows that the relationship between reform-oriented beliefs and teacher agency is more than linear. Teachers' agency to change is not activated or refrained by beliefs alone; other influential factors whose prior experiences, self-perceptions of transformative opportunities, accessible resources, and positive acting between pedagogical

reflections and academic research also profoundly influence teachers' contextual agentic choices and actions.

To challenge the constraints on agency, regular learning opportunities need to be offered at the school-based level to foster teachers' curriculum reform literacy and assessment literacy, enhance their agency to take on positive changes, coordinate inconsistencies between high-stakes examination preparation and holistic education, and produce more autonomy in class practices. It is vital that a variety of indigenous school-based reforms be implemented to echo the national reform, which supports teachers in creating a space for critical reflection and negotiating more favorable identities in their professional trajectories. It is acknowledged that teachers' agency dynamically develops during different periods of curriculum reform (Priestley et al., 2015). In the future, longitudinal-designed research is needed to explore language teacher agency and provide more explanatory power for understanding agency in educational changes. Moreover, there is a lack of a sufficient amount of class observation that tracks interviewed teachers' agency development trajectories within classrooms. It is also worth examining teachers' discourse and actions in responding to the reform and enabling the student voice to be heard at the classroom level.

AUTHOR'S NOTE

LW is an associate professor in the School of Foreign Language, Guizhou University of Finance and Economics,

Guizhou, China. LW interests include educational linguistics, language teacher development, and language education policy.

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

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by the Research Launch Project for Introducing Talents of Guizhou University of Finance and Economics (Grant No. 2020YJ029) and the Specialized Project of Key Discipline-Guizhou University of Finance and Economics (Grant No. 2020ZJXK08).

REFERENCES

- Ahearn, L. M. (2001). Language and agency. *Annu. Rev. Anthropol.* 30, 109–137. doi: 10.1146/annurev.anthro.30.1.109
- Bandura, A. (2001). Social cognitive theory: an agentic perspective. *Annu. Rev. Psychol.* 52, 1–26. doi: 10.1146/annurev.psych.52.1.1
- Bao, M., Ren, W., and Wang, D. (2020). Understanding the professional practice of teachers of Chinese as an additional language through the lens of teacher agency. *Sustainability* 12, 74–93. doi: 10.3390/su12187493
- Biesta, G., Priestley, M., and Robinson, S. (2015). The role of beliefs in teacher agency. *Teach. Teach.* 21, 624–640. doi: 10.1080/13540602.2015.1044325
- Biesta, G., and Tedder, M. (2007). Agency and learning in the life course: towards an ecological perspective. *Stud. Educ. Adults* 39, 132–149. doi: 10.1080/02660830.2007.11661545
- Bonner, S. M., Diehl, K., and Trachtman, R. (2019). Teacher belief and agency development in bringing change to scale. *J. Educ. Change* 21, 363–382. doi: 10.1007/s10833-019-09360-4
- Chu, W., Liu, H., and Fang, F. (2021). A tale of three excellent Chinese EFL teachers: unpacking teacher professional qualities for their sustainable career trajectories from an ecological perspective. *Sustainability* 13, 6721. doi: 10.3390/su13126721
- Cresswell, J. W., and Plano Clark, V. (2007). *Designing and Conducting Mixed Methods research*. London: Sage.
- Eteläpelto, A., Vähäsantanen, K., Hökkä, P., and Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educ. Res. Rev.* 10, 45–65. doi: 10.1016/j.edurev.2013.05.001
- Fang, X., and Warschauer, M. (2004). Technology and curricular reform in china: a case study. *TESOL Quarterly* 38, 301–323. doi: 10.2307/3588382
- Goller, M., and Harteis, C. (2017). "Human agency at work: towards a clarification and operationalisation of the concept," in *Agency at Work: An Agentic Perspective on Professional Learning and Development*, eds M. Goller and S. Paloniemi (Cham: Springer).
- Hamid, M. O., and Nguyen, H. T. M. (2016). Globalization, English language policy, and teacher agency: focus on Asia. *Int. Educ. J. Comp. Perspect.* 15, 26–44. Available online at: <http://openjournals.library.usyd.edu.au/index.php/IEJ/index>
- Kayi-Aydar, H. (2019). A language teacher's agency in the development of her professional identities: a narrative case study. *J. Lat. Educ.* 18, 4–18. doi: 10.1080/15348431.2017.1406360
- Kennedy, K. (2013). "Singapore's school curriculum for the future beyond: national development?" in *Globalisation and the Singapore Curriculum: From Policy to Classroom*, eds Z. Deng, S. Gopinathan, and C. K. Lee (Singapore, SN: Springer Publishing), 205–224.
- Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teach. Teach. Educ.* 21, 899–916. doi: 10.1016/j.tate.2005.06.003
- Le, M. D., Nguyen, H. T. M., and Burns, A. (2020). English primary teacher agency in implementing teaching methods in response to language policy reform: a vietnamese case study. *Curr. Issues Lang. Plan.* 22, 199–224. doi: 10.1080/14664208.2020.1741209
- Leijen, L., Pedaste, M., and Lepp, L. (2020). Teacher agency following the ecological model: how it is achieved and how it could be strengthened by different types of reflection. *Br. J. Educ. Stud.* 68, 295–310. doi: 10.1080/00071005.2019.1672855
- Liu, H., Chu, W., and Wang, Y. (2021). Unpacking EFL teacher self-efficacy in livestream teaching in the Chinese context. *Front. Psychol.* 12, 717129. doi: 10.3389/fpsyg.2021.717129
- Liu, H., Yan, C., and Fu, J. (2022). Exploring livestream English teaching anxiety in the Chinese context: An ecological perspective. *Teach. Teach. Educ.* 111, 103620. doi: 10.1016/j.tate.2021.103620

- Liu, S., Hallinger, P., and Feng, D., (2016). Supporting the professional learning of teachers in China: does principal leadership make a difference? *Teach. Teach. Educ.* 59, 79–91. doi: 10.1016/j.tate.2016.05.023
- Liyanage, I., Bartlett, B., Walker, T., and Guo, X. H. (2015). Assessment policies, curricular directives, and teacher agency: quandaries of EFL teachers in inner Mongolia. *Innov. Lang. Learn. Teach.* 9, 251–264. doi: 10.1080/17501229.2014.915846
- Miles, M. B., and Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks: Sage.
- Naraian, S., and Schlessinger, S. (2018). Becoming an inclusive educator: agentive maneuverings in collaboratively taught classrooms. *Teach. Teach. Educ.* 71, 179–189. doi: 10.1016/j.tate.2017.12.012
- Nguyen, H. T. M., and Bui, T. (2016). Teachers' agency and the enactment of educational reform in Vietnam. *Curr. Issues Lang. Plan.* 17, 88–105. doi: 10.1080/14664208.2016.1125664
- Poulton, P. (2020). Teacher agency in curriculum reform: the role of assessment in enabling and constraining primary teachers' agency. *Curric. Perspect.* 40, 35–40. doi: 10.1007/s41297-020-00100-w
- Priestley, M., Biesta, G., and Robinson, S. (2015). *Teacher Agency: An Ecological Approach*. London: Bloomsbury.
- Priestley, M., Edwards, R., Priestley, A., and Miller, K. (2012). Teacher agency in curriculum making: agents of change and spaces for manoeuvre. *Curric. Inq.* 42, 191–214. doi: 10.1111/j.1467-873X.2012.00588.x
- Pyhältö, K., Pietarinen, J., and Soini, T. (2012). Do comprehensive school teachers perceive themselves as active professional agents in school reforms? *J. Educ. Change* 13, 95–116. doi: 10.1007/s10833-011-9171-0
- Robinson, S. (2012). Constructing teacher agency in response to the constraints of education policy: adoption and adaptation. *Curric. J.* 23, 231–245. doi: 10.1080/09585176.2012.678702
- Tao, J., and Gao, X. S. (2017). Teacher agency and identity commitment in curricular reform. *Teach. Teach. Educ.* 63, 346–355. doi: 10.1016/j.tate.2017.01.010
- Tao, J., and Gao, X. S. (2021). *Language Teacher Agency*. London: Cambridge University Press.
- Tao, J., Zhao, K., and Chen, X. (2020). The collective agency of language teachers under the scheme of research excellence: using a social network approach. *Círculo de Lingüística Aplicada a la Comunicación*. 84, 13–25. doi: 10.5209/clac.71992
- Vähäsantanen, K. (2015). Professional agency in the stream of change: understanding educational change and teachers' professional identities. *Teach. Teach. Educ.* 47, 1–12. doi: 10.1016/j.tate.2014.11.006
- Wang, Q. (2007). "The national curriculum changes and their effects on english language teaching in the People's Republic of China," in *International Handbook of English Language Teaching*, eds J. Cummins and C. Davison (New York, NY: Springer), 87–105.
- Yang, H., and Clarke, M. (2018). Spaces of agency within contextual constraints: a case study of teacher's response to EFL reform in a Chinese university. *Asia Pacific J. Educ.* 38, 1–15. doi: 10.1080/02188791.2018.1460252

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.

Copyright © 2022 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.



OPEN ACCESS

EDITED BY

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

REVIEWED BY

Vanessa R. Simmering,
University of Kansas, United States
Naseem Al-Aidroos,
University of Guelph, Canada

*CORRESPONDENCE

María Quirós-Godoy
maria.quirosgodoy@uam.es

SPECIALTY SECTION

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

RECEIVED 29 March 2022

ACCEPTED 24 June 2022

PUBLISHED 22 July 2022

CITATION

Quirós-Godoy M, Gil-Gómez
de Liaño B and Perez-Hernandez E
(2022) I can look for it! Modulation of a
concurrent Visual Working Memory
task in Visual Search in development.
Front. Psychol. 13:907121.
doi: 10.3389/fpsyg.2022.907121

COPYRIGHT

© 2022 Quirós-Godoy, Gil-Gómez de
Liaño and Perez-Hernandez. 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.

I can look for it! Modulation of a concurrent Visual Working Memory task in Visual Search in development

María Quirós-Godoy^{1*}, Beatriz Gil-Gómez de Liaño¹ and
Elena Perez-Hernandez²

¹Department of Social Psychology and Methodology, Universidad Autónoma de Madrid, Madrid, Spain, ²Department of Development and Educational Psychology, Universidad Autónoma de Madrid, Madrid, Spain

Daily classroom activities that require children to perform visual search (VS) tasks are common across all educational levels: from searching for a missing piece of a puzzle in kindergarten to solving equations in college. However, VS tasks are often not performed in isolation, but rather students are maintaining information related to an ongoing task that loads working memory (WM). Unfortunately, it is still unclear how these processes interact and evolve in development. The present work aims to study how a concurrent visual WM (VWM) load can modulate VS performance based on the Developmental Model of Endogenous Mental Attention (Pascual-Leone and Johnson, 1999, 2005, 2021). A sample of kindergarten, elementary (2nd and 4th grades), middle school (6th grade), and college students looked for real-world photorealistic targets while maintaining similar objects in VWM in a dual-task paradigm. VWM load was manipulated using high and low memory load conditions. Additionally, looking for potential modulations related to individual differences, we studied the relationship between IQ, VWM span, and executive functions with VS efficiency. Finally, we also registered reported measures of potential strategies employed during the VS task. The results from a large sample of 147 participants between 5 and 25 years old revealed that even the youngest children could efficiently perform a VS task with a concurrent VWM load, replicating previous results found in adulthood. However, we found a slight increase in false alarms and commission errors when memory was highly loaded for all the participants regardless of age. As expected, we found positive correlations between VS efficiency and IQ and VWM span measures. Interestingly, the proportion of participants who used tracking organization strategies increased with age in all cases. However, although cognitive strategies to remember the target became more complex as age increased, it was only significant under the low VWM load conditions. The results seem relevant to understanding the development of VS based on the Model of Endogenous Mental Attention and the design of training

programs to improve attention. The implications in educational contexts are discussed and are especially relevant for students with learning disabilities or attention problems.

KEYWORDS

visual search, working memory, strategies, children, development, attention, executive functions, Theory of Constructive Operators

Introduction

Visual search is the basis of many activities daily performed in educational contexts. In kindergarten, children look for the necessary pieces to finish a tower of blocks or the red color among several paintings to draw the path in a maze. In primary school, children must look for spelling errors in grammar writing exercises or search for the isosceles triangle among other types of triangles in math exercises. Visual search is even present at advanced levels of education, i.e., when solving equations in college. These visual search tasks are often not performed in isolation, but rather students are maintaining information related to an ongoing task that can load working memory. These evolving processes improve as time passes and the cognitive system matures. Because of the importance of these two processes at all levels of education in school contexts, we will study how a concurrent working memory load task can modulate visual search efficiency in development within a large sample of 147 observers from 5 to 25 years old and structured on a neurodevelopmental perspective under the Model of Endogenous Mental Attention (Pascual-Leone and Johnson, 1999, 2005, 2021).

Visual search has usually been studied in cognitive development using two essential paradigms: feature and conjunction search tasks. In feature search (e.g., looking for a red pencil among blue pencils), the target is easily detected as it is defined by one feature that differentiates it from the distractors (color in our example: red vs. blue). Basic automatic processes capturing attention seem to be the base of feature search, and they have been found present in early stages of development (Gerhardstein and Rovee-Collier, 2002). However, for conjunction search, the target shares one or more features with the distractors (e.g., finding a red pencil among red pens and blue pencils), making it more challenging to search and thus requiring a more controlled and guided attentional process. Studies have shown variable results in conjunction search tasks, from set size functions clearly decreasing with age until adulthood (Donnelly et al., 2007; Michael et al., 2013; Woods et al., 2013; Brennan et al., 2017) to smoother non-dramatic set size modulations in the lifespan (Hommel et al., 2004). According to results, most researchers have postulated executive functions, specifically attentional

control, as the base to explain the results. Michael et al. (2013) and Woods et al. (2013) concluded that the executive functions needed to perform a visual search efficiently are in development during childhood, explaining the modulations. Namely, they must rely on inhibitory-attentional control, mental flexibility, planning, and working memory processes to perform a conjunction search-like, more complex search task. Indeed, a recent developmental study using conjunction search-like inefficient although more ecological, unique object search has shown visual search as a powerful paradigm to understand executive function development. Gil-Gómez de Liaño et al. (2020) found different developmental trajectories for accuracy, search slopes, and intercepts in visual search closely related to underlying executive functions described in Anderson's (2002) neuropsychological executive function model. The developmental course of accuracy essentially overlapped attentional control in Anderson's model, while slope functions overlapped goal setting, and intercepts followed the time course of cognitive flexibility and information processing.

Working memory seems to be a key executive function process in visual search. The attributes of the target that guide search, conceptualized as what Wolfe (2021) calls the “guiding template,” must be stored in working memory. Although with distinct hints other researchers have supported a similar viewpoint. Olivers et al. (2011) argue that the “target template” in visual search is probably maintained in working memory. Since working memory seems to play a crucial role in understanding guidance in visual search complex tasks like conjunction search or unique object searches, comprehending its development could be essential to understand differences in visual search among children of varying ages. How does performing a concurrent working memory task might affect visual search efficiency throughout development? Could working memory load differently affect performance in visual search in diverse developmental stages?

Working memory is known to be under development between 4 and 12 years old. Its development is closely related to the maturation of the prefrontal cortex, with significant changes between 6 and 10 years old and the age of 9–10 being a hallmark in working memory development

(Luciana and Nelson, 1998; Brooking et al., 2012; Simmering, 2012; Vuontela et al., 2013). However, most of these studies are based on models conceptualized from research on adults' working memory (e.g., models from Baddeley and Hitch, 1974; Cowan, 1988; Baddeley, 2000). These models have a limitation of not offering a developmental perspective. As Karmiloff-Smith (1995) proposed, we should focus on the “*changes over time*” of a developing mind rather than only when the child reaches adult levels in a given cognitive process.

An influential model studying attention in children that follows the developmental viewpoint suggested by Karmiloff-Smith is the Model of Endogenous Mental Attention proposed in the Theory of Constructive Operators (Pascual-Leone, 1970, 1995; Pascual-Leone and Johnson, 2005, 2021). Interestingly, our study pays special attention to understanding the development of controlled and effortful attention, executive attention, from a neurodevelopmental perspective.

According to the Theory of Constructive Operators, schemes are the basic unit of information, and they are expressed as neural networks in the brain (Pascual-Leone and Johnson, 2005; Arsalidou et al., 2019). Schemes can be of different types (figurative, operative, or executive), and all of them are under the regulation of domain-free operators, “functional mechanisms of brain hardware” (Pascual-Leone and Johnson, 2005, 2021). Each operator is related to a particular brain region, has a specific function (Arsalidou et al., 2019), and can be applied to schemes in any content (e.g., visual or auditory, Pascual-Leone and Johnson, 2005, 2021). Although Pascual-Leone and Johnson (2021) described eleven operators, we will only mention those applicable to the Model of Endogenous Mental Attention and directly related to the so-called Mental Attention potentially affecting the processes involved in a visual search task, that is, the M, I, E, and F operators (Pascual-Leone, 1995; Pascual-Leone and Johnson, 1999, 2005, 2021; Arsalidou et al., 2019).

Considering a visual search, we show in **Figure 1** a diagram of the Endogenous Mental Attention and operators that impact the efficiency of the task. The M-operator (mental attentional activation) is responsible for the effort of fully hyper-activating the necessary schemes to perform a task. In our visual search case, it is the figurative scheme of the “target” and the operative scheme of “scan and find.”

On the contrary, the I-operator (mental attentional inhibition) inhibits unwanted schemes that could lead to a task error. The Theory of Constructive Operators distinguishes between effortful inhibition and automatic inhibition (Howard et al., 2014; Pascual-Leone and Johnson, 2021). The automatic inhibition deactivates schemes effortlessly, especially in facilitating tasks. The effortful inhibition suppresses hyperactivated schemes that are incompatible with successful performance and that, otherwise, would remain active within the focus of the Mental Attention. Specifically, in the visual search task, the I-operator must inhibit the distractors (the

more features shared with the target, the greater the effort) or the schemes related to the concurrent tasks (in the case of the present experiment, schemes related to the concurrent working memory task). Poor functioning of one of the two operators, the M or the I, would cause a decrease in accuracy in the visual search task, that is, in the number of hits. Specifically, omissions would be related to a deficit in the M-operator due to insufficient activation of the operating scheme of the target itself. On the contrary, false alarms or commissions (confounding a distractor with the target) would be due to a problem with the I-operator, deactivating schemes incompatible with the task goal.

The E-operator controls activated executive schemes useful to achieve the task directly related to plans, goals, or strategies to perform the visual search. Therefore, a change in response speed or search slopes could be an indicator of the functioning of the E-operator. Moreover, studies on task efficiency (when controlling speed/accuracy trade-offs) can serve as a global measure of the performance of the I-, M-, and E-operators as a unit.

In the model, the E-, M-, and I-operators are related to the prefrontal lobe's function. The fourth operator that might apply in a visual search task, the F-operator, organizes the *field of mental attention*, unifying mental representations from a neo-Gestaltist point of view. Finally, the Schematic Overdetermination of Performance principle synthesizes all the most dominant schemes to produce a task-directed behavior (yes/no motor response). The coordinated functioning of the M-, I-, E-, and F-operators is essential, especially in complex situations (like conjunction searches or unique object search) when the cognitive demand increases because of salient or irrelevant information that keeps activated schemes potentially advocating an inadequate response in a task (refer to Pascual-Leone and Johnson, 2005, 2021, for a deeper explanation about the Model of Endogenous Mental Attention). As we can see, the Model of Endogenous Mental Attention allows for us to study the cognitive processes underlying the visual search and their development.

On the other hand, the interaction between working memory and visual search has been a key research question in psychological science. A typical paradigm followed in many studies testing adults in the field studies how performance in visual search can vary under different conditions of working memory load in dual-task paradigms. However, the results of these studies with adult population are variable. In a review, Soto et al. (2008) showed how the contents of a concurrent memory load task can attract attention in attentional tasks like visual search. Also, Lavie et al. (2004) and Lavie and de Fockert (2005) found that high loads in a concurrent working memory task could impair efficiency in visual search probably because of competition between control cognitive resources to avoid distractors in the attentional task and memory contents. However, other studies have found somehow the opposite result: better performance in the attentional task

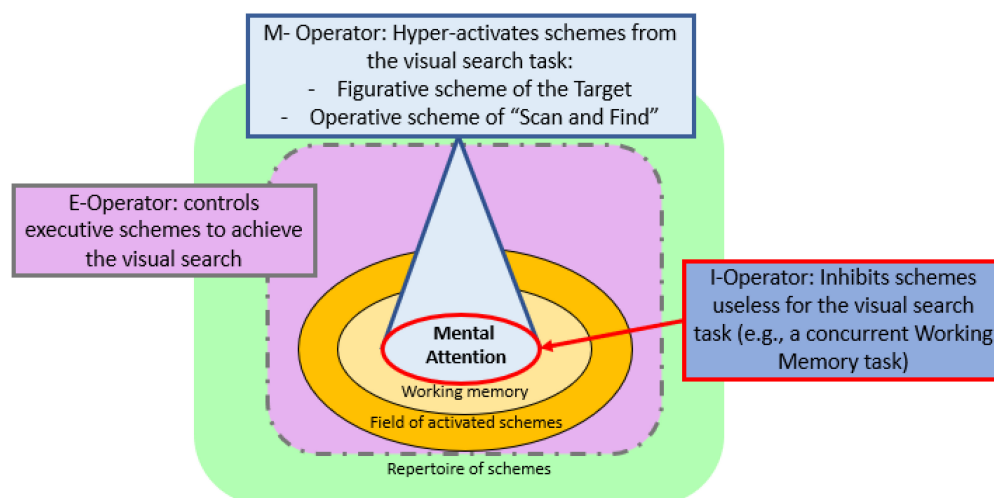


FIGURE 1

Diagram of the functioning of the mental attention during a visual search task, based on the visual representations of the Model of Endogenous Mental Attention from Arsalidou et al. (2019) and Pascual-Leone and Johnson (2021).

under high visual working memory loads. Smilek et al. (2006) studied how cognitive strategies could influence visual search performance during similar dual-task paradigms. They found better search performance by increasing the difficulty of the memory task when the search strategy was instructed to be a “passive” search letting the target “pop out.” This effect disappeared when the search strategy was to actively search for the target. Also, Gil-Gómez de Liaño et al. (2011) and Gil-Gómez de Liaño et al. (2014) found similar results using Rapid Serial Visual Presentation tasks as a way of search in time rather than in space. However, many other studies have just simply not found any modulation of working memory loads/contents in attentional tasks, particularly in visual search tasks. Downing and Dodds (2004) and Gil-Gómez de Liaño et al. (2011) tested the effects of working memory contents overlapping the distracters during visual search tasks but failed to find any modulation of memory contents or any difference between high and low loads in search performance. Both studies support theories in which working memory is fractionated, allowing for the maintenance of non-essential items in the visual search task. Vogel et al. (2001) and Woodman and Luck (2007) also found similar results and argued that knowing that the items stored in working memory could never be a target in the search task, participants can strategically inhibit them during the search. Moreover, Gil-Gómez de Liaño et al. (2016) showed again similar “lack of effects” results non-replicating previous studies in a set of experiments increasing power and pointing toward the necessity of replication studies to increase the reliability of studies in the field. That study led to a meta-analysis (Quirós-Godoy et al., 2017) concluding that working memory load did not interact with visual search (and other attentional tasks like the flanker

task or Stroop-like tasks) and supporting again that lack of interaction.

However, to our knowledge, there are no similar dual-task articles studying working memory and visual search interactions in children populations. A recent correlational study, though, points toward a potential modulation. Guilbert et al. (2020) found that better working memory abilities measured by Backward Digit Span let children from 8 to 11 years be more organized during a visual search in the Bells cancellation task. However, they did not study how different memory loads might affect visual search in children and how that could vary in development at different ages.

The Model of Endogenous Mental Attention (Pascual-Leone and Johnson, 2005, 2021) offers a neurodevelopmental perspective that could allow for us to understand working memory and visual search interaction development in different age stages. Under this model, working memory encompasses all hyperactivated schemes (e.g., schemes boosted by other operators like motivational or affective schemes) including those boosted by Mental Attention in coordinated studies on the M- and I-operators (Pascual-Leone, 2000). That is why the development of the M-operator and the I-operator plays a relevant role in explaining working memory development. The capacity of the M-operator is limited. It has a maximum number of schemes that it can keep effortfully active at a time. Empirical data show that M-capacity increases by one scheme every 2 years from about 3 years old to adolescence (15–16 years). In this developmental stage, the M-operator reaches its maximum of seven simultaneous schemes at once. Added to this is “e,” which refers to the capacity developed during the sensorimotor stage before 3 years of age (Pascual-Leone and Johnson, 2005, 2021; Arsalidou et al., 2019).

On the other hand, the development of the I-inhibitor also distinguishes between the automatic inhibition and the effortful inhibition, as previously mentioned: while at 7-years of age one can automatically inhibit as effectively as adults, in the case of the effortful inhibition, there is a gradual improvement in adulthood until older ages (even 12-year-olds perform significantly worse than adults; Howard et al., 2014).

Finally, and although secondary to the main objectives of this study, it is important to mention that other cognitive processes might interact with visual search performance (Michael et al., 2013; Woods et al., 2013) that we should consider. In the Model of Endogenous Mental Attention framework, Arsalidou et al. (2019) described a relationship between Mental Attention and fluid intelligence. Navarro et al. (2006) and Howard et al. (2013) found that gifted children had higher M-measurements than their non-gifted peers. A higher M-capacity would free up space for executive schemes and be more efficient in complex visual search tasks under higher working memory load situations. Thus, if higher intelligence quotient (IQ) is related to better Mental Attention performance, would children with higher IQ levels also be more efficient in visual search tasks under high working memory loads? Also, metacognitive strategies are associated with better performance in cognitive tasks (Bewick et al., 1995) and better academic performance (Pintrich and de Groot, 1990; Palos et al., 2011). During a visual search, there are two moments when participants could use metacognitive strategies: when coding the target in working memory (for instance by repetition or by mental imagery) and organizing visual tracking (looking in reading-order, up-down, left-right). However, not all children can self-generate strategies or use fewer effective strategies. It can depend on their developmental moment. For example, younger children of 4–5 years tend to prefer visual coding in working memory, while older children above 8–10 years tend to prefer verbal codes to maintain temporal information in working memory (Palmer, 2000; Henry et al., 2012). In the Model of Endogenous Mental Attention, the executive schemes are metacognitive strategies controlled by the E-operator but hyperactivated by the M-operator.

This study aims to understand visual search modulations under different working memory load conditions in a dual-task paradigm in a developmental study. To our knowledge, we are the first to study these interactions under those classic dual-task paradigms. We will use the Pirate-Treasure visual search task designed by Gil-Gómez de Liaño et al. (2020) in a sample of kindergarten to college students. Observers must look for real-world photorealistic targets while maintaining similar objects in visual working memory. We will refer to working memory without specifying the kind of content stored because the Model of Endogenous Mental Attention describes the operators as content-free. We will manipulate working memory load by including two memory load conditions: high load with four images to remember and low load with one. We will measure

accuracy, reaction times, search slopes, and efficiency in the visual search task to find differences under different memory load conditions. Although in our experimental design we would not expect to find differences in adults considering the evidence in the field, we expect to find a larger impairment in visual search for the youngest children, especially under high load conditions, as both attentional control and working memory capacity are still in maturation. In terms of the Model of Endogenous Mental Attention, if the I-operator is not yet developed in the early developmental stages to deactivate the schemes of the concurrent working memory task, there should be a decrease in accuracy and a general decrement in efficiency performing the visual search task. Also, we will control and study how other secondary potentially related individual differences in intelligence, executive functions, and working memory span can affect visual search efficiency. We expect to find a positive correlation between visual search efficiency and intelligence, working memory span, and executive function capacity. Finally, we will ask the participants about the type of strategy followed during the task to understand how metacognitive strategies could affect potential changes in visual search performance and whether they could be related to age strategy changes during development. We will discuss the implications in educational contexts.

Materials and methods

Participants

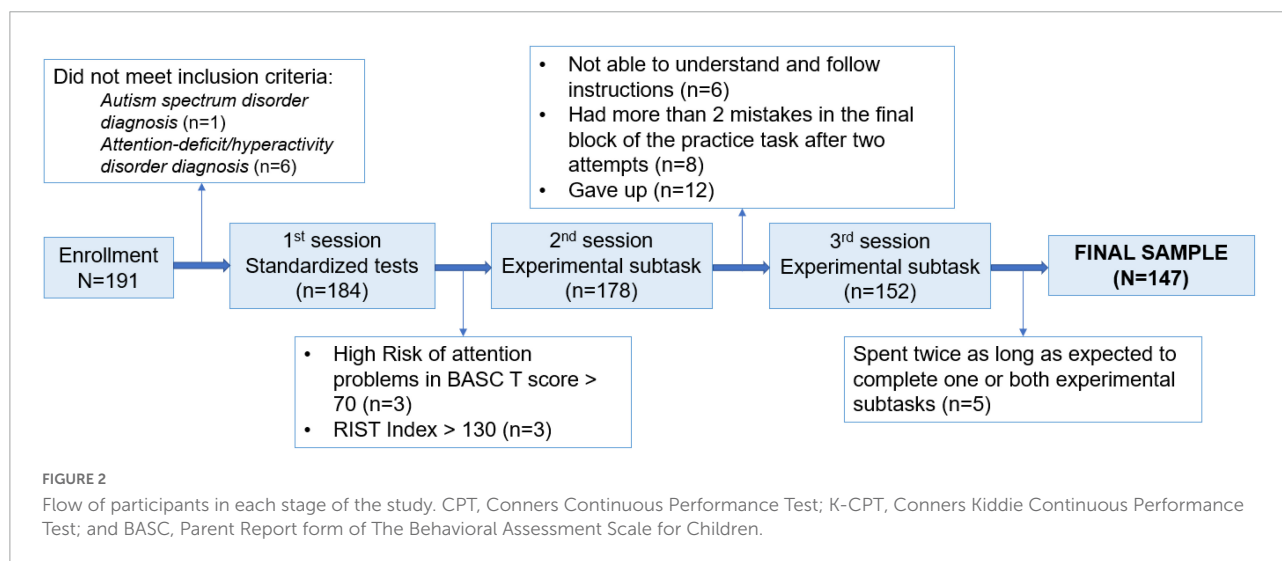
We recruited 191 participants from public and private elementary and middle schools in Madrid, Spain, and college students from Universidad Autónoma de Madrid. From previous studies on the lifespan in visual search (Gil-Gómez de Liaño et al., 2020), looking for age differences showed that with the alpha set to 0.05 and 1-beta (power) of over 0.9, we could detect significant effects (partial eta-square $\eta^2 = 0.01$) if we run between 21 and 33 participants per group of age. Thus, our sample size allowed for sizeable cohorts in each age, with a minimum of 21 for each age group (refer to Table 1). After missing several participants for different reasons explained in Figure 2, the final sample was composed of 115 children divided into four groups at different school levels and 32 psychology undergrad students (refer to Table 1 for sociodemographic descriptive statistics).

All the participants had a normal or corrected-to-normal vision and no history of neurological, motor impairment, or generalized developmental disorders. They gave verbal and/or written assent and written informed consent signed by parents or guardians in the case of minors. The children got “Pirate” diplomas as a reward, and the college students got credits for courses. The study was approved in advance by the ethics committee of Universidad Autónoma de Madrid (CEI-84-1553).

TABLE 1 Sociodemographic descriptive statistics of the final sample for age and IQ.

Grade		n	Age				IQ			
			M	SD	Min	Max	M	SD	Min	Max
Kindergarten	TOTAL	25	5.6	0.5	5	6	104.42	11.11	84	129
	Male	11	5.64	0.505	5	6	102.18	7.92	84	110
	Female	14	5.57	0.51	5	6	106.31	13.26	88	129
2nd elementary	TOTAL	28	7.14	0.36	7	8	109.86	10.45	77	122
	Male	16	7.06	0.25	7	8	109.81	13.03	77	121
	Female	12	7.25	0.452	7	8	109.92	6.05	101	122
4th elementary	TOTAL	31	9.16	0.37	9	10	109.65	11.94	74	130
	Male	13	9	0	9	9	111.77	11.25	91	130
	Female	18	9.28	0.46	9	10	108.11	12.49	74	125
6th middle school	TOTAL	31	10.97	0.41	10	12	104.9	10.02	83	122
	Male	8	10.88	0.35	10	11	108.75	10.91	88	122
	Female	23	11	0.43	10	12	103.5	9.56	83	121
College students	TOTAL	32	19.78	1.84	18	25	102.09	12.37	77	129
	Male	11	19.91	2.39	18	25	103.73	14.14	77	129
	Female	21	19.71	1.55	18	25	101.24	11.61	79	123

M, mean; SD, standard deviation; Min, minimum; Max, maximum; IQ, intelligence quotient measured by Reynolds Intellectual Screening Test (RIST).



Instruments and procedure

Each participant completed three sessions, the children at schools and the college students at the university. We applied several standardized tests to eliminate potentially non-typical developing individuals in the first session and control and study potential effects based on individual differences.

Session 1: The standardized tests

Intelligence quotient (IQ) was measured by *Reynolds Intellectual Screening Test* (RIST; Reynolds and Kamphaus, 2003). RIST includes a verbal subtest (*Guess What*, crystallized intelligence screening) and a non-verbal subtest (*Odd-Item Out*,

fluid intelligence screening), and it takes between 20 and 30 min to complete.

We assessed visual working memory by the *Picture Span* subtest from *Wechsler Intelligence Scale for Children 5th Edition* (WISC-V, Spanish edition; Wechsler et al., 2015). The participants had to remember a set of pictures to recognize later in the same order with difficulty increasing in each trial. Standardized scores are only available for children from 6 to 16 years old. Therefore, we used the eldest scale range (16 years and 11 months) for the young adults' sample.

Parents filled out two standardized questionnaires for the children: first, the *Parent Report* form of *The Behavioral Assessment Scale for Children* (BASC;

Reynolds and Kamphaus, 2004) measures adaptive and behavioral problems in the community and at-home settings. The test can identify and differentiate between attention problems and hyperactivity. Second, we used the *Parent Report of The Behavior Rating Inventory of Executive Function* (BRIEF; Gioia et al., 2015), which detects potential difficulties in executive functions at home and school.

Finally, the young adults and parents filled out a short questionnaire about the participants' basic development and medical background.

Sessions 2 and 3: The experimental tasks

The experiment was coded using *E-prime 3.0* (Psychology Software Tools Inc., 2016). Stimuli were 2,469 real-word images provided by Brady et al. (2008). We divided the images into four sets: two for the visual search (targets and distractors pool of images) and two for the working memory concurrent task (images to remember and show as distractors in the working memory probe and testing phase). We aimed to avoid any potential confounding/overlapping of the images between the tasks. Thus, any visual search image would never appear in the working memory set and vice versa. For the target set of images in the visual search task, we included only child-friendly images to increase motivation for the participants. They responded with a touch-screen computer (*Microsoft surface Pro i5*) with an 800 pixel × 600 pixel monitor resolution. Image size was 96 pixels × 96 pixels each.

The participants must follow some basic instructions before starting the task: both hands had to be at the sides of the tablet on the table, and they could only respond with their dominant hand. After every tap-answer, the hand must return to rest on the tablet's side again. If a child was too young to follow the basic instructions, having problems inhibiting their non-dominant hand, they should put the non-dominant hand under their thighs on the chair.

The experimental task included two subtasks: a subtask with low working memory load (low-load, one image) and another subtask with high working memory load (high-load, four images). Low-load took 10–20 min, and high-load took around 20–30 min. The subtasks were counterbalanced, with half of the participants running low-load/high-load and the other half high-load/low-load. We run them on different session days, sessions 2 and 3 of the whole testing. There were no more than 3 days between sessions.

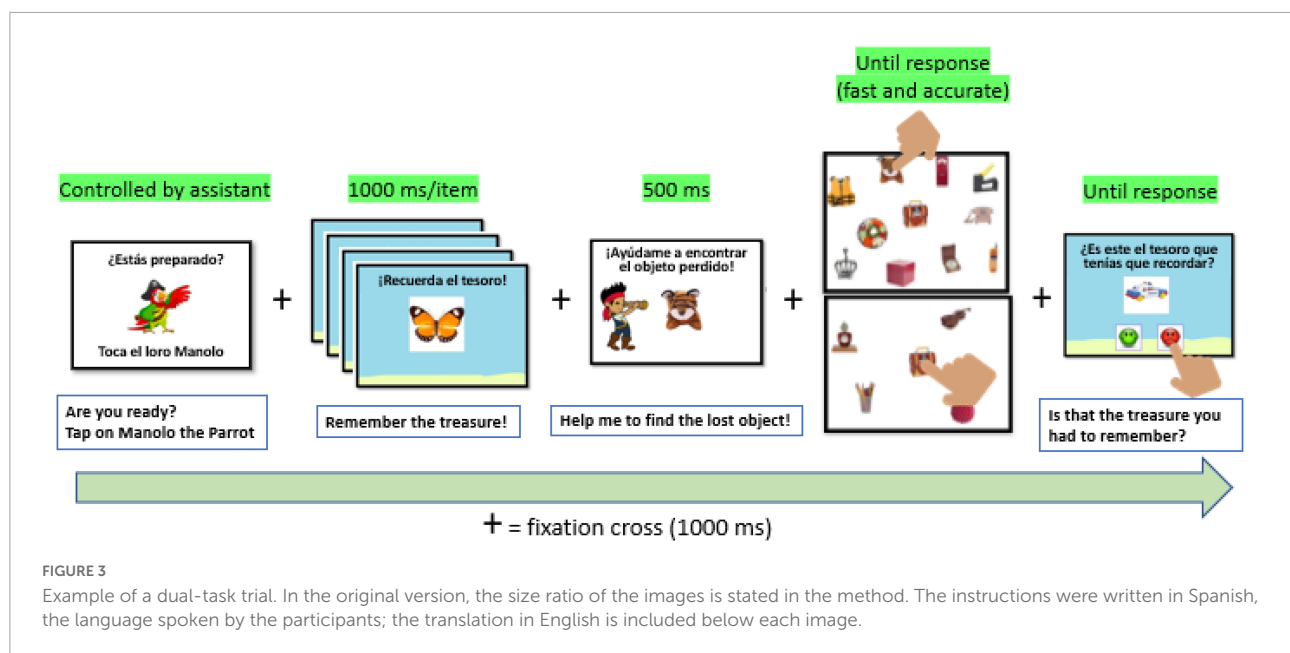
We told the participants they were pirates with two missions: look for treasures stolen by evil pirates and remember the treasures buried on the beach to prevent further thefts. An example of a full trial is shown in Figure 3. A trial started with the picture of a parrot that remained until response (tapping anywhere on the screen) to prepare for the subsequent trial. That way, every participant could adjust the speed of the task according to the age and control they were attending to the task. Then, a white background with a centered cross remained on

the screen for 1,000 ms before the beach-treasure/s (one for the low-load and four for the high-load) were presented in the beach background for the working memory maintenance task. Images for the concurrent working memory task appeared on the center of the screen one at a time for 1,000 ms, one after another. For the low-load, there was only one image lasting 1,000 ms, too. Then, the target for the visual search appeared for 500 ms in the same position (refer to Figure 3). Right after, the search display appeared.

The participants had to look for the target among distractors and respond as quickly and accurately as possible. Set size could be 8 or 32 items, including the target, randomly presented. Half of the trials contained the target, and the other half did not. They were presented in random order. Thus, there were 7 distractors plus the target for the target-present trials, while 8 items were distractors for the target-absent trials. The same rationale was applied for set size 32. The target position for each trial was also randomized. If the target was present, the participants must tap on it. For the target-absent trials, they had to tap on the pirate chest in the center of the screen (which appeared in all the trials regardless if the target was or was not present). Thus, it was both an identification and a localization visual search task. Search displays remained on the screen until a response. Finally, the beach background appeared again for the working memory test. The test consisted of an image that was the one shown in the low-load condition or one of those shown in the high-load condition for half of the test trials. The other half contained a different image randomly selected from the working memory pool. If the testing image was one of the high-load or the low-load images, the observer must tap on a green smiling face shown at the bottom of the screen, while if it was not, they must tap on the sad red face also at the bottom (refer again to Figure 3). The order of green/red correct responses was again randomized. The beach-test display remained on the screen until a response with no emphasis on speeded responses as in the visual search task.

The final design included set size (8/32), target (present/absent), and working memory load (low-load/high-load) as within-subjects factors, while grade was the between-subject factor (kindergarten, KG; 2nd elementary, 2nd; 4th elementary, 4th; 6th middle school, 6th; and college students, college). Each experimental subtask in sessions 2 and 3 consisted of 120 trials, with 240 for the whole task. Every 40 trials, a "progress bar" showed up as a reinforcement to keep the participants motivated throughout the task. A congratulatory message for recovering all the pirate treasures was also displayed at the end of each subtask. The participants initially performed a 24-trial practice block with feedback and the experimenter present to make sure they understood the task for every subtask. There was no feedback in the experimental blocks.

At the end of each experimental session, we asked the participants about the cognitive strategies potentially used during the visual search. First, we asked them if they used any "tricks" to remember the objects in each trial (potential



maintenance strategies). We asked them to choose between using “no trick; look and that’s it,” no strategy, a perceptual strategy (“color,” “shape,” or “picture”), and a verbal strategy (“naming/repeating the object out loud or in your mind”). Then, we asked them an open-response question, if they had used any “trick” to search faster (looking for potential visual tracking strategies). Because of the small number of participants reporting using tracking strategies (e.g., searching from left to right or up-down), we divided them into those using them and those not using them for the analysis.

Data analysis

The analyses were conducted using R version 4.0.3 (R Core Team, 2020) and SPSS Statistics 26 (IBM Corp, 2019).

Visual search performance

Trials with null responses in the visual search or the working memory task, that is, when the participants tapped anywhere outside the response areas (the images), were eliminated (<0.01%). In the case of the visual search task, trials with responses shorter than 200 ms and longer than 9,000 ms were also eliminated (<0.04% of the data), following Gil-Gómez de Liaño et al. (2020)’s criteria. We performed visual search data analysis only for trials with a correct response in the working memory task (85% of the trials) to ensure that our working memory manipulation effectively loaded memory during the search. The results were similar when all the trials were included (also working memory task errors). To ensure that the

manipulation of the working memory task has been sufficient to cause two different load levels, interested readers can find the working memory accuracy analysis in the [Supplementary Material](#).

For each condition of the visual search task, we collected latency (reaction times, RTs) and accuracy measures (correct responses; false alarms, tapping a distractor when the target is absent; commissions errors, tapping a distractor when the target is present; omissions errors, tapping the pirate chest when the target is present). Note that like in Gil-Gómez de Liaño et al. (2020), the design of our experiment allows for us to analyze commission errors when the target is present since the observers must localize the target. Thus, the observers can identify it as a non-target and tap instead on a distractor identified as the target in those error trials. We used the raw data for correct responses, false alarms, commissions, and omissions, analyzing them as a dichotomous dependent variable (yes/no response). Mean RTs (in milliseconds, ms) and search slopes were calculated for each participant’s different combinations of variables (refer later on the use of General Linear Mixed Models, GLMMs).

In addition to the latency and accuracy measures in visual search, we calculated an efficiency measure that controlled the trade-off between speed and accuracy that usually occurs in visual search tasks: the observers adjust their response speed to make fewer mistakes based on the task’s difficulty. The Inverse Efficiency Score (Townsend and Ashby, 1983) is one of the most widespread indices. However, we need to meet two assumptions: the proportion of correct responses must be greater than 0.90 (Bruyer and Brysbaert, 2011), and there must be a strong positive linear correlation between RT and the proportion of errors (Townsend and Ashby, 1983). According to the present study’s data, we cannot assume these two assumptions. As

an alternative, we calculated the Efficiency Score proposed by Roncadin et al. (2007). The formula is as follows:

$$\text{Efficiency Score} = \frac{RT(RT - 100n)/RT(RT + 100n)}{RT(RT - PCn)/RT(RT + PCn)}$$

where RT is the individual's mean reaction time, PC is the individual's mean proportion of correct responses, and n is the sample's minimum RT divided by twice the sample's maximum proportion of correct responses. The minimum RT for the present sample was 706.34 ms, and the maximum proportion was 100 ($n = 3.532$). The distribution of the Efficiency Score falls between 0 (worst performance, the proportion of correct responses is 0) and 1 (perfect performance). Thus, higher scores reflect better efficiency regardless of observers' response styles.

We conducted accuracy analyses using GLMM for the log-odds ratio of correct responses, false alarms, commissions, and omissions to control individual variability. We used linear mixed-effects models for continuous variables (RT , search slopes, and efficiency score).

The hierarchical structure of the data was analyzed under a two-level random intercept model. Independent variables were included in level 1 and organized within observers (level 2). In level 1, for efficiency score, correct responses and RT -dependent variables, working memory load, target, and set size were the independent factors. In the case of errors, the factors were working memory load and set size. Finally, working memory load and target were the independent variables for the search slope analysis. In all the analyses, grade was added as a predictor to level 2 to study the developmental effect, which is one of our main objectives of the present study.

The estimation of fixed regression coefficients was based on maximum likelihood to allow for model comparisons (Field et al., 2012). We conducted Hommel corrections (Hommel, 1988) for *post hoc* comparisons between conditions in the case of the linear mixed-effects models. For comparisons in the GLMM, and due to the difficulty usually found in interpreting the odds ratios in this type of model, we used plots that include their transformation to estimated proportions and the confidence intervals of each estimation, as recommended by Gelman and Hill (2007).

To build the final models, we started from the null model, not including parameters, in which the intraclass correlation coefficients (ICCs) were between 0.017 and 0.33. This verifies the individual variability among the participants and the pertinence to use such models instead of conducting regular ANOVAs. We added to the null model the parameters, main effects as well as interactions, one by one to the null model, testing the fit to the data and the improvement of the model in each step by comparing changes in the log-likelihood between models, as well as Akaike's information criterion (AIC) and Schwarz's Bayesian Criterion (BIC). Parameters that did not significantly improve the fit were discarded. For interested readers, the building progression of the final models is shown in

Supplementary Material, including comparisons with the null model and the maximal model with all the parameters included.

Individual differences and visual search efficiency

To study the relationship between individual differences and efficiency in visual search, we conducted a partial correlation analysis controlling for age (in months) to eliminate developmental effects. Specifically, we correlated the average efficiency score in each of the eight conditions resulting from combining the three within-subjects experimental factors in visual search (working memory load, set size, and target) with the RIST T-Scores (both for the two subtests as well as for the general index), the scalar scores provided by the *Picture span* test from WISC-V, and the BRIEF and BASC T-Scores. Correlations between ES and BRIEF, BASC, and *Picture span* were only analyzed for the children since the tests were not scaled for the population over 17 years of age.

Metacognitive strategies change

In the working memory low-load session, 9 participants did not answer the memory strategy question, and 7 did not answer the visual tracking one. In the working memory high-load condition, they were 13 participants who did not answer the memory strategy question and 11 who did not respond to the visual tracking question. We did not consider these participants for the final analysis in the study on metacognitive strategies. We performed chi-square tests of independence to assess whether memory and visual tracking strategy changes were related to the different age groups. We calculated Cramer's V to measure effect size.

Results

Visual search performance

In **Figure 4**, we can see the mean proportion of correct responses (A), reaction times (B), search slopes (C), and efficiency score (D) for the visual search task as a function of the target (present/absent), working memory load (low-load/high-load), set size (8–32, except for search slopes), and grade (KG to college).

The M-operator and the I-operator functioning: accuracy, false alarms, commissions, and omissions

As we can see in the upper panels of **Figure 4A** (target-present conditions), the proportion of correct responses in

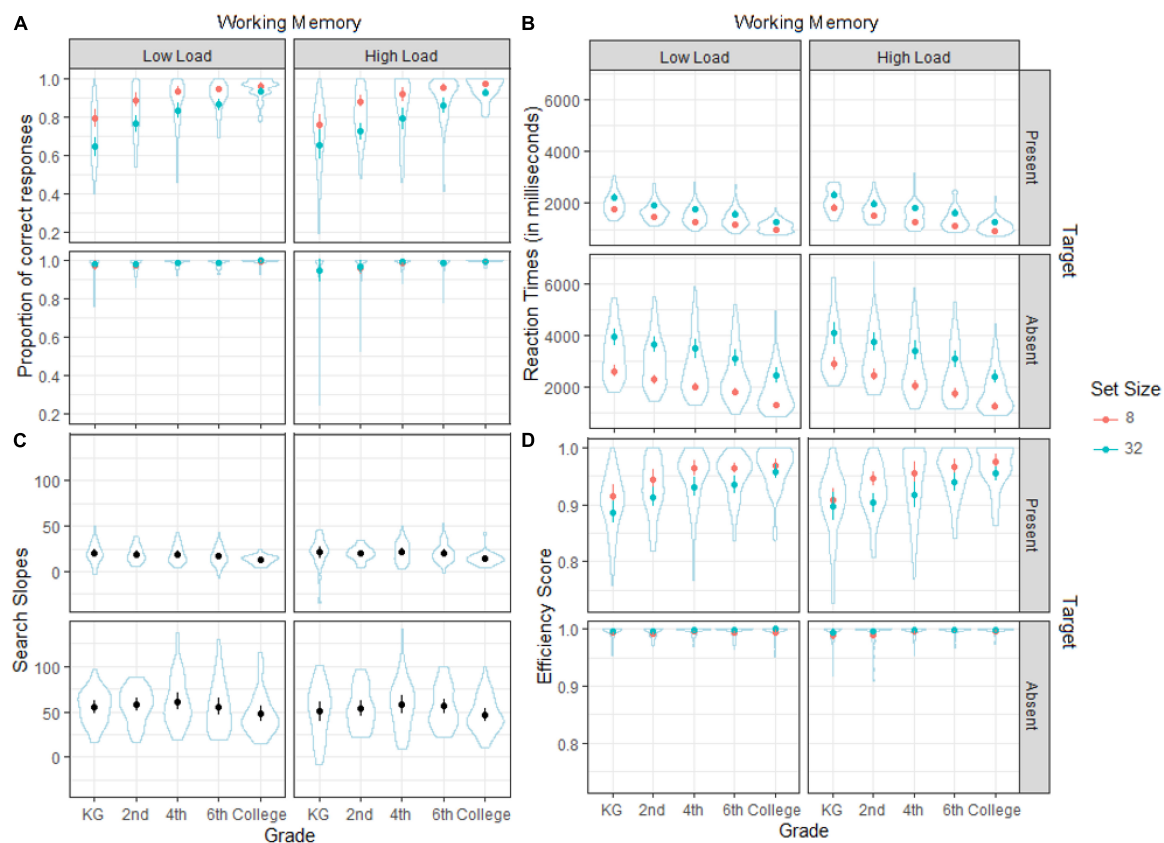


FIGURE 4

Mean proportion of (A) correct responses, (B) reaction time, (C) search slopes, and (D) efficiency score as a function of grade, set size, working memory load, and target. Mean of search slopes as a function of grade, working memory load, and target. Bars represent the confidence intervals for each mean represented by dots.

the visual search task seems higher for the low-load than for the high-load, as corroborated by the significant effect of working memory shown in the GLMM analysis in [Table 2](#). The probability of a hit in the visual search is higher for the low-load condition. However, interactions between working memory and target or set size were not significant (refer again to [Table 2](#)).

Regarding the rest of the factors, the main effect of target was significant: The likelihood of a correct response was higher when the target is absent (showing a ceiling effect on all grades regardless of set size, as shown in [Figure 4A](#)). The main effect of set size was also significant, showing worse accuracy as distractors increase, although this occurred only when the target was present (refer to confidence intervals in [Figure 5A](#)).

For grade, as expected the significant main effect indicated an increase in accuracy with age. Its interaction with target was also significant (refer again to [Table 2](#)): for the target-present conditions, there was a gradual improvement in correct responses with age, especially pronounced at younger ages, while for target-absent the ceiling effects appeared from 4th grade onward (although even younger children showed very high accuracy; see confidence intervals in [Figure 5B](#)). There was also

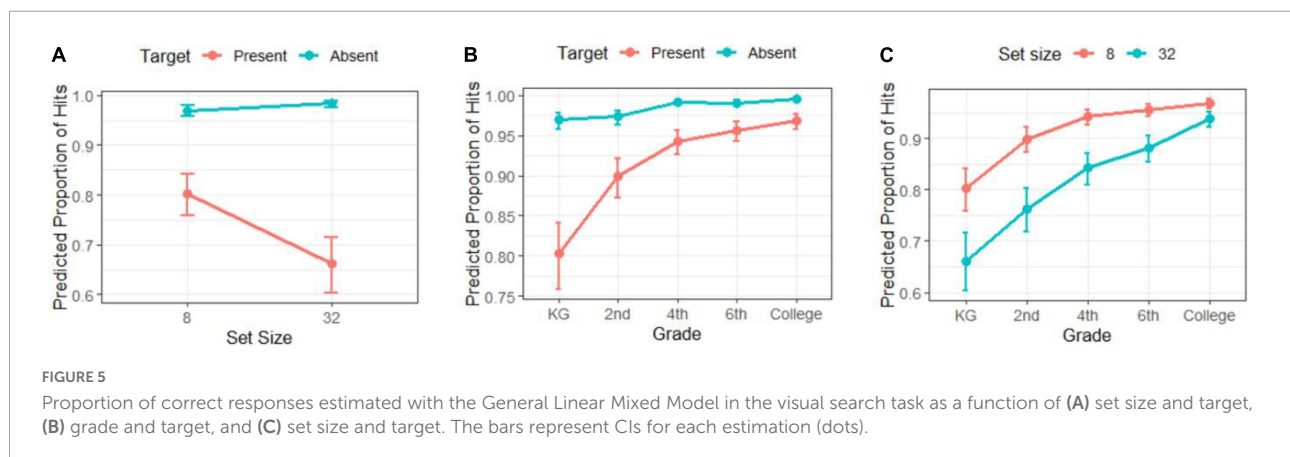
a main effect for the interaction between set size and grade (refer to [Table 2](#)). Changes between grades are smoother with fewer distractors (set size 8) from 4th grade observers onward, while under search conditions with more distractors (set size 32), there are still significant changes in the higher grades (see confidence intervals in [Figure 5C](#)). Also, the difference in accuracy between set size 8 and 32 was minor in college observers compared to the rest of grades. Although with minor changes, these results essentially replicate those found in [Gil-Gómez de Liaño et al. \(2020\)](#) under no working memory load conditions.

To study if the effect of working memory load that we found in accuracy affects the M-operator and the I-operator in the same way, we analyzed the error patterns separately. Note that false alarms and commissions are related to failure of the I-operator to deactivate schemes irrelevant to the task, while omissions would be due to lack of target hyper-activation by the M-operator. For false alarms and commissions, as we can see in [Table 3](#), the probability of committing those types of errors was higher when working memory is highly loaded and as grade decreases. However, the main effect of grade was not affected by working memory (not significant interaction), showing that the

TABLE 2 Estimated coefficients for correct responses (odds ratios) in the visual search task.

Fixed effects – predictors	OR	SE	CI	<i>t</i>	<i>p</i>
Intercept (KG, working memory low load, target present, set size 8)	4.08	0.55	3.13 – 5.30	10.47	<0.001
Set size (32)	0.48	0.05	0.40 – 0.58	−7.58	<0.001
Target (absent)	7.90	1.24	5.81 – 10.74	13.17	<0.001
Working memory (high load)	0.89	0.04	0.81 – 0.98	−2.36	0.018
Grade (2nd)	2.19	0.41	1.52 – 3.16	4.20	<0.001
Grade (4th)	4.07	0.78	2.79 – 5.92	7.33	<0.001
Grade (6th)	5.42	1.07	3.68 – 7.96	8.59	<0.001
Grade (college)	7.71	1.58	5.16 – 11.53	9.95	<0.001
Setsize (32) × Target (absent)	3.85	0.59	2.85 – 5.20	8.78	<0.001
Target (absent) × Grade (2nd)	0.53	0.10	0.36 – 0.77	−3.26	0.001
Target (absent) × Grade (4th)	0.93	0.22	0.58 – 1.49	−0.31	0.754
Target (absent) × Grade (6th)	0.62	0.15	0.39 – 0.99	−2.01	0.045
Target (absent) × Grade (college)	0.88	0.27	0.48 – 1.60	−0.43	0.670
Set size (32) × Grade (2nd P)	0.75	0.10	0.58 – 0.99	−2.05	0.040
Set size (32) × Grade (4th P)	0.67	0.10	0.50 – 0.90	−2.64	0.008
Set size (32) × Grade (6th P)	0.70	0.11	0.51 – 0.95	−2.26	0.024
Set size (32) × Grade (college)	1.01	0.18	0.71 – 1.43	0.07	0.946
Random effects					
σ^2			3.29		
τ_{00} Subject			0.29		
Intraclass correlation coefficient			0.08		
Marginal R^2 /Conditional R^2			0.376/0.427		

OR, odds ratio; SE, standard error; CI, confidence interval. Refer to the building progression of the model in [Supplementary Material](#). The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.



likelihood of showing these errors was lower with age regardless of memory load. The effect of set size was only significant for the false alarms: it was more likely to have a false alarm with 32 objects on the screen than with 8 objects.

For omissions, however, the working memory load did not reach significance in the final model (refer to [Table 3](#)), nor did the interactions with set size and grade. Grade was significant, though, showing that the probability of omissions decreased with age. Nonetheless, differences were minimal for older children from the 4th grade onward for set size 8, as we can see comparing the confidence intervals in [Figure 6](#), while the

younger ones showed more pronounced effects between them and the rest. The probability of missing a target was higher when the search display was 32 for all the groups (refer again to [Table 3](#) and [Figure 6](#)).

The E-operator functioning: reaction times and search slopes

As we can see in [Table 4](#), the main effect of working memory load is again significant, with larger RTs under high-load conditions and in the same way as the proportion of correct responses. However, there were no interactions between

TABLE 3 Estimated coefficients for false alarms, commissions, and omissions (odds ratios) in the visual search task.

Fixed effects – predictors	False alarms					Commissions					Omissions				
	OR	SE	CI	<i>t</i>	<i>p</i>	OR	SE	CI	<i>t</i>	<i>p</i>	OR	SE	CI	<i>t</i>	<i>p</i>
Intercept	0.02	0.00	0.01 – 0.03	−14.13	<0.001	0.00	0.00	0.00 – 0.01	−11.69	<0.001	0.24	0.03	0.19 – 0.32	−10.40	<0.001
Working memory (high load)	1.38	0.20	1.04 – 1.84	2.25	0.025	2.49	0.63	1.52 – 4.07	3.63	<0.001					
Set size (32)	0.69	0.10	0.52 – 0.91	−2.57	0.010						1.97	0.20	1.61 – 2.40	6.58	<0.001
Grade (2nd)	0.99	0.35	0.49 – 2.00	−0.03	0.978	0.59	0.35	0.18 – 1.91	−0.88	0.380	0.44	0.09	0.30 – 0.65	−4.21	<0.001
Grade (4th)	0.35	0.13	0.16 – 0.73	−2.77	0.006	0.36	0.22	0.11 – 1.19	−1.67	0.096	0.24	0.05	0.16 – 0.35	−7.27	<0.001
Grade (6th)	0.35	0.13	0.17 – 0.75	−2.72	0.006	0.27	0.17	0.08 – 0.93	−2.08	0.038	0.19	0.04	0.12 – 0.28	−8.22	<0.001
Grade (college)	0.16	0.07	0.07 – 0.36	−4.37	<0.001	0.07	0.06	0.02 – 0.35	−3.28	0.001	0.13	0.03	0.08 – 0.19	−9.71	<0.001
Set size (32) × Grade (2nd)											1.46	0.22	1.09 – 1.97	2.52	0.012
Set size (32) × Grade (4th)											1.69	0.27	1.23 – 2.31	3.25	0.001
Set size (32) × Grade (6th)											1.55	0.26	1.11 – 2.16	2.55	0.011
Set size (32) × Grade (college)											1.17	0.22	0.81 – 1.70	0.82	0.410
Random effects															
σ^2	3.29					3.29					3.29				
τ_{00} Subject	1.01					2.29					0.30				
Intraclass correlation coefficient	0.24					0.41					0.08				
Marginal R^2 /Conditional R^2	0.112/0.321					0.147/0.497					0.160/0.231				

OR = odds ratio, SE = standard error, CI = confidence interval. Intercepts for omissions were KG and set size 8 and for false alarms and commissions KG and low working memory load. The rows without data correspond to parameters that did not improve the models and were not included. Refer to the building progression of the models in [Supplementary Material](#). The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.

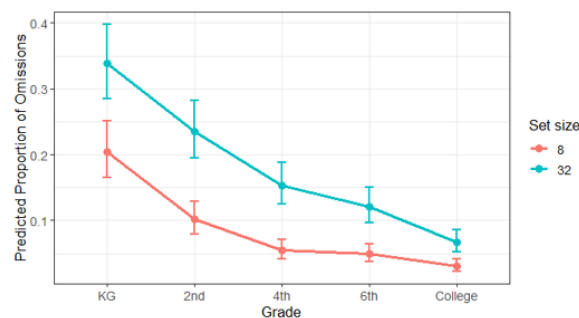


FIGURE 6

Proportion of omissions estimated with the General Linear Mixed Model in the visual search task as a function of grade. Bars represent confidence intervals for each estimation (dots).

working memory load and the rest of the visual search factors (target or set size) or a triple or quadruple interaction among all the factors. However, the interaction between working, memory load and grade showed a tendency for the youngest children (KG and 2nd) with higher RTs for high-load conditions compared to the older ones. *Post hoc* comparisons with no corrections showed significant differences between load conditions for KG ($p = 0.007$) and 2nd ($p = 0.056$), with these younger children spending more time when the target was present in the high-load condition. Indeed, if we consider RT differences between both levels in working memory load in different ages, KG obtained significant RT differences compared to the rest of the groups, except for the 2nd grade (refer to [Table 4](#) again). There were no differences among the rest of the groups.

Again, as expected from previous studies, RTs were larger for set size 32 and target-absent conditions (refer to [Table 4](#)). The interaction was also significant: set size effects were larger under target-absent conditions, replicating previous results again (Gil-Gómez de Liaño et al., 2020).

As for the correct responses, the effect of grade produced a gradual decrease in RT. The difference was significant even for the 6th graders and college students, as shown in [Table 5](#), in the *post hoc* comparisons.

Interestingly, the differences in RT between set sizes 8 and 32 were significantly lower for the college students than for the rest of the grades ($\beta = 142.3$, $p = 0.054$ for KG; $\beta = 171.6$, $p = 0.017$ for 2nd; $\beta = 223.9$, $p = 0.001$ for 4th; $\beta = 160.3$, $p = 0.022$ for 6th). There were no differences among all the children, indicating that the interference due to increase in the distraction (set size manipulation) was indeed lower for the young adults than for the rest of the observers in the sample. This effect for older children in the 4th and 6th grades (9–12 years old) was not found in Gil-Gómez de Liaño et al. (2020). Although we have not seen an interaction between working memory load and set size for the proportion of correct responses or RTs, maybe the presence of a concurrent working memory task has made a difference.

Finally, we analyzed the search slopes to differentiate whether the effect of memory load was caused by a problem in the E-operator or it was related to the processing speed. In this case, only the main effect of target was significant in the final model (refer to [Table 6](#)): slopes were steeper when the target was absent, indicating that the participants spent more time per item searching under these conditions, again replicating previous results in visual search with children (Gil-Gómez de Liaño et al., 2020), with no effect of working memory load.

Global functioning of the M, I and E-operators: efficiency score

We calculated the efficiency score (Roncadin et al., 2007) to determine to what extent the coordinated functioning of the M-, I, and E-operators might be affected by working memory load in visual search. The final model for efficiency score showed essentially significant main effects for grade, target, and set size (refer to [Table 7](#)). However, working memory load was not a significant parameter, neither were any of the interactions with the rest of the variables: Working memory load did not affect the visual search task's efficiency.

As expected, for set size 32, efficiency was lower, while for target-absent conditions, efficiency was higher (refer to the main effect in [Table 7](#)). However, the differences between set size conditions were only significant under target-present conditions ($p < 0.0001$) and disappeared under target-absent ones ($p = 0.165$). Finally, although there was a significant effect of grade with higher efficiency as age increases, it only arose again for target-present conditions: the KG children achieved the lowest efficiency followed by the 2nd grade children ($p < 0.0001$, for all the comparisons with the other grades and between them, except between 2nd and 4th where $p = 0.034$). There were no efficiency differences between 4th and 6th ($p = 0.755$) or between 6th and the College students ($p = 0.15$). However, the 4th-grade children were less efficient than the college group ($p \leq 0.001$).

TABLE 4 Estimated coefficients for reaction times in the visual search task.

Fixed effects – predictors	Estimates	SE	CI	<i>t</i>	<i>p</i>
Intercept (KG, working memory low load, target present, set size 8)	1738.10	96.98	1548.02 – 1928.17	17.92	<0.001
Set size (32)	446.65	60.21	328.64 – 564.66	7.42	<0.001
Target (absent)	909.94	60.21	791.92 – 1027.95	15.11	<0.001
Working memory (high load)	150.74	55.66	41.64 – 259.84	2.71	0.007
Grade (2nd)	–305.11	132.49	–564.78 – –45.43	–2.30	0.021
Grade (4th)	–466.60	129.43	–720.27 – –212.92	–3.61	<0.001
Grade (6th)	–564.17	129.43	–817.85 – –310.50	–4.36	<0.001
Grade (college)	–769.49	128.52	–1021.39 – –517.59	–5.99	<0.001
Set size (32) × Target (absent)	855.82	45.91	765.84 – 945.81	18.64	<0.001
Set size (32) × Grade (2nd)	28.32	76.58	–121.78 – 178.41	0.37	0.712
Set size (32) × Grade (4th)	80.62	74.81	–66.01 – 227.26	1.08	0.281
Set size (32) × Grade (6th)	16.97	74.81	–129.66 – 163.60	0.23	0.821
Set size (32) × Grade (college)	–143.32	74.29	–288.93 – 2.29	–1.93	0.054
Target (absent) × Grade (2nd)	–21.59	76.58	–171.69 – 128.50	–0.28	0.778
Target (absent) × Grade (4th)	–144.26	74.81	–290.89 – 2.38	–1.93	0.054
Target (absent) × Grade (6th)	–302.07	74.81	–448.70 – –155.44	–4.04	<0.001
Target (absent) × Grade (college)	–600.08	74.29	–745.69 – –454.48	–8.08	<0.001
Working memory (high) × Grade (2nd)	–50.77	76.58	–200.87 – 99.33	–0.66	0.507
Working memory (high) × Grade (4th)	–151.04	74.81	–297.67 – –4.41	–2.02	0.044
Working memory (high) × Grade (6th)	–180.39	74.81	–327.02 – –33.76	–2.41	0.016
Working memory (high) × Grade (college)	–176.01	74.29	–321.61 – –30.40	–2.37	0.018
Random effects					
σ^2			154923.09		
τ_{00} Subject			154370.24		
Intraclass correlation coefficient			0.50		
Marginal R^2 /Conditional R^2			0.702/0.851		

SE, standard error; CI, confidence interval. Refer to the building progression of the model in [Supplementary Material](#). The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.

TABLE 5 Post hoc comparisons for RT (in ms) in the visual search task between different grade levels under different target and set size conditions.

Comparison	Target present		Target absent		Set size 8		Set size 32	
	Mean difference	<i>P</i> -value	Mean difference	<i>P</i> -value	Mean difference	<i>P</i> -value	Mean difference	<i>P</i> -value
KG – 2nd	316	0.054	338	0.03	341.3	0.052	313	0.08
KG – 4th	502	<0.001	646	<0.0001	614.2	<0.0001	533.6	<0.001
KG – 6th	646	<0.0001	948	<0.0001	805.4	<0.0001	788.4	<0.0001
KG – college	929	<0.0001	1529	<0.0001	1157.5	<0.0001	1300.9	<0.0001
2nd – 4th	185	0.28	308	0.04	273	0.12	220.6	0.19
2nd – 6th	330	<0.05	610	<0.0001	464.1	<0.01	475.5	<0.01
2nd – college	613	<0.0001	1191	<0.0001	816.2	<0.0001	987.9	<0.0001
4th – 6th	144	0.28	302	0.042	191.2	0.28	254.8	0.15
4th – college	427	<0.01	883	<0.0001	543.3	0.0001	767.2	<0.0001
6th – college	283	0.06	581	0.0001	352.1	0.02	512.4	<0.001

The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.

Individual differences and visual search efficiency

We found significant positive correlations between the RIST and the *Picture span* tests with efficiency in the visual search

task, but only when the target was present (refer to [Table 8](#)). The RIST general index was positively correlated with the remaining four efficiency measures: the higher the capacity (IQ), the higher the efficiency. Specifically, the high scores in the *Odd-Item Out* subtest were related to higher efficiency under

TABLE 6 Estimated coefficients for search slopes in the visual search task.

Fixed effects – predictors	Estimates	SE	CI	<i>t</i>	<i>p</i>
Intercept (target present)	18.39	1.29	15.86 – 20.93	14.23	<0.001
Target (absent)	35.66	1.20	33.3 – 38.01	29.68	<0.001
Random effects					
σ^2			212.17		
τ_{00} Subject			139.64		
Intraclass correlation coefficient			0.40		
Marginal R^2 /Conditional R^2			0.475/0.683		

SE, standard error; CI, confidence interval. Refer to the building progression of the model in [Supplementary Material](#). The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.

TABLE 7 Estimated coefficients for efficiency score in the visual search task.

Fixed effects – predictors	Estimates	SE	CI	<i>t</i>	<i>p</i>
Intercept (KG, target present, set size 8)	0.92	0.00	0.91 – 0.92	224.59	<0.001
Set size (32)	−0.03	0.00	−0.03 – −0.02	−11.26	<0.001
Target (absent)	0.08	0.00	0.07 – 0.08	16.96	<0.001
Grade (2nd)	0.02	0.01	0.01 – 0.04	4.60	<0.001
Grade (4th)	0.04	0.01	0.03 – 0.05	7.69	<0.001
Grade (6th)	0.05	0.01	0.04 – 0.06	9.44	<0.001
Grade (college)	0.06	0.01	0.05 – 0.07	11.92	<0.001
Setsize (32) × Target (absent)	0.03	0.00	0.02 – 0.04	8.95	<0.001
Target (absent) × Grade (2nd)	−0.02	0.01	−0.04 – −0.01	−4.28	<0.001
Target (absent) × Grade (4th)	−0.04	0.01	−0.05 – −0.03	−6.54	<0.001
Target (absent) × Grade (6th)	−0.05	0.01	−0.06 – −0.03	−8.22	<0.001
Target (absent) × Grade (college)	−0.06	0.01	−0.07 – −0.05	−10.37	<0.001
Random effects					
σ^2			0.00086		
τ_{00} Subject			0.00016		
Intraclass correlation coefficient			0.16		
Marginal R^2 /Conditional R^2			0.523/0.599		

SE, standard error; CI, confidence interval. Neither working memory nor any of its interactions significantly improved the fit of the model, and they were not included. Refer to the building progression of the model in [Supplementary Material](#). The values indicate that $p < 0.05$ and therefore, the effect to which they refer is significant.

low-load conditions. For the *Guess What* subtest, the higher scores were also related to higher efficiency but only for high-load conditions, and for the low-load under set size 8. *Picture span* scalar score was only correlated with efficiency in high-load for set size 8. Finally, we did not find significant correlations between efficiency score and BRIEF and BASC *T*-scores.

Metacognitive strategies change

Figures 7A,B shows the percentage of participants in each age group who used strategies during target encoding in visual search in each session (low-load and high-load). While the percentage of participants that fell into each of the three memory strategy categories (no strategy, perceptual, or verbal) did not significantly differ by group of age in the high-load session, $\chi^2(8, N = 134) = 10.6, p = 0.225$, they did differ for the low-load, $\chi^2(8, N = 138) = 19.53, p = 0.012, V = 0.27$. As shown in **Figure 7A**, the

percentage of participants who did not use strategies decreased with age, especially for the 6th and college groups.

Considering the visual tracking strategies, the relationship between the age groups and the use of an organized search strategy was significant, both in the low-load session [$\chi^2(4, N = 140) = 26.99, p > 0.0001, V = 0.44$] and in the high-load session [$\chi^2(4, N = 136) = 28.02, p > 0.0001, V = 0.45$]. As shown in **Figures 7C,D**, the trend is similar for both working memory conditions: The percentage of participants who used tracking strategies increased with age. There was only a small percentage of children who used these strategies, while in the college students, the percentage was close to 50%.

Discussion

Visual search is present in many daily activities at all educational levels. They are often not performed in isolation,

TABLE 8 Partial correlations matrix (controlling for age in months) between efficiency scores in each of the eight conditions resulting from combining the 3 experimental variables in visual search (working memory load, set size, and target) with T-scores from RIST and scalar scores from Picture Span (WISC-V).

		Efficiency score							
		LL, TP, S8	LL, TP, S32	LL, TA, S8	LL, TA, S32	HL, TP, S8	HL, TP, S32	HL, TA, S8	HL, TA, S32
Guess what – RIST subtest	Pearson's <i>r</i>	0.213	0.131	0.051	0.024	0.223	0.208	−0.04	0.04
	<i>p</i> -value	0.01	0.119	0.544	0.774	0.007	0.012	0.632	0.632
Odd-item out – RIST subtest	Pearson's <i>r</i>	0.315	0.213	−0.036	0.038	0.119	0.226	−0.028	−0.016
	<i>p</i> -value	<0.0001	0.011	0.665	0.648	0.157	0.006	0.74	0.85
RIST general index	Pearson's <i>r</i>	0.327	0.205	0.021	0.04	0.228	0.27	−0.037	0.021
	<i>p</i> -value	<0.0001	0.014	0.805	0.632	0.006	0.001	0.661	0.798
Picture span – WISC-V	Pearson's <i>r</i>	0.159	0.119	0.024	−0.095	0.228	0.142	0.018	0.08
	<i>p</i> -value	0.095	0.212	0.802	0.318	0.015	0.135	0.849	0.401

LL, low working memory load; HL, high working memory load; TP, target present; TA, target absent; S8, set size 8; S32, set size 32. Correlations controlled by age (in months). The values indicate that $p < 0.05$ and the accompanying correlation values are significant. Therefore, the effect to which they refer is significant.

and working memory plays a critical role in most visual search tasks. This study aimed to better understand working memory modulations in visual search from a developmental perspective based on the Model of Endogenous Mental Attention, which is part of the Theory of Constructive Operators (Pascual-Leone and Johnson, 2005, 2021). Our study is the first cross-sectional one from childhood to adulthood investigating the effects of a concurrent memory load in a dual-task paradigm in visual search from a developmental perspective. Since to our knowledge this is the first study manipulating how working memory load can modulate visual search during childhood, we avoided any potential overlapping of images between working memory and visual search tasks. Under these circumstances, our results essentially show that the efficiency of visual search is not affected by the concurrent working memory load at any age from kindergarten to college, replicating previous findings in adulthood (Downing and Dodds, 2004; Woodman and Luck, 2007; Quirós-Godoy et al., 2017). Surprisingly, even for our youngest children, the load of a concurrent visual working memory task did not make an essential difference in visual search efficiency. However, the different use of strategies in different educational stages, individual difference results, and lack of working memory modulation in visual search efficiency provide novel insights into understanding working memory and visual search interactions during development that we discuss by following the Theory of Constructive Operators.

According to the Model of Endogenous Mental Attention, the M-operator keeps the figurative scheme of the target hyper-activated to “scan and find” operative schemes during visual search. The pirate chest is always shown in the center of the screen in our task and remains active during the task, that is, visible on the screen, and therefore also in the Focus of Mental Attention. Thus, along with the distractors, they are perceptually available without consuming the M-Capacity. The results, mainly related to omissions, support this assumption: the M-operator was not overloaded because of a high load in working memory. That way, even our KG children have enough resources to efficiently perform the task with an M-capacity of $2+e$.

We did not find memory effects either on the target factor or the set size factor, replicating previous findings in adulthood again (Downing and Dodds, 2004; Woodman and Luck, 2007; Quirós-Godoy et al., 2017). The interaction between target and set size is consistent with previous developmental studies (Donnelly et al., 2007; Michael et al., 2013; Woods et al., 2013; Brennan et al., 2017; Gil-Gómez de Liaño et al., 2020). However, the lack of modulation of working memory loads in visual search efficiency does not mean that there is no cost of a concurrent task for the visual performance. We found a slight increase in the probability of committing errors, specifically false alarms, and commissions under high load conditions. Former studies have not found such results. Indeed, other research paradigms studying the effects of working memory load on visual search



using dual-task paradigms essentially used detection but not localization visual search tasks with relative ceiling effects for correct responses basing the results on RTs. They could not study false alarms and commissions as we did.

The Model of Endogenous Mental Attention can account for this interesting modulation of working memory load on false alarms and commissions. Indeed, the I-operator could fail to effortfully deactivate the schemes of the working memory items during visual search in the high-load condition. These schemes are unnecessary for visual search but remain active within the focus of Mental Attention and cause interferences in the form of false alarms and commissions. However, we must be cautious since the percentage of these errors was very small (less than 0.012% of all the observations), leading to a potential lack of power. Actually, we found no age-group modulations, probably as we did not have errors enough to find reliable differences for the groups. However, all together, they show sufficient power to lead to significance when analyzed for the whole sample. These results are consistent with previous studies on effortful cognitive inhibition (Howard et al., 2014). This inhibition is related to the capacity of inactivating schemes that were activated previously.

Howard et al. (2014) also found no age differences in some cases, and if there were, they could be attributed to diverse strategies used by children and adults.

Also, the results found for RTs show working memory modulations: the youngest children needed more time in high-load conditions, especially the KG group showing a tendency to spend more time looking for the target than the 2nd-grade children who were less affected. The trend completely disappeared in the 4th grade. However, the RT differences might have come from speed differences in information processing rather than attentional changes, as we found no slope effects for age either. The myelination process at the neuronal level is under maturation from age 5–6 to adulthood, specifically in the prefrontal cortex (Kolb et al., 2012; Miller et al., 2012).

Many studies on visual search with children suggest that executive functions might be the base of most differences in development in those tasks (e.g., Woods et al., 2013). Looking for empirical data supporting this assumption, we tested whether better executive function performance was related to better efficiency in the visual search task regardless of age. However, we did not find significant correlations between

visual search efficiency and any variable from the BASC and BRIEF questionnaires. We only found an exceptional positive correlation between visual search efficiency and performance in the working memory span from the WISC-V test, similar to Guilbert et al. (2020). Part of these insignificant effects could be due to the homogeneity of the sample. All extreme values in the neuropsychological tests were eliminated as part of the inclusion criteria, involving only typical development participants.

However, we found clear evidence of a positive relationship between visual search efficiency and IQ levels, consistent with the previous literature (Navarro et al., 2006; Howard et al., 2013). The higher the IQ, the higher the visual search efficiency but only when the target was present. The low variance results across participants could explain the lack of modulation for target-absent conditions. Although high scores on the verbal and no verbal subtests of the RIST were related to better performance, it seems more pronounced for the verbal one. Certainly, language is part of numerous cognitive processes, and it is critical in the so-called *Metacognitive Knowledge*, that is, the knowledge of higher cognitive process functioning (Flavell, 1979; Efklides, 2008). Language allows for us to describe the task we perform in terms of execution, factors that can improve or worsen its achievement, or how and when to use strategies efficiently. Indeed, Vygotsky (1962) claimed that language has a critical role in developing cognitive abilities. Then, higher verbal performance could be related to better *Metacognitive Knowledge* functioning in the visual search task.

Finally, we studied how diverse strategies might affect the consecution of dual tasks. The results show an increase in the percentage of participants who used visual-tracking organization strategies with age. While the KG children barely used visual tracking strategies, 50% of the college participants used them. Reading skills (up-down; left-right) might help develop these search strategies, consistent with the increase in using them found with age (Ólafsdóttir et al., 2021). Also, from the Model of Endogenous Mental Attention, an increase in the number of schemes that the M-operator can hyperactivate at once is related to age (Pascual-Leone and Johnson, 2005, 2021; Arsalidou et al., 2019). Therefore, older groups have a greater capacity to hyperactivate and use strategies' executive schemes, which are closely related to the speed of information processing increasing with age to manipulate more information in the brain (Fry and Hale, 2000). More free resources release the capacity to create cognitive strategies. Note that the instructions before the experimental task did not explicitly suggest any strategy. However, the participants can generate them at will. Children from 4th grade onward can use strategies, but younger children cannot develop them independently (Pérez and Capilla, 2008; Rossignoli-Palomeque et al., 2020), although we can train them (Rossignoli-Palomeque et al., 2019).

For memory strategies (remember, Figures 7A,B), we only found differences for the low working memory load condition.

The youngest children (KG and 2nd) reported using (or not) strategies in a similar proportion. However, the older children and college students' proportion using memory strategies was higher than those who did not use any strategy. Fewer free cognitive resources to use strategies in the high-load conditions could explain the lack of modulations found for this more demanding condition.

However, we must consider that the way we studied strategies in the present work could also be related to the awareness of using them rather than using those strategies themselves. Maybe younger children are using these or other types of strategies to perform a task, but they simply are not aware of it or do not know how to explain their use verbally. This is consistent with the previous idea about the development of metacognitive knowledge dependent on language and other reasoning capacities under maturation during the early ages we tested here (Flavell, 1979; Efklides, 2008).

Notably, the present study's findings have critical implications in educational contexts. One of the most relevant results is that even young children can perform a pretty inefficient object visual search task under high loads of working memory. However, we should adapt the time needed to complete the task successfully, and it depends on their stage of development in light of the present results. This is especially relevant in the case of the first years of formal education (elementary school) when time adaptations are essential for achievement.

Furthermore, regardless of age, all children will be more likely to commit more errors under high working memory loads, especially coming from distractors in a visual search. Therefore, we can try to reduce both working memory load and distractions/distractors by adaptations in activities in the classroom. Some strategies that a teacher can use in the classroom would be to let students write information to remember on paper. They may also choose to make information irrelevant to visual search but essential to other concurrent tasks available using pictograms and pictures or in writing. This way, students can refer to it when they need it without overloading their working memory. Moreover, these findings can be the basis for designing strategy training programs that will help children to perform search more efficiently, especially in the case of younger children who are not capable of self-generating themselves.

These recommendations are particularly important in children with executive function impairments, i.e., in children with language (Im-Bolter et al., 2006; Pascual-Leone and Johnson, 2021) or reading (Koltermann et al., 2020) impairments, arithmetic learning disabilities (Abreu-Mendoza et al., 2018), or attention deficit disorder with hyperactivity (Koltermann et al., 2020; Soto et al., 2021). The Theory of Constructive Operators and, specifically,

the Model of Endogenous Mental Attention (Pascual-Leone and Johnson, 2005, 2021) have proven to be valid models to explain higher cognitive functioning and its development, as we have seen here, and for visual search. It would be helpful to explore their applications in tasks like those we used here on samples with atypical development or executive function problems (e.g., ADHD, dyslexia, specific language impairment, or learning disabilities).

Although this first study sheds some light on the role of working memory in visual search throughout childhood under a dual-task paradigm, additional research is needed to explore other relationships between the two processes. As we mentioned before, we did not manipulate any relationship between working memory contents and distractors/target in the visual search task like other studies have done in adulthood (e.g., Downing and Dodds, 2004; Gil-Gómez de Liaño et al., 2011). Allowing for the overlapping of stimuli between both tasks would be an interesting manipulation to test during childhood as both working memory and attention are under development between 5 and 12 years old (e.g., Brooking et al., 2012; Gil-Gómez de Liaño et al., 2020). It would also be interesting to study how search strategies like those used by Smilek et al. (2006) would apply to a developing brain.

Conclusion

In summary, our study provides an explanation of visual working memory load modulation in the visual search task based on the neurodevelopmental perspective of the Model of Endogenous Mental Attention, replicating previous results in adulthood with some new effects. Moreover, our findings contribute to a better understanding of the use of metacognitive strategies during the development of visual search and the role of the individual differences that could mediate performance. The results provide some hints to consider in the educational context.

Data availability statement

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

Ethics statement

The study was approved by the Ethics Committee of Universidad Autónoma de Madrid (CEI-84-1553). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

MQ-G, BG-GL, and EP-H conceived and planned the design, performed the data analysis, and wrote the manuscript. MQ-G programmed the experiments and carried out the data collection. All authors contributed to the article and approved the submitted version.

Funding

The present research has been supported by Research Grant FPI-UAM 2016 (Universidad Autónoma de Madrid) given to MQ-G, as well as Research Grant Project PSI2015-69358-R ("Ministerio de Economía y Competitividad de España," MINECO, and "Fondo Europeo de Desarrollo Regional," FEDER) given to BG-GL as PI.

Acknowledgments

This project would not have been possible without all the children, families, teachers, counselors, and principals who collaborated with us from Corazón Inmaculado School (Madrid, Spain), Alborada School (Alcalá de Henares, Spain), and Los Ángeles School (Pozuelo de Alarcón, Spain). We would also like to thank Eduardo García-Garzón and Ricardo Olmos Albacete for their help during the data analysis.

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.2022.907121/full#supplementary-material>

References

- Abreu-Mendoza, R. A., Chamorro, Y., Garcia-Barrera, M. A., and Matute, E. (2018). The contributions of executive functions to mathematical learning difficulties and mathematical talent during adolescence. *PLoS One* 13:e0209267. doi: 10.1371/journal.pone.0209267
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724
- Arsalidou, M., Pascual-Leone, J., and Johnson, J. (2019). The constructive operators of the working mind: a developmental account of mental-attentional capacity. *Russ. J. Cogn. Sci.* 6, 44–55.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2
- Baddeley, A. D., and Hitch, G. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1
- Bewick, K. C., Raymond, M. J., Malia, K. B., and Bennett, T. L. (1995). Metacognition as the ultimate executive: techniques and tasks to facilitate executive functions. *NeuroRehabilitation* 5, 367–375. doi: 10.3233/NRE-1995-5411
- Brady, T. F., Konkle, T., Alvarez, G. A., and Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14325–14329. doi: 10.1073/pnas.0803390105
- Brennan, A. A., Bruderer, A. J., Liu-Ambrose, T., Handy, T. C., and Enns, J. T. (2017). Lifespan changes in attention revisited: everyday visual search. *Can. J. Exp. Psychol.* 71, 160–171. doi: 10.1037/cep0000130
- Brooking, L., Uehara, E., Charchat-Fichman, H., and Landeira-Fernandez, J. (2012). Memory performance in Brazilian school-age children. *Psychol. Neurosci.* 5, 165–173. doi: 10.3922/j.psns.2012.2.06
- Bruyer, R., and Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychol. Belgica* 51, 5–13. doi: 10.5334/pb-51-1-5
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi: 10.1037/0033-2909.104.2.163
- Donnelly, N., Cave, K., Greenway, R., Hadwin, J. A., Stevenson, J., and Sonuga-Barke, E. (2007). Visual search in children and adults: top-down and bottom-up mechanisms. *Q. J. Exp. Psychol.* 60, 120–136. doi: 10.1080/17470210600625362
- Downing, P., and Dodds, C. (2004). Competition in visual working memory for control of search. *Vis. Cogn.* 11, 689–703. doi: 10.1080/13506280344000446
- Efklides, A. (2008). Metacognition. *Eur. Psychol.* 13, 277–287. doi: 10.1027/1016-9040.13.4.277
- Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. Thousand Oaks, CA: SAGE Publications.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am. Psychol.* 34, 906–911. doi: 10.1037/0003-066X.34.10.906
- Fry, A. F., and Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biol. Psychol.* 54, 1–34. doi: 10.1016/S0301-0511(00)00051-X
- Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, MA: University Press.
- Gerhardstein, P., and Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *J. Exp. Child Psychol.* 81, 194–215. doi: 10.1006/jecp.2001.2649
- Gil-Gómez de Liaño, B., Botella, J., and Pascual-Ezama, D. (2011). The types of stimuli loaded in memory can modulate its effects on visual search. *J. Cogn. Psychol.* 23, 531–542. doi: 10.1080/20445911.2011.542747
- Gil-Gómez de Liaño, B., Potter, M. C., and Rodríguez, C. (2014). Working memory effects in speeded RSVP tasks. *Psychol. Res.* 78, 124–135. doi: 10.1007/s00426-013-0479-7
- Gil-Gómez de Liaño, B., Quirós-Godoy, M., Pérez-Hernández, E., and Wolfe, J. M. (2020). Efficiency and accuracy of visual search develop at different rates from early childhood through early adulthood. *Psychon. Bull. Rev.* 27, 504–511. doi: 10.3758/s13423-020-01712-z
- Gil-Gómez de Liaño, B., Stablum, F., and Umiltà, C. (2016). Can concurrent memory load reduce distraction? A replication study and beyond. *J. Exp. Psychol. Gen.* 145, e1–e12. doi: 10.1037/xge0000131
- Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2015). *BRIEF-2: Behavior Rating Inventory of Executive Function*, Second Edn. Lutz, FL: PAR Inc.
- Guilbert, A., Perguilhem, S., and Guiraud-Vinatea, H. (2020). Visual search strategies in children: a reflection of working memory processes? *J. Clin. Exp. Neuropsychol.* 42, 975–981. doi: 10.1080/13803395.2020.1830033
- Henry, L. A., Messer, D., Luger-Klein, S., and Crane, L. (2012). Phonological, visual, and semantic coding strategies and children's short-term picture memory span. *Q. J. Exp. Psychol.* 65, 2033–2053. doi: 10.1080/17470218.2012.672997
- Hommel, B., Li, K. Z. H., and Li, S. C. (2004). Visual search across the life span. *Dev. Psychol.* 40, 545–558. doi: 10.1037/0012-1649.40.4.545
- Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified bonferroni test. *Biometrika* 75, 383–386.
- Howard, S. J., Johnson, J., and Pascual-Leone, J. (2013). Measurement of mental attention: assessing a cognitive component underlying performance on standardized intelligence tests. *Psychol. Test Assess. Model.* 55, 250–273.
- Howard, S. J., Johnson, J., and Pascual-Leone, J. (2014). Clarifying inhibitory control: diversity and development of attentional inhibition. *Cogn. Dev.* 31, 1–21. doi: 10.1016/j.cogdev.2014.03.001
- IBM Corp (2019). *IBM SPSS Statistics for Windows, Version 26.0*. Armonk, NY: IBM Corp.
- Im-Bolter, N., Johnson, J., and Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: the role of executive function. *Child Dev.* 77, 1822–1841. doi: 10.1111/j.1467-8624.2006.00976.x
- Karmiloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: The MIT Press.
- Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D. O., and Gibb, R. (2012). Experience and the developing prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109(Suppl. 2), 17186–17193. doi: 10.1073/pnas.1121251109
- Koltermann, G., Becker, N., Lopes-Silva, J. B., Gomides, M. R. A., Paiva, G. M., Haase, V. G., et al. (2020). Are "cool" executive function impairments more salient in ADHD symptoms than in reading disability? *Dement. Neuropsychol.* 14, 47–55. doi: 10.1590/1980-57642020dn14-010008
- Lavie, N., and de Fockert, J. (2005). The role of working memory in attentional capture. *Psychon. Bull. Rev.* 12, 669–674. doi: 10.3758/BF03196756
- Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (2004). Load theory of selective attention and cognitive control. *J. Exp. Psychol. Gen.* 133, 339–354. doi: 10.1037/0096-3445.133.3.339
- Luciana, M., and Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia* 36, 273–293. doi: 10.1016/S0028-3932(97)00109-7
- Michael, G. A., Lété, B., and Ducrot, S. (2013). Trajectories of attentional development: an exploration with the master activation map model. *Dev. Psychol.* 49, 615–631. doi: 10.1037/a0028410
- Miller, D. J., Duka, T., Stimpson, C. D., Schapiro, S. J., Baze, W. B., McArthur, M. J., et al. (2012). Prolonged myelination in human neocortical evolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16480–16485. doi: 10.1073/pnas.1117943109
- Navarro, J. I., Ramiro, P., Lopez, J. M., Aguilar, M., Acosta, M., and Montero, J. (2006). Mental attention in gifted and nongifted children. *Eur. J. Psychol. Educ.* 21, 401–411. doi: 10.1007/BF03173510
- Ólafsdóttir, I. M., Gestsdóttir, S., and Kristjánsson, Á (2021). The development of foraging organization. *Attent. Percept. Psychophys.* 83, 2891–2904. doi: 10.3758/s13414-021-02328-8
- Olivers, C. N. L., Peters, J., Houtkamp, R., and Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334. doi: 10.1016/j.tics.2011.05.004
- Palmer, S. (2000). Working memory: a developmental study of phonological recoding. *Memory* 8, 179–193. doi: 10.1080/096582100387597
- Paloş, R., Munteanu, A., Costea, I., and Macsinga, I. (2011). Motivational and cognitive variables with impact on academic performance Preliminary study. *Proc. Soc. Behav. Sci.* 15, 138–142. doi: 10.1016/j.sbspro.2011.03.063
- Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychol.* 32, 301–345. doi: 10.1016/0001-6918(70)90108-3
- Pascual-Leone, J. (1995). Learning and development as dialectical factors in cognitive growth. *Hum. Dev.* 38, 338–348. doi: 10.1159/000278340
- Pascual-Leone, J. (2000). Reflections on working memory: are the two models complementary? *J. Exp. Child Psychol.* 77, 138–154. doi: 10.1006/jecp.2000.2593

- Pascual-Leone, J., and Johnson, J. (1999). *Methods of Task Analysis in Cognitive Development*. Toronto, ON: York University.
- Pascual-Leone, J., and Johnson, J. (2005). "A dialectical constructivist view of developmental intelligence," in *Handbook of Understanding and Measuring Intelligence*, eds O. Wilhelm and R. W. Engle (Thousand Oaks, CA: SAGE Publications, Inc).
- Pascual-Leone, J., and Johnson, J. M. (2021). *The Working Mind: Meaning and Mental Attention in Human Development*. Cambridge, MA: The MIT Press.
- Pérez, E., and Capilla, A. (2008). "Neuropsicología infantil," in *Manual de Neuropsicología*, eds J. Tirapu, M. Ríos, and F. Maestú (Madrid: Viguera).
- Pintrich, P. R., and de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *J. Educ. Psychol.* 82, 33–40. doi: 10.1037/0022-0663.82.1.33
- Psychology Software Tools, Inc. (2016). *E-Prime 3*. Sharpsburg, PA: Psychology Software Tools, Inc.
- Quirós-Godoy, M., Botella, J., and Gil-Gómez de Liaño, B. (2017). A concurrent working memory load task does not impair visual selective attention: a meta-analysis. *J. Vis.* 17:967. doi: 10.1167/17.10.967
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reynolds, C. R., and Kamphaus, R. W. (2003). *RIST: Reynolds intellectual screening test – Interpretative Manual*. Torrance, CA: Western Psychological Services.
- Reynolds, C. R., and Kamphaus, R. W. (2004). *BASC: Behavior Assessment System for Children*, 2nd Edn. Circle Pines, MN: American Guidance Service.
- Roncadin, C., Pascual-Leone, J., Rich, J. B., and Dennis, M. (2007). Developmental relations between working memory and inhibitory control. *J. Int. Neuropsychol. Soc.* 13, 59–67. doi: 10.1017/S1355617707070099
- Rossignoli-Palomeque, T., Perez-Hernandez, E., and González-Marqués, J. (2020). Training effects of attention and EF strategy-based training "Nexxo" in school-age students. *Acta Psychol.* 210:103174. doi: 10.1016/j.actpsy.2020.103174
- Rossignoli-Palomeque, T., Quiros-Godoy, M., Perez-Hernandez, E., and González-Marqués, J. (2019). Schoolchildren's compensatory strategies and skills in relation to attention and executive function app training. *Front. Psychol.* 10:2332. doi: 10.3389/fpsyg.2019.02332
- Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *J. Exp. Child Psychol.* 111, 695–707. doi: 10.1016/j.jecp.2011.10.007
- Smilek, D., Enns, J. T., Eastwood, J. D., and Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Vis. Cogn.* 14, 543–564. doi: 10.1080/13506280500193487
- Soto, D., Hodsoll, J., Rotshtein, P., and Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends Cogn. Sci.* 12, 342–348. doi: 10.1016/j.tics.2008.05.007
- Soto, E. F., Irwin, L. N., Chan, E. S. M., Spiegel, J. A., and Kofler, M. J. (2021). Executive functions and writing skills in children with and without ADHD. *Neuropsychology* 35, 792–808. doi: 10.1037/neu0000769
- Townsend, J. T., and Ashby, F. G. (1983). *Stochastic Modelling of Elementary Psychological Processes*. New York, NY: Cambridge University Press.
- Vogel, E. K., Woodman, G. F., and Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 92–114. doi: 10.1037/0096-1523.27.1.92
- Vuontela, V., Carlson, S., Troberg, A. M., Fontell, T., Simola, P., Saarinen, S., et al. (2013). Working memory, attention, inhibition, and their relation to adaptive functioning and behavioral/emotional symptoms in school-aged children. *Child Psychiatry Hum. Dev.* 44, 105–122. doi: 10.1007/s10578-012-0313-2
- Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, MA: MIT Press.
- Wechsler, D., Rainford, S., and Holdnack, J. (2015). *WISC-V: Escala de inteligencia de Wechsler para niños-V*. Madrid: Pearson PsychCorp.
- Wolfe, J. M. (2021). Guided Search 6.0: an updated model of visual search. *Psychon. Bull. Rev.* 28, 1060–1092. doi: 10.3758/s13423-020-01859-9
- Woodman, G. F., and Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *J. Exp. Psychol.* 33, 363–377. doi: 10.1037/0096-1523.33.2.363
- Woods, A. J., Göksun, T., Chatterjee, A., Zelonis, S., Mehta, A., and Smith, S. E. (2013). The development of organized visual search. *Acta Psychol.* 143, 191–199. doi: 10.1016/j.actpsy.2013.03.008



Development of an Online Multilingual Educational Programme for Parents of Dual-Career Athletes: A Participatory Design

Laura Capranica^{1,2*}, Flavia Guidotti¹, Carlos Gonçalves³, Laurence Blondel⁴, Matteo Bovis⁵, Rute Costa⁶, Nadine Debois⁴, Antonio Figueiredo³, Ciaran MacDonncha^{7,8}, Viktorija Pecnikar-Oblak⁹, Jean-Luc Patoret⁴, Andrej Pišl¹⁰, Eoin Rheinisch¹¹, Ana Rolo⁶, Gary Ryan^{7,8}, Anne Templet⁴, Antonio Tessitore¹, Giles Warrington^{7,8} and Mojca Doupona⁹

¹ Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome, Italy, ² European Athlete as Student, European Athlete as Student (EAS) Network, Ghaxaq, Malta, ³ Faculty of Sport Science and Physical Education, University of Coimbra, Coimbra, Portugal, ⁴ National Institute of Sport, Expertise and Performance (INSEP), Paris, France, ⁵ CONI, Rome, Italy, ⁶ Ginásio Clube Figueirense, Figueira da Foz, Portugal, ⁷ Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland, ⁸ Health Research Institute, University of Limerick, Limerick, Ireland, ⁹ Department of Sport Sociology, Faculty of Sport, University of Ljubljana, Ljubljana, Slovenia, ¹⁰ EUSA Institute, Ljubljana, Slovenia, ¹¹ Sport Ireland Institute, Dublin, Ireland

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the
Americas—Nevis, United States

Reviewed by:

Corrado Lupo,
University of Turin, Italy
Cristina O. Mosso,
University of Turin, Italy

*Correspondence:

Laura Capranica
laura.capranica@uniroma4.it

Specialty section:

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

Received: 15 January 2022

Accepted: 16 June 2022

Published: 22 July 2022

Citation:

Capranica L, Guidotti F, Gonçalves C, Blondel L, Bovis M, Costa R, Debois N, Figueiredo A, MacDonncha C, Pecnikar-Oblak V, Patoret J-L, Pišl A, Rheinisch E, Rolo A, Ryan G, Templet A, Tessitore A, Warrington G and Doupona M (2022) Development of an Online Multilingual Educational Programme for Parents of Dual-Career Athletes: A Participatory Design. *Front. Psychol.* 13:855531. doi: 10.3389/fpsyg.2022.855531

There is a need for educational support structures to empower parents in sustaining talented athletes pursuing academic and sports careers (e. g., dual career). The present work describes the participatory design used to develop a series of educational resources and the subsequent iterations used to evaluate the content of the EMPATIA online education programme for parents of dual-career athletes. Following an ethnographic approach, the project team (18 dual-career experts) engaged in four iterations (i.e., rounds) planned to develop (rounds 1 and 2) educational material from preliminary evidence (systematic literature review) and eminence (focus groups and concept mapping) knowledge and to evaluate the educational programme (round 3 and 4) engaging end users ($n = 76$) and other stakeholders (9 dual-career experts). The EMPATIA programme was developed and organized in four modules labeled after macro-aspects, parents could ask about dual career: "Why" (the definition and challenges of dual career for athletes and their parents), "What" (insights, guidelines, and suggestions on the parental role in support of student-athletes), "How" (practical advice on planning dual career at sports and academic levels, and post-sports careers), and "Where" (finding legal information or counseling on dual career). Despite overall positive evaluations of the programme, parents of dual-career athletes attributed higher scores with respect to those of dual-career experts. The participatory approach presented in this work enables developers to apply a systematic and multidisciplinary approach toward the creation of educational programmes for parents. The cooperation among dual-career researchers, experts from high-performance centers, Olympic bodies, sports clubs, and parents of elite student-athletes of different sports and nationalities created an educational programme suitable for end users to support parenting athletes in combining their sports and academic careers.

Keywords: dual career, online education, parenting athletes, student-athletes, experts

INTRODUCTION

In Europe, 120,000 talented young European athletes are estimated to dedicate several years to achieving elite level in sports (European Commission, 2016). In general, a sports career starts at a young age, lasts into adulthood, and during its lifespan encompasses increasing volume, intensity, and organization of training and competitions organized at national and international levels. Parallel to sports careers, academic careers significantly increase the educational demands from mandatory education (e.g., primary and high school) to university. In considering that in the European Member States competitive sports are mainly structured at club level, talented and elite athletes are at risk of academic or sport drop outs when educational and sport systems present cultural and organizational divergences (Aquilina and Henry, 2010; European Commission, 2016). In fact, a sports engagement of around 20 h per week⁻¹ and an academic commitment of around 30 h per week⁻¹ have been reported (Sheehan et al., 2018; Condello et al., 2019). To support the athletes' right to pursue their holistic development through the conciliation of their sports and academic careers, in 2007, the European white paper on sports introduced the term "dual career," which is considered one of the priorities in European sports strategy and policy (European Commission, 2007, 2012, 2016, 2022; European Parliament, 2015, 2017, 2021).

From various perspectives, numerous studies have examined the factors that influence dual careers (Guidotti et al., 2015; Stambulova and Wylleman, 2019). Although, sportspersons are mainly responsible for their dual careers, several stakeholders have been identified as influencing the micro—(e.g., the individual athlete), the meso—(e.g., parents, peers, teachers/employers, coaches, sports managers), the macro—(e.g., sports clubs/federations, educational institutions, and labor market), and the policy (e.g., national and European governing bodies) levels of dual-career paths (European Commission, 2007, 2012; Guidotti et al., 2015; Capranica and Guidotti, 2016; Stambulova and Wylleman, 2019; Stambulova et al., 2021).

Actually, Member States adopt different approaches to dual careers, which determine a wide range of support measures and services also in relation to the sports-specific and education/work-specific needs of athletes (Aquilina and Henry, 2010; Capranica and Guidotti, 2016; European Commission, 2016). Independent from nationality and sport, at a personal level, elite athletes highlighted the crucial role of their parents, who provide emotional, motivational, instrumental, material, social, informational, and financial support (Condello et al., 2019). However, the personal experiences and opinions of parents regarding their supportive role of their dual-career children have been presented, and the complex parenting demands at individual and interindividual levels in relation to the family, sport, and academic environments have been systematized recently (Harwood and Knight, 2009, 2015; Harwood et al., 2010; Bean et al., 2014; Clarke and Harwood, 2014; Dorsch et al., 2017; Gjaka et al., 2021; Tessitore et al., 2021; Varga et al., 2021). To help parents facilitate relationships with dual-career stakeholders of family, sport, and educational entourages, the elaboration of a specific education programme appeared as a strategic

necessity (Thrower et al., 2016; Capranica et al., 2018; Gjaka et al., 2021). With the support of the European ERASMUS+ Programme (<https://erasmus-plus.ec.europa.eu>), a consortium of 10 universities and sports institutions from six Member States (Ireland, Italy, France, Malta, Portugal, and Slovenia) having a consolidated relationship at national and European levels in dual career aimed to structure an Education Model for Parents of Athletes in Academics (EMPATIA, 590437-EPP-1-2017-1-SI-SPO-SCP).

To understand the values, beliefs, behaviors, and needs of parents of dual-career athletes from different countries and sports environments, from an epistemological point of view, the adoption of an ethnographic stance appears plausible to gain a comprehensive understanding of the group members' perceptions, as they shaped by their social and cultural settings. Therefore, the multidisciplinary EMPATIA team used an ethnographic research approach, which helps to describe a group or culture, strongly relying on personal experiences to make additional decisions about which approaches are appropriate for the situation at hand (Genzuck, 2003). Thus, a participatory approach was deemed a key feature for ensuring an active engagement of parents in the developing process of a standalone digital educational resource, self-directed in nature, freely and openly available in the form of reusable learning objects, and based on the small, independent, reusable, aggregation-ready educational units (McGreal, 2004). Indeed, a participatory design could ensure that a programme having usable, simple, and intelligible contents, aligned to end-users' needs, could ultimately empower the people involved (Ferguson et al., 2018; Hansen et al., 2019). According to the ethnographic research approach, to gather a comprehensive understanding of parenting dual-career athletes through the key aspects, their structure, and interrelationships, a variety of information gathered from different perspectives is crucial (Genzuck, 2003). First, an analysis of documents encompassed a systematic literature review, which highlighted fragmented information presenting a two-level construct of individual and interindividual aspects of parental support strategies of dual-career athletes (Tessitore et al., 2021). Second, to collect a variety of information on the personal experience of parents from different perspectives, the research included 12 focus groups involving a total of 115 parents of elite dual-career athletes, who discussed themes related to the athletes' needs, sports environment, academic environment, dual-career-related policies and services, and educational methods for parenting dual-career athletes (Gjaka et al., 2021). Thus, the parents condensed into 80 statements their perceived needs and opinions on the most relevant content of an educational programme for parenting dual-career athletes, mainly related to the parental support athletes' needs, the parents' relationships with the sports and academic entourages of the athletes, the parents' need of information on dual-career policies/services, and educational resources for parenting dual-career athletes. Finally, to uncover the cultural perspectives on which parenting dual-career athletes is based, the 80 statements were further integrated and visually and numerically represented in a concept mapping, which resulted in the composite thinking of 334 groups of parents of elite athletes as students around the complex dual-career

phenomenon (Varga et al., 2021). The findings presented a five-cluster framework on the parents' roles, needs, and awareness to support athletes; requirements for effective planning of dual-career pathway; educational opportunity; policy and provision for dual career; and athletes' lifestyle and self-management.

Thus, the main objective of this paper was to describe the participatory design that is used to develop a series of educational resources and the subsequent iterations that are used to provide formative feedback on the content of the EMPATIA education programme for parents of dual-career athletes as a part of the verification process. It was hypothesized that core information and key priorities could lead to relevant and high-quality educational materials aligned to end-users' needs and a positive and structured feedback from a selected expert panel (Genzük, 2003; Hsu and Sandford, 2007; Diamond et al., 2014).

METHODS

Experimental Approach to the Study

Based on the sound triangulation of the different data collected in the background phase presented in the introduction (Gjaka et al., 2021; Tessitore et al., 2021; Varga et al., 2021), during a 2-year period, four iterations (i.e., rounds) were considered necessary to develop the EMPATIA education programme through qualitative approaches and to reach an agreed degree of consensus through qualitative and quantitative methods (Genzük, 2003; Hsu and Sandford, 2007). In particular, the EMPATIA team was entitled to select wisely the key informants from the parents' perspectives and to use them carefully for developing a manageable online educational programme, whereas the parents of athletes as students and the dual-career experts represented the insiders to cross-evaluate that the researchers organized effectively educational material without omitting to include the essential aspects for empowering parents of dual-career athletes (Genzük, 2003).

As schematically presented in **Figure 1**, rounds 1 and 2 were dedicated to the developmental phase, whereas during the second year, a consensus was reached through rounds 3 and 4.

Round 1

A total of two face-to-face meetings among the EMPATIA team members were organized to identify the most relevant inputs deriving from the five clusters (i.e., policy and provision for dual career; parent/guardian roles, needs and awareness to support athletes' health, wellbeing, sports, and education; athletes' lifestyle and self-management; requirements for effective planning of DC pathway; and educational opportunities) of the EMPATIA framework on parenting dual-career athletes (Varga et al., 2021). This interpretive process required a consensus on three challenges: first, to move from the logic deriving from a concept mapping to a conceptual logic of interconnected constructs that could respond to the educational needs of parents; second, to build coherent pedagogical units, conceptually coherent, to be used in sequence or separately that fit the constraints of an internet page, maintaining an appealing and user-friendly structure; third, to ensure a viable, interactive internet structure for allowing a separate and easy search of specific topics of interest.

Round 2

The EMPATIA team members established the preliminary priorities among educational items and agreed the learning goals encompassed the following pedagogical components: (i) the presentation of the concept or procedure to support the learning objectives; (ii) an activity for the learner to engage with the content; (iii) self-assessment to test mastery of the content; and (iv) links to further resources to reinforce the learning. Then, a consensus on two distinct phases was required: First, the prototype is developed to obtain a full specification of the future programme. Second, the formal educational programme is provided in English, French, Italian, Portuguese, and Slovenian.

Round 3

National workshops were organized to present the EMPATIA education platform to parents of dual-career athletes. After allowing the participants to navigate individually the programme, an evaluation questionnaire including ratings and open judgments was administered. This round provided an opportunity for clarifications on the relative importance and comprehensibility of the educational items, as well as provided an opportunity to revise the educational programme, if needed.

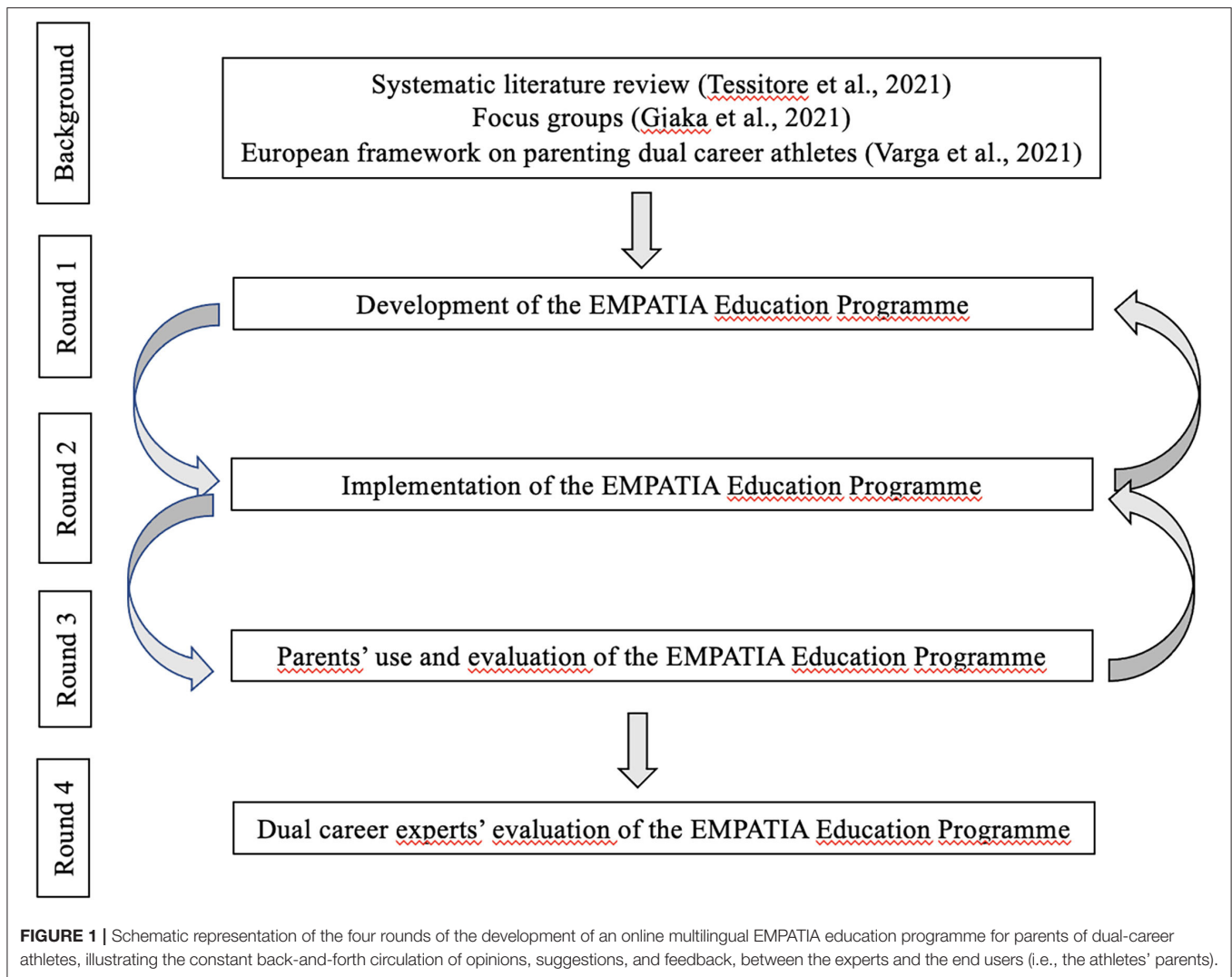
Round 4

In the final round, a workshop with international dual-career experts was organized to present the EMPATIA education platform. After allowing the participants to navigate the programme individually, an evaluation questionnaire including ratings and open judgments was administered. This round provided a final opportunity to revise the educational programme, if needed.

Participants

Three groups of well-qualified experts were identified: (1) 18 dual-career experts from three universities with sports science programmes, two high-performance centers, Olympic bodies, sports clubs, an internet designer, and all members of the EMPATIA team. This group was entitled to rounds 1 and 2; (2) a sample of 79 parents ($F = 47\%$, $M = 53\%$; education: university = 69% , high school = 31%) of dual-career athletes (range: 14–24 years) competing in team (72%) and individual (28%) sports at national (highest championships: 53%) and international (European and World competitions; 47%) levels was drawn from the national databases of parents who previously contributed to the development of the EMPATIA framework on parenting dual-career athletes (Varga et al., 2021).

This group was engaged in round 3. All potential participants received a notification email requesting to contribute to the evaluation of the EMPATIA education programme based on the results of the previous concept mapping consultation of the EMPATIA framework, the time required and the nature of the online evaluation exercise, and information that their participation was entirely voluntary, that they could withdraw at any time without providing a reason, that their consent to participate was implied by submitting the online survey, and that all responses were anonymous and confidential in nature; and (3) nine dual-career expert delegates of European institutions members of the European dual-career network



(EAS; www.dualcareer.eu). This group was engaged in round 4. All potential participants were invited to the 2021 Annual Conference of the European Athlete as Student, and they were informed about the EMPATIA project and the development of the educational platform, the time required and the nature of the online evaluation exercise, and information that their participation was entirely voluntary, that they could withdraw at any time without providing a reason, that their consent to participate was implied by submitting the online survey, and that all responses were anonymous and confidential in nature.

The Instrument

Due to a lack of a validated scale for the assessment of an educational programme for parents of dual-career athletes, an electronic and multilingual (i.e., English, French, Italian, Portuguese, and Slovenian) evaluation questionnaire was developed (**Appendix A**) to allow a time and geographic flexibilities, as well as multimedia and self-administration modalities (Callegaro et al., 2015). From an applied point of view, this self-report measurement instrument has been developed as probes and follow-up questions (Genzuck, 2003). The questionnaire included a preliminary section to gather

socio-demographic information from the sample, encompassing the sport practiced by the child, and the country, sex, and educational level of the parent. Then, 14 questions were developed to gather the parents' feedback on their first personal impression of the EMPATIA education programme (Q1–Q5), their perceived relevance of the educational modules for parental support of dual-career athletes, the management of the child's sports and academic environments and the child's dual-career transitions (Q6–Q10), their opinions regarding the texting and the visual structure, their overall evaluation score (Q11–Q13), and suggestions for further improvements (Q14). For the dual-career experts' evaluation, only the questions related to the quality of the programme (Q1–Q5, Q12–Q14) were considered.

To collect data on a large sample of a heterogeneous population, close-ended questions were chosen (e.g., single or multiple response checklist type) even though respondents were also allowed to elaborate further on their answers.

Statistical Analysis

Data were analyzed using the Statistical Package for the Social Science, version 24.0 (SPSS Inc., Chicago Illinois). Descriptive statistics was applied, and the frequency of occurrence (expressed

in absolute values or percentages) was calculated for the questions for which a single response or multiple responses were allowed. Except for neutral responses, data reduction has been conducted by collapsing very positive and positive responses, as well as negative and very negative responses. Then, the chi-square tests were applied to verify the distribution of answers between parents and dual-career experts, with a level of statistical significance set at $p < 0.05$ for all computations. Finally, the sum of positive or neutral responses $>70\%$ was considered a cutoff for agreement.

RESULTS

Developmental Phase

During the initial developmental phase, an agreement defined the overall structure of the educational programme for guiding the users in their search and interpretation of information. It was deemed crucial to address the demands of dual careers and the role of parents, and to suggest resources for parents in search of supplementary information (videos, papers, books, online sites, etc.). Then, the best-ranked statements/items included in the concept mapping of the EMPATIA framework were organized in four modules labeled after macro-aspects, and parents could ask about dual careers: “Why,” “What,” “How,” and “Where.” Finally, the online platform’s planning and design were approved to ensure an effective data indexing to enhance the search operation. The implementation phase tested the developed system regarding the expected results and the defect-free requirements of the system. Furthermore, an agreement was reached regarding the content of the four modules for quality education. In particular, the “Why” module encompassed the definition of dual career and its structure, the importance of sports and educational careers, the challenges of dual careers for athletes and their parents, information on European dual-career guidelines, and policies in place in the Member States. Furthermore, video recordings of athletes, coaches, dual-career policy officers, and parents were provided. The “What” module offered some insights, guidelines, and suggestions on the parental role to support athletes in balancing the dual demand of maintaining a sporting and academic career pathway and potential sources of stress for both the athletes and the whole family. The “How” module included practical advice on planning dual careers, lifestyles, self-management, and dual-career transitions to new environments, at sports and academic levels, and to post-sports careers. Finally, the “Where” module was meant to assist parents in finding legal information or counseling about their rights and doubts about dual careers. While offering mainly the information on specific legislation or provisional services and resources for France, Ireland, Italy, Portugal, and Slovenia, it also extended the information on dual careers in Europe, official European documents and studies, and suggested further readings.

The educational content was developed in English and translated into French, Italian, Portuguese, and Slovenian by two native experts in each language, who independently performed the forward translation and agreed on a combined version. To avoid inaccuracies in producing a conceptually and

semantic equivalent translation, an English reviewer performed a backward translation with a blind procedure (Su and Parham, 2002).

Evaluation Phases

All the evaluations of both the parents and the dual-career experts’ groups reached a consensus, with significantly higher ($97.5 \pm 6.8\%$) positive evaluations for the parents ($89.3 \pm 10.9\%$) with respect to those of their dual-career counterparts (**Table 1**). For both groups, no negative response emerged for the first impression of the programme and the relevance of the four education modules, which received very high positive responses, especially from parents (range 95–98%). To note, a lack of difference between groups emerged only for the relevance of the module WHY. While the majority of parents considered the clarity and simplicity of the text very positive (range: 93–94%) or adequate (range 3–5%), the provided information received lower scores (positive: 49%; adequate: 26%). Interesting, only for this issue, the sum of positive and neutral evaluations of dual-career experts was higher values (86%) with respect to those of parents (76%). While parents attributed high positive evaluations (91–99%) to the visual structure of the programme considered very easy to navigate and well-structured, the dual-career experts attributed their lowest positive scores (51%) and the highest frequency of negative ones (14–29%). Nonetheless, the overall score of the EMPATIA education programme was positive (parents: 95%; dual-career experts: 71%).

Also, the evaluations of the parents’ perceived relevance of the educational modules for parental support of dual-career athletes, the management of the child’s sport and academic environments, and the child’s dual-career transitions resulted in high positive evaluations regarding the information for understanding the parental role in dual career (94%) to help the athletes managing their time (89%), financial issues (88%), healthy lifestyles (93%), and nutrition (84%), and preventing doping (84%), the latter presenting one negative response. The EMPATIA education programme was perceived as very helpful for the parental role in supporting the athletes in their sports environment by means of effective strategies (88%), managing difficult conversations (86%), recognizing bad (81%), and friendly (86%) sports environments, and balancing sports commitments (95%). Similarly, very positive feedback emerged relative to the academic environment, with high percentages for effective strategies (83%), managing difficult conversations (84%), recognizing bad (88%), and friendly (90%) academic environments, and in balancing academic commitments (93%). Finally, the EMPATIA programme was considered helpful for the parental role (89%) in support of athletes during academic (86%) and sports (86%) transitions, to adapt to transitions (86%) and relocations (80%), and during injuries and post-sports careers.

DISCUSSION

To improve the usability and satisfaction of sound educational programmes for the implementation of competencies of end users, European research funding bodies encourage capacity building based on sound theory and design principles, as well

TABLE 1 | Frequency of occurrence (%) of negative, neutral, and positive evaluations of the EMPATIA programme and statistical comparisons ($p < 0.05$) between parents of dual-career athletes and European dual-career experts.

		Parents	Experts	
Questions		%	%	P
First impression of the EMPATIA programme				
	Negative	0	0	<0.001
	Neutral	0	18	
	Positive	100	82	
Relevance of the module “HOW”				
	Negative	0	0	0.001
	Neutral	4	18	
	Positive	96	82	
Relevance of the module “WHERE”				
	Negative	0	0	0.009
	Neutral	2	11	
	Positive	98	89	
Relevance of the module “WHAT”				
	Negative	0	0	<0.001
	Neutral	2	22	
	Positive	98	78	
Relevance of the module “WHY”				
	Negative	0	0	0.083
	Neutral	5	11	
	Positive	95	89	
Text of the EMPATIA programme				
Easy to understand	Negative	3	14	<0.001
	Neutral	3	29	
	Positive	94	57	
Simple language	Negative	1	14	<0.001
	Neutral	5	29	
	Positive	94	57	
Quantity of information	Negative	24	14	<0.001
	Neutral	27	43	
	Positive	49	43	
Visual structure of the EMPATIA programme				
Well-structured	Negative	0	29	<0.001
	Neutral	5	14	
	Positive	95	57	
Easy to navigate	Negative	1	29	<0.001
	Neutral	8	14	
	Positive	91	57	
Clear structure	Negative	1	14	<0.001
	Neutral	8	29	
	Positive	91	57	
Overall score of the EMPATIA programme				
	Negative	0	14	<0.001
	Positive	5	14	
	Neutral	95	71	

as end users and stakeholders viewed as co-constructors of the educational tools at the core of the development and evaluation phases (European Commission, 2022). Furthermore, the adopted ethnographic research approach allowed a comprehensive

understanding of European parents' needs as shaped by their national sociocultural dual-career settings and different sports environments (Genzuk, 2003). In line with the above principles, the main findings of this study substantiated the value of a participatory development process of the online multilingual EMPATIA education programme based on evidence and eminence knowledge on the educational needs of parents of dual-career athletes, which was positively evaluated through a process encompassing stakeholder consultations (Capranica et al., 2018; Hansen et al., 2019; Gjaka et al., 2021; Tessitore et al., 2021; Varga et al., 2021). Interesting to note, the parents of dual-career athletes (e.g., end users) appreciated very much the overall quality of the programme (i.e., Q1–Q5 and Q11–Q14), whereas dual-career experts tended to be positive, yet more moderately. This partial mismatch might highlight the parents' feelings to be considered relevant actors in need of information on dual careers, a relevant phenomenon heavily conditioning the life of their entire family. It is also possible to speculate that parenting dual-career athletes is a relatively novel issue also for dual-career experts, who might give the parental support of athletes as granted. Indeed, there is a need to establish a valuable dual-career alliance also including parents to facilitate sharing of the information and experiences on specific dual-career demands, development programmes, and processes (Gjaka et al., 2021; Tessitore et al., 2021).

The positive evaluations of parents mirror their appreciation of information strictly related to the priorities they highlighted during the collective conceptualization of parental support of dual-career athletes (Varga et al., 2021). According to the principles of participative approaches (Hansen et al., 2019), the parents were required to evaluate the aspects related to the programme's outcomes and impact, especially those related to their personal gains in supporting the dual career of their talented child. Overall, the parents perceived the educational content of the EMPATIA programme as helpful at the individual level not only for clarifying their role in supporting dual careers but also in helping the athletes manage healthy lifestyles (range 83–94%). Starting from a relevant responsibility in providing motivation, encouragement, and support toward the holistic development of their talented children, parents need to be properly informed on short- and long-term plans for sound educational and sports decisions, which could have immediate and/or longer-term implications for their children and their family. In fact, intergenerational transfer of skills, knowledge, and social ability needs time and specific competencies, so parents have to be aware and ready to fine-tune their support in relation to the different challenging aspects of dual-career pathways during the developmental years of the athletes. Indeed, dual careers have to be considered as a lifelong process aiming to ensure the wellbeing of student-athletes, which includes their health-related needs to achieve and maintain an optimal mental and physical condition, to prevent and manage injuries and burnouts, and to avoid risky behaviors reported in athletes, such as unhealthy dietary habits, illegal supplement, alcohol, drugs, bullying, depression, gambling, match-fixing, violence, and sexual harassment (Aunola et al., 2018; Knight et al., 2018; Sorkkila et al., 2020; Gjaka et al., 2021; Stambulova et al., 2021).

In addressing the emphasis, stress, and perspectives on sports and/or academic achievements, the parents considered the EMPATIA education programme also beneficial for their interindividual relationship with the athletes and other dual-career stakeholders at family, academic, and sports levels (range 81–90%). Actually, parents share the responsibility of accompanying talented athletes with a multiplicity of stakeholders (e.g., the partner, other children, peers, teachers/employers, coaches, sports managers, etc.) having direct and strong influences on dual-career paths (Capranica and Guidotti, 2016). Despite the differences in national policies, services, and sociocultural contexts at sports and academic levels have been reported to challenge especially the interindividual relationships of parents, the information gathered based on cross-national and cross-sports research on the parents' views of their actual educational needs proved to be useful for a fine identification of relevant information to be included in the EMPATIA education units (Gjaka et al., 2021; Tessitore et al., 2021; Varga et al., 2021). In particular, the parents valued the useful information on the strategies and quality of communication to manage difficult conversations and conflicts in the family, sports, and academic environments and in uncovering signs that characterize negative and positive academic and sports environments. Indeed, good and regular parents-athlete-teacher/coach communication is deemed crucial to facilitate a successful dual-career alliance as well as to foster effective parental interventions at academic and/or sports levels, which could also lessen potential stress for parents. Parents educated in supporting dual careers could be valuable human resources for both sports organizations and educational bodies to ensure a continuous communication regarding the progress of the athlete's holistic development, especially when there is a need to negotiate flexibility, tutoring, and services and to design sound educational and sport short- and long-term wellbeing programmes. In going beyond the mastery of sport-specific skills as the athlete ages, the mentorship role of a coach is especially relevant during transitions to new environments requiring a variety of contextual changes that include people, culture, interactions, and social settings (Cosh and Tully, 2014; Holl and Burnett, 2014; Knight et al., 2016; Condello et al., 2019; Owiti and Hauw, 2021). Thus, positive interaction with coaches for athletic and academic achievements is required to encourage athletes enduring education requirements as well. A positive interaction between parents and academic staff is also crucial not only to help athletes manage major issues related to balancing timetables, tiredness, feeling of school belonging, goal settings, and personal development to prepare for post-sporting careers but also to foster a leadership role of student-athletes as role models in supporting active and healthy lifestyles promotion in the academic population, which could benefit from the close connection between learning abilities and physical movement or sports activities (O'Neill et al., 2017; Hills et al., 2019; Tomporowski and Pesce, 2019; López-Flores et al., 2021; Pesce et al., 2021). In fact, in considering education programmes as a capacity-building process, it is crucial to recognize and share the accumulated and valuable knowledge and experiences of the athletes' families to build impactful and sustainable collaborative partnerships.

CONCLUSION

Although sports organizations and academic institutions are mainly responsible for their dual-career policies, if any, in the last decade, the European Parliament and Commission highly promoted the establishment and management of friendly environments for student-athletes through the sports call of the ERASMUS+ programme funding topic "Promote education in and through sport with special focus on skills development, as well support the implementation of the EU Guidelines on Dual Careers of Athletes" (European Commission, 2012, 2022; European Parliament, 2015, 2017). Despite these efforts in fostering a European dual-career discourse, the development of dual-career environments for the entire dual-career pathway (from school-age to university and vocation) in Member States is still limited, particularly in relation to the empowerment of parents of athletes (Guidotti et al., 2015; Capranica and Guidotti, 2016; Stambulova and Wylleman, 2019; López-Flores et al., 2021; Morris et al., 2021; Tessitore et al., 2021). The present work was the first EU-funded collaborative partnership addressing the issues of sustaining high school and university student-athletes by educating the supporting capabilities of their parents (Capranica et al., 2018). However, there is a need for further investigations to provide empirical validation of an assessment tool for the lifelong learning competencies of parents of dual-career athletes. Moreover, future studies are needed to investigate sports and academic bodies developing and supporting inclusive schemes and initiatives to comply with the European Guidelines, also including communication or formal arrangements with parents (European Commission, 2012). To bridge the gap between dual-career policies and practice, the online multilingual EMPATIA education programme can be considered a valuable resource and good practice for encouraging further initiatives based on a participatory and a user-centered design approach to describe implementation context and apply that information. In illuminating and understanding the needs of the dual-career actors, and in translating the information on actual needs in a free online education, the EMPATIA education programme could generate gains at individual levels and enhance the implementation of the European dual-career culture.

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 Research Ethics Committee of the University of Ljubljana of the EMPATIA project (9:2018). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by the European Commission under the Erasmus + Programme (number 590437-EPP-1-2017-1-SI-SPO-SCP).

REFERENCES

- Aquilina, D., and Henry, I. (2010). Elite athletes and university education in Europe: a review of policy and practice in higher education in the European Union Member States. *Int. J. Sport Policy Polit.* 2, 25–47. doi: 10.1080/19406941003634024
- Aunola, K., Sorkkila, M., Viljaranta, J., Tolvanen, A., and Ryba, T. V. (2018). The role of parental affection and psychological control in adolescent athletes' symptoms of school and sport burnout during the transition to upper secondary school. *J. Adolesc.* 69, 140–149. doi: 10.1016/j.adolescence.2018.10.001
- Bean, C. N., Fortier, M., Post, C., and Chima, K. (2014). Understanding how organized youth sport may be harming individual players within the family unit: a literature review. *Int. J. Environ. Res. Publ. Health* 11, 10226–10268. doi: 10.3390/ijerph111010226
- Callegaro, M., Manfreda, K. L., and Vehovar, V. (2015). *Web Survey Methodology*. Thousand Oaks, CA: Sage. doi: 10.4135/9781529799651
- Capranica, L., and Guidotti, F. (2016). *Research for Cult Committee: Qualifications/Dual Careers in Sports. European Parliament: Directorate-General for Internal Policies. Policy Department. Structural and Cohesion Policies: Cultural and Education*. Available online at: [http://www.europarl.europa.eu/RegData/etudes/STUD/-2016/573416/IPOL_STU\(2016\)573416_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/-2016/573416/IPOL_STU(2016)573416_EN.pdf)
- Capranica, L., MacDonncha, C., Blondel, L., Bozzano, E., Burlot, F., Costa, R., et al. (2018). Towards the construction of an educational model for dual career parenting: the EMPATIA project. *Kinesiol. Sloven.* 24, 19–30.
- Clarke, N. J., and Harwood, C. G. (2014). Parenting experiences in elite youth football: a phenomenological study. *Psychol. Sport Exerc.* 15, 528–537. doi: 10.1016/j.psychsport.2014.05.004
- Condello, G., Capranica, L., Doupona, M., Varga, K., and Burk, V. (2019). Dual-career through the elite university student-athletes' lenses: the international FISU-EAS survey. *PLoS ONE* 14:e0223278. doi: 10.1371/journal.pone.0223278
- Cosh, S., and Tully, P. J. (2014). “All I have to do is pass”: a discursive analysis of student athletes' talk about prioritising sport to the detriment of education to overcome stressors encountered in combining elite sport and tertiary education. *Psychol. Sport Exerc.* 15, 180–189. doi: 10.1016/j.psychsport.2013.10.015
- Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., et al. (2014). Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol.* 67, 401–409. doi: 10.1016/j.jclinepi.2013.12.002
- Dorsch, T. E., King, M. Q., Dunn, C. R., Osai, K. V., and Tulane, S. (2017). The impact of evidence-based parent education in organized youth sport: a pilot study. *J. Appl. Sport Psychol.* 29, 199–214. doi: 10.1080/10413200.2016.1194909
- European Commission (2007). *White Paper on Sport*. European Commission. Available online at: http://www.aop.pt/upload/tb_content/320160419151552/35716314642829/whitepaperfullen.pdf
- European Commission (2012). *Guidelines on Dual Careers of Athletes-Recommended Policy Actions in Support of Dual Careers in High-Performance*

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the parents of dual-career athletes and to the European dual-career experts for their support when carrying out the workshops for the evaluation phases of the EMPATIA education model.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.855531/full#supplementary-material>

- Sport*. Available online at: https://ec.europa.eu/assets/eac/sport/~library/documents/dual-career-guidelines-final_en.pdf
- European Commission (2016). *Directorate General for Education and Culture., Amsterdam University of Applied Sciences, Birch Consultants, Talented Athlete Scholarship Scheme, Vrije Universiteit Brussel. Study on the Minimum Quality Requirements for Dual Career Services: Final Report*. European Commission: Publications Office. Available online at: <https://data.europa.eu/doi/10.2766/345818>
- European Commission (2022). *2022 Annual Work Programme “Erasmus+”: the Union Programme for Education, Training, Youth and Sport*. European Commission. Available online at: <https://erasmus-plus.ec.europa.eu/it/node/399>
- European Parliament (2015). *EU Sport Policy. An Overview*. Brussels: European Parliamentary Research Service (978-92-823-7903-5). Available online at: [http://www.europarl.europa.eu/RegData/etudes/IDAN/2015/565908/EPRS_IDA\(2015\)565908_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/IDAN/2015/565908/EPRS_IDA(2015)565908_EN.pdf)
- European Parliament (2017). *European Parliament Resolution of 2 February 2017 on an Integrated Approach to Sport Policy: Good Governance, Accessibility and Integrity*. Available online at: <http://www.europarl.europa.eu/sides/getDoc.do?type=T&andreference=P8-T&A=2017-0012&language=EN&ndring=A8-2016-0381>
- European Parliament (2021). *Motion for a European Parliament Resolution on EU Sports Policy: Assessment and Possible Ways Forward*. Available online at: https://www.europarl.europa.eu/doceo/document/A-9-2021-0318_EN.html
- Ferguson, M., Leighton, P., Brandreth, M., and Wharrad, H. (2018). Development of a multimedia educational programme for first-time hearing aid users: a participatory design. *Int. J. Audiol.* 57, 600–609. doi: 10.1080/14992027.2018.1457803
- Genzuk, M. (eds.). (2003). *A Synthesis of Ethnographic Research*. Occasional Papers Series. Los Angeles, CA: Center for Multilingual, Multicultural Research; Rossier School of Education; University of Southern California, 1–10.
- Gjaka, M., Tessitore, A., Blondel, L., Bozzano, E., Burlot, F., Debois, N., et al. (2021). Understanding the educational needs of parenting athletes involved in sport and education: the parents' view. *PLoS ONE* 16:e0243354. doi: 10.1371/journal.pone.0243354
- Guidotti, F., Cortis, C., and Capranica, L. (2015). Dual career of European student-athletes: a systematic literature review. *Kinesiol. Sloven.* 21, 1318–2269.
- Hansen, N. B., Dindler, C., Halskov, K., Iversen, O. S., Bossen, C., Basballe, D. A., et al. (2019). “How participatory design works: mechanisms and effects,” in *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, (Fremantle, WA), 30–41. doi: 10.1145/3369457.3369460
- Harwood, C., Drew, A., and Knight, C. J. (2010). Parental stressors in professional youth football academies: a qualitative investigation of specialising stage parents. *Qual. Res. Sport Exerc.* 2, 39–55. doi: 10.1080/19398440903510152
- Harwood, C., and Knight, C. J. (2009). Understanding parental stressors: an investigation of British tennis-parents. *J. Sports Sci.* 27, 339–351. doi: 10.1080/02640410802603871

- Harwood, C., and Knight, C. J. (2015). Parenting in youth sport: a position paper on parenting expertise. *Psychol. Sport Exerc.* 16, 24–35. doi: 10.1016/j.psychsport.2014.03.001
- Hills, S., Walker, M., and Barry, A. E. (2019). Sport as a vehicle for health promotion: a shared value example of corporate social responsibility. *Sport Manage. Rev.* 22, 126–141. doi: 10.1016/j.smr.2018.10.001
- Holl, L., and Burnett, C. (2014). Changing relationships with significant others: reflections of national and international level student-athletes. *South Afr. J. Res. Sport Phys. Educ. Recreat.* 36, 115–128.
- Hsu, C. C., and Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Pract. Assess. Res. Eval.* 12:10. doi: 10.7275/pdz9-th90
- Knight, C. J., Dorsch, T. E., Osai, K. V., Haderlie, K. L., and Sellars, P. A. (2016). Influences on parental involvement in youth sport. *Sport Exerc. Perform. Psychol.* 5:161. doi: 10.1037/spy0000053
- Knight, C. J., Harwood, C. G., and Sellars, P. A. (2018). Supporting adolescent athletes' dual careers: the role of an athlete's social support network. *Psychol. Sport Exerc.* 38, 137–147. doi: 10.1016/j.psychsport.2018.06.007
- López-Flores, M., Hong, H. J., and Botwina, G. (2021). Dual career of junior athletes: Identifying challenges, available resources, and roles of social support providers. *Cult. Cienc. Dep.* 16, 117–129.
- McGreal, R. (2004). *Online Education Using Learning Objects*. New York, NY: Routledge. doi: 10.4324/9780203416082
- Morris, R., Cartigny, E., Ryba, T. V., Wylleman, P., Henriksen, K., Torregrossa, M., et al. (2021). A taxonomy of dual career development environments in European countries. *Eur. Sport Manage. Q.* 21, 134–151. doi: 10.1080/16184742.2020.1725778
- O'Neill, M. M., Calder, A. A., and Hinz, B. (2017). Student-athletes in my classroom: Australian teachers' perspectives of the problems faced by student-athletes balancing school and sport. *Austr. J. Teach. Educ.* 42, 160–178. doi: 10.14221/ajte.2017v42n9.10
- Owiti, S., and Hauw, D. (2021). The problematic experience of players' mutations between clubs: discovering the social adaptability skills required. *Front. Sports Active Living* 3:591438. doi: 10.3389/fspor.2021.591438
- Pesce, C., Vazou, S., Benzing, V., Álvarez-Bueno, C., Anzeneder, S., Mavilidi, M. F., et al. (2021). Effects of chronic physical activity on cognition across the lifespan: a systematic meta-review of randomized controlled trials and realist synthesis of contextualized mechanisms. *Int. Rev. Sport Exerc. Psychol.* 1–39. doi: 10.1080/1750984X.2021.1929404
- Sheehan, R. B., Herring, M. P., and Campbell, M. J. (2018). Longitudinal relations of mental health and motivation among elite student-athletes across a condensed season: plausible influence of academic and athletic schedule. *Psychol. Sport Exerc.* 37, 146–152. doi: 10.1016/j.psychsport.2018.03.005
- Sorkkila, M., Ryba, T. V., Selänne, H., and Aunola, K. (2020). Development of school and sport burnout in adolescent student-athletes: a longitudinal mixed-methods study. *J. Res. Adolesc.* 30, 115–133. doi: 10.1111/jora.12453
- Stambulova, N. B., Ryba, T. V., and Henriksen, K. (2021). Career development and transitions of athletes: the international society of sport psychology position stand revisited. *Int. J. Sport Exerc. Psychol.* 19, 524–550. doi: 10.1080/1612197X.2020.1737836
- Stambulova, N. B., and Wylleman, P. (2019). Psychology of athletes' dual careers: a state-of-the-art critical review of the European discourse. *Psychol. Sport Exerc.* 42, 74–88. doi: 10.1016/j.psychsport.2018.11.013
- Su, C. T., and Parham, L. D. (2002). Generating a valid questionnaire translation for cross-cultural use. *Am. J. Occup. Ther.* 56, 581–585. doi: 10.5014/ajot.56.5.581
- Tessitore, A., Capranica, L., Pesce, C., De Bois, N., Gjaka, M., Warrington, G., et al. (2021). Parents about parenting dual career athletes: a systematic literature review. *Psychol. Sport Exerc.* 53:101833. doi: 10.1016/j.psychsport.2020.101833
- Thrower, S. N., Harwood, C. G., and Spray, C. M. (2016). Educating and supporting tennis parents: a grounded theory of parents' needs during childhood and early adolescence. *Sport Exerc. Perform. Psychol.* 5:107. doi: 10.1037/spy0000054
- Tompowski, P. D., and Pesce, C. (2019). Exercise, sports, and performance arts benefit cognition via a common process. *Psychol. Bull.* 145:929. doi: 10.1037/bul0000200
- Varga, K., MacDonncha, C., Blondel, L., Bozzano, E., Burlot, F., Costa, R., et al. (2021). Collective conceptualization of parental support of dual career athletes: the EMPATIA framework. *PLoS ONE* 16:e0257719. doi: 10.1371/journal.pone.0257719

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 Capranica, Guidotti, Gonçalves, Blondel, Bovis, Costa, Debois, Figueiredo, MacDonncha, Pecnikar-Oblak, Patoret, Pišl, Rheinisch, Rolo, Ryan, Templet, Tessitore, Warrington and Doupona. 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 Effects of Teacher Feedback and Automated Feedback on Cognitive and Psychological Aspects of Foreign Language Writing: A Mixed-Methods Research

Zehua Wang^{1†} and Feifei Han^{2*†}

¹ School of Foreign Languages, Shaanxi Xueqian Normal University, Xi'an, China, ² Institute for Learning Sciences and Teacher Education, Australian Catholic University, Sydney, QLD, Australia

OPEN ACCESS

Edited by:

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

Reviewed by:

Chian-Wen Kao,
Chihlee University of Technology,
Taiwan
Sophia Skoufaki,
University of Essex, United Kingdom

*Correspondence:

Feifei Han
feifei.han@acu.edu.au

[†] 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: 31 March 2022

Accepted: 15 June 2022

Published: 28 July 2022

Citation:

Wang Z and Han F (2022) The
Effects of Teacher Feedback
and Automated Feedback on
Cognitive and Psychological Aspects
of Foreign Language Writing:
A Mixed-Methods Research.
Front. Psychol. 13:909802.
doi: 10.3389/fpsyg.2022.909802

Feedback plays a crucial role in the writing processes. However, in the literature on foreign language (FL) writing, there is a dearth of studies that compare the effects of teacher feedback and automated feedback on both cognitive and psychological aspects of FL writing. To fill this gap, the current study compared the effects of teacher feedback and automated feedback on both revision quality and writing proficiency development (i.e., the cognitive aspects), and perceived usefulness and perceived ease of use of the feedback (i.e., the psychological aspects) in English writing among English learners as an FL (EFLs) in China. It also investigated students' perceptions of the strengths and weaknesses of the two types of feedback. The study adopted a mixed-methods design. The quantitative method collected the data through (1) a pre-test and a post-test, which measured the participants' English writing proficiency development; (2) a writing task, which received either teacher feedback or automated feedback; and (3) a close-ended questionnaire, which examined students' perceived usefulness and perceived ease of use of the feedback. The qualitative method collected the data through an open-ended questionnaire, which examined the participants' perceptions of the strengths and weaknesses of teacher feedback or automated feedback depending on the type of feedback they received. Chinese university EFLs in two English classes ($n = 35$ in each class) taught by the same English teacher participated in the study: one class received teacher feedback while the other received automated feedback using Pigaiwang. While the students in the two classes did not differ significantly on the pre-test of students' writing proficiency, students who received teacher feedback scored significantly higher on revision than those who received automated feedback. Students in the teacher feedback class also had significantly higher ratings on perceived usefulness and perceived ease of use of the feedback than those in the automated feedback class. However, students in the automated feedback class obtained significantly higher scores on the post-test of the writing proficiency. The

qualitative results identified three themes of strengths and two themes of weaknesses for the teacher feedback and the automated feedback, respectively. The results suggest that while teacher feedback has a more positive effect on the psychological aspect of FL writing, automated feedback may be more effective in developing FL writing proficiency in the long run.

Keywords: teacher feedback, automated feedback, foreign language writing, revision quality, writing proficiency development, perceived usefulness, perceived ease of use, perceptions of the feedback

INTRODUCTION

In early 2020, the COVID-19 pandemic spread rapidly worldwide, which brought great challenges to all walks of life, including educational institutions. In order to ensure the normal progress of learning and at the same time to prevent the spread of COVID-19, institutions across countries were required to redeploy more learning and teaching activities to virtual learning spaces in order to maintain physical distancing. As a result, the vast numbers of face-to-face courses have been delivered either as blended courses or as purely online courses (Tang et al., 2021). Under such circumstances, technology-supported learning, particularly freely accessible web-based tools of high quality have been playing an important role in the learning and teaching processes. In foreign language (FL) writing, one such useful tool is web-based automated writing feedback systems.

A web-based automated writing feedback system is also known as online automated writing feedback, which is a type of an online platform that can be flexibly and easily accessed by students and provide immediate feedback to them (Warschauer and Grimes, 2008). The online automated feedback can also enable teachers to identify an individual student's level of writing ability in relation to the whole class when all students' essays are entered into the system (Bai and Hu, 2017). Because of these benefits, online automated feedback platforms have been increasingly adopted by English teachers around the world to fully or partially replace teacher feedback in English writing classes (Warschauer and Ware, 2006).

The functions of automated feedback have been recognized as both an assessment and a learning tool. Research on the effectiveness of automated feedback as an assessment tool has demonstrated that automated feedback is advantageous for its ability to significantly reduce logistics for marking a large number of written scripts and to evaluate writing more objectively than using human raters (Shermis and Burstein, 2003; Wilson et al., 2014). These merits make automated feedback particularly favorable for scoring written scripts of test-takers in large-scale standardized tests (Ramineni, 2013). As a learning tool, automated feedback can provide feedback on learners' writing in various aspects, including mechanics, vocabulary, grammar, and content and organization, which not only can assist learners in improving the quality of writing products, but may also help them acquire additional linguistic knowledge, such as learning new words by reading the synonyms offered by automated feedback (Wilson and Cziki, 2016).

However, there is little research comparing the effects of teacher feedback and automated feedback on both FL learners' revision quality and their writing proficiency development (the cognitive aspects of FL writing) in a single study. Moreover, from a psychological perspective, it is also important to know how learners perceive the usefulness and ease of use of automated feedback compared to teacher feedback, as operating automated feedback requires a certain level of knowledge of using computers. From a teaching point of view, in order to help teachers make informed decisions as to how much the two types of feedback should be used in the FL writing classes, students' perceptions of the strengths and weaknesses (the psychological aspects of FL writing) of the two types of feedback also should be understood. To address these research gaps, the current study compares the effects of teacher feedback and automated feedback on both the cognitive and psychological aspects of FL writing among learners of English as an FL (EFL learners) in China. The following sections review relevant literature on both teacher feedback and automated feedback.

LITERATURE REVIEW

Feedback in Foreign Language Writing

Revising is an important part of the writing process, especially when writing in an FL. FL learners who routinely revise inadequacies in their texts after receiving feedback tend to develop better writing skills than those who do not (Ferris and Roberts, 2001; Bitchener, 2008). To help FL learners achieve a desirable revision quality in writing, providing various types of feedback about their writing is of great importance. There are a variety of types of feedback, including corrective, non-corrective, direct, indirect, local, and global (Luo and Liu, 2017). The feedback can also focus on different features in the writing, such as mechanics, grammatical errors, vocabulary and collocations, and content and structure of the writing (Lee, 2008; Boubekeur, 2015). Researchers suggest that the good practice of writing feedback should cover both form and content in writing (Biber et al., 2011).

One major issue that has been consistently addressed in FL writing is how to provide effective feedback (Lee, 2017). Researchers and educators have proposed both unfocused and selective approaches to provide feedback; however, the two approaches tend to target learners with different levels of FL proficiency (Ferris et al., 2013). In the unfocused approach, teachers give comprehensive feedback to students, responding

to every single error, whereas in the selective approach, teachers only target a selected type of error in their feedback (Lee, 2013). Hence, the unfocused approach is more suitable for advanced learners, as their writings tend not to have a large number of errors. On the other hand, the selective approach is more appropriate for less proficient learners, because selective feedback can avoid overwhelming learners, at the same time enable them to notice the focused types of errors (Sheen, 2007; Bitchener, 2008; Bitchener and Ferris, 2012). As to whether direct or indirect feedback is more effective for FL learners, there is no conclusive evidence. Some researchers believe that indirect feedback provides learners opportunities to use their existing knowledge to correct the errors themselves, which not only can engage learners in the revising processes but tends to be more effective to promote accuracy in writing in the long term (e.g., Ferris, 2006; Bitchener and Ferris, 2012). More recent studies, however, reported that it seemed to be more effective to provide direct corrective feedback involving metalinguistic explanations, through which learners' cognitive engagement could also be enhanced (Bitchener and Knoch, 2008, 2009, 2010; Esfandiari et al., 2014).

Teacher Feedback on the Revision Quality and the Writing Proficiency Development

A large number of studies have examined the effect of teacher feedback on FL learners' revision, which demonstrate that different types of teacher feedback have different effects on learners' revision (Hyland and Hyland, 2006). Research has reported that teachers' positive comments and encouraging language could build students' confidence, which is considered important in the revising processes (Ferris, 1997). Studies have also suggested that in order for teacher feedback to be effective in the revising process, direct, specific, and content-related feedback should be given (Ferris, 2011). Of the three types of feedback, namely advice, criticism, and praise, research has shown that the advice type of feedback is more likely to prompt students to revise (Silver and Lee, 2007).

Compared to the research on the effect of teacher feedback on FL learners' revision, fewer studies have investigated the effect of teacher feedback on FL learners' writing proficiency development. The effect of teacher feedback on learners' writing proficiency development may not be as effective as on the revision quality, as learners may simply directly copy teacher corrective feedback without understanding the errors (Hyland, 1998; Lee, 2007; Zhao, 2010). As a result, they will still make the same mistake in their subsequent writings. Zhao (2010) has suggested that teacher feedback "that is used/copied but misunderstood may help to improve writing quality but does not necessarily contribute to the development of learners' long-term writing proficiency" (p. 4). Seeing these problems, researchers have proposed that the examination of the effects of teacher feedback should also be examined in subsequent instances of writings beyond just the revised drafts of the same text (Ruegg, 2015). Thus, it is necessary to investigate the effects of teacher

feedback on students' revision and their writing proficiency development simultaneously.

Students' Perceptions of Teacher Feedback

In general, FL writers attach great importance to teacher feedback, as they believe that teachers are more authoritative in giving writing feedback than their peers (Ferris, 1995; Hyland and Hyland, 2006; Biber et al., 2011). FL learners believe that teacher feedback not only helps them avoid making similar mistakes in their subsequent writing (Chandler, 2003), but also strengthens their confidence and motivation in FL writing, particularly when teachers use positive language in the feedback (Weaver, 2006).

However, teacher feedback focusing on different aspects of writing is perceived differently by learners. Some FL writers perceive that the comments on contents and structures are the most important and useful feedback (Semke, 1984; Zamel, 1985; Zhan, 2016), whereas others value feedback on form over content (Saito, 1994; Hedgcock and Lefkowitz, 1996; Ashwell, 2000; Lee, 2005). Some learners even expect to receive feedback on all aspects of their writing, including language problems, contents, and organizational structure (Radecki and Swales, 1988; Ferris, 1995; Lee, 2005). Learners may favor different types of feedback based on different reasons. Some FL writers prefer indirect feedback on the basis that the indirect type gives them more agency to actively participate in the revision processes (Hyland, 2001). In contrast, other learners welcome detailed and personalized feedback with clear explanations of the errors (Ferguson, 2011; Dawson et al., 2019).

The FL learners do not always hold positive perceptions toward teacher feedback. Some students mentioned that they either did not understand teacher feedback or they found teachers' language in the feedback was ambiguous, hence they had to ignore these comments in the revising processes (Ferris, 1997; Conrad and Goldstein, 1999; Rollinson, 2005; Kim, 2015). These problems were particularly prominent among students with low self-efficacy in FL writing and with poor writing proficiency (Lee, 2008). While past studies have investigated students' perceptions of teacher feedback, relatively little research has compared students' perceptions of the strengths and weaknesses of teacher feedback and automated feedback.

Automated Feedback on the Revision Quality and the Writing Proficiency Development

With the development of educational technology, automated feedback has been increasingly applied in English writing evaluation and instruction (Chen and Cheng, 2008; Warschauer and Grimes, 2008; Grimes and Warschauer, 2010). The initial aim of the development of automated feedback was on scoring a large number of essays in standardized writing assessments (Page, 2003). In recent years, automated feedback has been also employed in FL writing classrooms to provide timely feedback in classes with large enrolments (Stevenson and Phakiti, 2014; Liao, 2016; Wilson and Andrada, 2016). The apparent strength of automated feedback, especially web-based automated

feedback, lies in its efficiency and flexibility, as it can identify errors and provide immediate feedback merely by a click on the web page (Chen and Cheng, 2008; Cotos, 2011). It is particularly effective to provide corrective feedback on the aspects of mechanics and grammatical errors (Wilson et al., 2014; Li et al., 2015). Bai and Hu (2017) reported that the correction rates of automated feedback were 57 and 42% for grammatical and collocation errors, respectively. Others found that the success rate of error corrections could be as high as 70% (Chapelle et al., 2015; Liao, 2016). Moreover, automated feedback also has the potential to reduce the burden for English teachers in terms of managing, storing, and marking FL learners' writing samples (Manap et al., 2019).

Despite these benefits, automated feedback has been criticized for its low-quality feedback on the content and organization of the writing (Warschauer and Grimes, 2008; Wang, 2013). For instance, some popular automated feedback systems, such as *Criterion* and *My Access!*, predominantly focus on detecting language errors (Dikli, 2010), but are limited in identifying high-level problems, such as content and logic (Deane, 2013). The predominant foci of the mechanical and linguistic features generated by automated feedback may mislead FL learners to think that FL writing practice is all about language aspects, neglecting the content and rhetoric aspects of the writing (Cheville, 2004). Another concern for using automated feedback is that it requires learners to have some levels of learning autonomy so that they can sustainably interact with the machine (Warschauer and Ware, 2006). Thus, automated feedback may not be suitable for younger FL learners due to their lack of learning autonomy (Lee, 2017).

Students' Perceptions of Automated Feedback

As to students' perceptions of automated feedback, the majority of the existing research has explored students' perceptions of automated feedback in the context of first language writing (Calvo and Ellis, 2010; Grimes and Warschauer, 2010). In the FL writing, the limited research has produced mixed results (Chen and Cheng, 2008; Calvo and Ellis, 2010; Dikli and Bleyle, 2014; Bai and Hu, 2017). While some researchers have reported that students hold negative perceptions of automated feedback as they believe that automated feedback does not provide sufficient information on the contents of the writing (Chen and Cheng, 2008), others have positive attitudes due to the flexibility in accessing automated feedback (Fang, 2010). Students also have different perceptions as to the automated feedback on different aspects of FL writing. While most students have positive perceptions of the feedback mechanics and grammar provided by the automated feedback, they showed some concerns about the reliability of the feedback on collocations (Bai and Hu, 2017).

Studies investigating students' perceived usefulness and ease of use of automated feedback seemed scarce. Perceived usefulness refers to the degree to which a user perceives that using a particular technology system would enhance his/her performance (Davis, 1989), whereas perceived ease of use is defined as the

degree to which a user expects that using a particular technology system is free of effort (Davis, 1989; Venkatesh and Davis, 2000). These two constructs are the most important constructs in the Technology Acceptance Model (Davis, 1989; Neo et al., 2015), and have been widely researched in users' experience of using e-learning and technology systems (Zyad, 2016). As the web-based automated feedback system is also an e-learning system, examination of FL writers' perceived usefulness and ease of use of the automated feedback system is important.

Comparing Teacher Feedback and Automated Feedback in Foreign Language Writing

To date, only a small number of studies have compared the effects of teacher feedback and automated feedback on FL writing (Warden, 2000; Wang and Wang, 2012; Dikli and Bleyle, 2014; Wilson and Czik, 2016; Link et al., 2020). However, these studies suffer from some design issues. Warden reported a better effect of automated feedback on reducing learners' error rates than teacher feedback. In his study, the feedback from automated feedback was on specific errors, whereas teacher feedback was general comments. Similarly, the participant in the teacher feedback condition in Wang and Wang's study also only received the global comments on his writing, whereas the participant in the automated feedback condition received the specific comments on grammar, spelling, and collocations. Moreover, this study only had two participants, which severely limited the generalizability of the findings. In Link et al.'s (2020) and Wilson and Czik's (2016) study, students in the automated feedback conditions received a combination of automated feedback and teacher feedback. Hence, the comparison was not purely between teacher feedback and automated feedback. To address these methodological issues, the current study will (1) have pure teacher feedback and a pure automated feedback condition, and (2) require the teacher to provide feedback by covering all the aspects in writing, which are covered in automated feedback.

The Current Study and Research Questions

The literature review shows that there is a lack of research comparing the effects of teacher feedback and automated feedback on both cognitive and psychological aspects of FL writing. To fill this gap, the current study compared the effects of teacher feedback and automated feedback on both revision quality and writing proficiency development (i.e., the cognitive aspects), and perceived usefulness and perceived ease of use of the feedback (i.e., the psychological aspects) in English writing among English learners as an FL (EFLs) in China. It also investigated students' perceptions of the strengths and weaknesses of the two types of feedback. The current study sought to answer the following three research questions:

- (1) To what extent do revision quality and writing proficiency development differ between Chinese EFLs who receive teacher feedback and automate feedback?

- (2) To what extent do perceived usefulness and ease of use differ between Chinese EFLs who receive teaching feedback and automated feedback?
- (3) What are the strengths and drawbacks of teacher feedback and automated feedback perceived by Chinese EFLs?

METHOD

Research Design

The study adopted a mixed-methods design: the quantitative method provided the answers to the first and the second research questions, whereas the qualitative method provided the answer to the third research question. For the quantitative method, we conducted a quasi-experiment as it was not possible to randomly assign the participants into two groups due to the university's policy. Hence, we designed the quasi-experiment on the basis of the two intact classes: one class received teacher feedback and the other received web-based automated feedback on their English essay drafts. The quantitative part also collected students' responses to perceived usefulness and ease of use of either teacher feedback or automated feedback, depending on which one they received, through a Likert-scale questionnaire. The qualitative method obtained students' perceptions of the strengths and drawbacks of teacher feedback and automated feedback through an open-ended questionnaire. In the following sections, details regarding the participants, instruments, data collection, and analysis methods are explained.

Participants

A total of 70 Chinese freshmen, who majored in English Education in Early Childhood participated in the study. Among them, 67 were women and only three were men. The uneven gender distribution was largely attributable to the fact that the major of Early Childhood Education generally attracts female students in China. The 70 students attended two English classes taught by the same English teacher, with each class having 35 students. The participants were aged between 18 and 21 years, with an average of 19.5 years. All the participants had studied English as a compulsory subject for 9 years, from grade three in primary school to completion of high school. At the time of the data collection, students had just commenced their university studies, hence they did not have opportunities to take part in any national examinations for college students. Therefore, we gathered students' English scores from the National College Entrance Examinations as an indicator of their English proficiency. The total score of the National College Entrance English Examination is 150. Of the 70 participants, 34 had scores ranging from 85 to 100, 28 ranging from 100 to 110, and 8 ranging from 110 to 120. Therefore, their English proficiency could be placed at the lower intermediate to intermediate levels.

Instruments

The Writing Tasks

Three writing tasks were used in the study. The first writing task served as a pre-test of participants' English writing proficiency in order to examine if students in the two English classes had

similar English writing proficiency before the quasi-experiment. As the participants had just commenced their university life, we used a writing task titled "My first day at the university," which was considered appropriate and relevant to students' life experiences. A one-way ANOVA was conducted to examine if the students in the two feedback conditions had the same initial writing proficiency. Levene's test found that the assumption of homogeneity of variances was met [$F(1,68) = 0.76, p = 0.38$]. The results showed that there was no significant difference [$F(1, 68) = 0.31, p = 0.58, \eta^2 = 0.13$] in the pre-test of English writing proficiency between students receiving teacher feedback ($M = 66.66, SD = 7.73$) and those receiving automated feedback ($M = 67.63, SD = 6.75$).

The second writing task, which was titled "The most impressive classmate in my university," was used to provide feedback to students for them to revise. The third writing task was called "The most successful thing I have done," which was used to test students' writing proficiency after the quasi-experiment. All three writing tasks required students to produce an approximate 150-word English text following a structure of three compulsory parts, namely, an introduction, a body text, and a conclusion. As past research suggests that text type can affect FL writers' writing performance (Li, 2014), we therefore used a single text type, that is, narratives for all three tasks. We purposefully chose to use narratives rather than expositions or other text types for the writing tasks, because the participants were familiar with this text type. The topics of the three writing tasks came from the category of the daily practice of narratives for English major students in China in the bank of the web-based automated feedback used in the study in order to ensure that the three writing tasks had similar difficulty levels.

The Writing Feedback

The online automated feedback used in this study was called "Pigaiwang," whose word-to-word English translation is "Marking Website" (see <http://www.pigai.org/> for an example of the interface of the Pigaiwang). Entering the market in 2011, this platform has registered as a patent in China (ZL2012 10049982.4). It is both a corpus-based and a cloud-computing-based online service for automated evaluation and feedback of English writing by Chinese EFLs. The reliability and the validity of the Pigaiwang for English essay scoring were established by calibrating with a large corpus of human-scored English essays (Hu, 2015; Yang and Dai, 2015). It has been reported that the correlation between essay scores in the Pigaiwang and human raters was high and satisfying (He, 2013; Wang, 2016). With years of development, the platform has evolved into the most popular and most widely subscribed web-based automated feedback platform in China, with more than 20 million registered users by early 2018.

The Pigaiwang provides four main functions for learners:

- Scoring the essay: this function computes a score immediately upon the submission of an essay to indicate the quality of English writing. Each essay is scored on the four dimensions of vocabulary, grammar, structure and organization, and content and relevance, each of which is scored by comparing the quality of the submitted essays

with a large human-scored essay corpus. The possible writing scores range from 0 to 100. In addition to the scores, when students are registered under one class, the system will also generate a rank of their writing quality relative to the whole class performance.

- Providing immediate holistic feedback: this function uses both bar graphs and comments to demonstrate the strengths of an essay in terms of vocabulary, grammar, structure and organization, and content and relevance.
- Providing corrective feedback at the sentence level: this function provides diagnostic comments by pointing out the errors in mechanics (e.g., spelling, punctuations, and capitalization), vocabulary (e.g., word choice and collocation), grammar, and content and relevance. It also gives suggestions for revision at the level of individual sentences and recommendations for collocations. The recommended collocations are listed in ranks according to the frequency of the appearance in corpora.
- Listing suggestions for synonyms at the level of a word: this function offers the writers multiple synonyms in order to enhance the vocabulary diversity of the writing. For each synonym, it also supplies hyperlinks for further information on the meaning of the synonyms and detailed explanations as to the differences between the synonyms and the word appearing in the submitted texts.

The English teacher's feedback matched the format of the feedback generated in the Pigaiwang. **Table 1** provides some examples of the teacher feedback by types of errors.

The Likert-Scale Questionnaire on Perceived Usefulness and Ease of Use of the Feedback

To measure students' perceived usefulness and ease of use of the feedback they received, two 5-point Likert scales were adapted from the existing scales. The items were adapted from Venkatesh et al. (2003) and Venkatesh and Bala (2008), which were originally developed and reported by Davis (1989) and Davis et al. (1989). The wording of the scales used for students who received teacher feedback and automated feedback was exactly the same, except for the words "teacher feedback" and "automated feedback." The perceived usefulness scale had four items, and its reliability was 0.68 for teacher feedback and 0.67 for automated feedback. The perceived ease of use scale had three items, and its reliability was 0.69 for both teacher feedback and automated feedback.

The Open-Ended Questionnaire on the Strengths and Weaknesses of the Feedback

The open-ended questionnaire asked students to list three aspects of both strengths and weaknesses of either teacher feedback or automated feedback depending on which one they received.

Ethics Consideration

Prior to the study, the students in the two classes were informed about the purposes of the study and were invited to participate in the study voluntarily. Before the data collection, an ethical application was submitted to the ethics committee of the School

of Foreign Languages, Shaanxi Xueqian Normal University. The committee evaluated the nature of the study and believed that the study would be a component of classroom teaching. Hence, the participants were not required to sign a written consent form. However, all the participants needed to agree verbally for the voluntary participation. The ethics committee recorded the participants' verbal consent.

Procedure of the Data Collection

The research was conducted in four English sessions. In the first English session, participants in the two classes completed the first writing task. The scores were used to represent students' initial English writing proficiency. In the second English session, both groups completed the second writing task. The students in the automated feedback class submitted their essays in the Pigaiwang, whereas students in the teacher feedback class submitted their essays to the English teacher. In the third English class, the students were instructed to revise their essays using either teacher feedback or automated feedback. The scores of the revised essays were used to represent the revision quality. In the fourth English session, they were given the third writing task. The scores of the third writing task were used to assess students' post-English writing proficiency. Upon completion of the third writing task, they were also given both the Likert-scale questionnaire as well as the open-ended questionnaire to fill. The essays for the first writing task (pre-test of English writing proficiency), the revised essays for the second writing task, and the essays for the third writing task (post-test of English writing proficiency) were scored in the Pigaiwang to prevent marking bias of the human raters. The procedure of the data collection is summarized in **Table 2**.

Data Analysis

To answer the first research question, comparing students' revision quality and post-test of English writing proficiency, a mixed-design 2 (within-subjects factor: revision quality and post-test of writing proficiency) \times 2 (between-subjects factor: teacher feedback vs. automated feedback) ANOVA was conducted. To answer the second research question, comparing students' perceived usefulness and ease of use of the feedback, a MANOVA was used. To examine students' perceptions of the strengths and weaknesses of the two types of feedback, a thematic analysis of the students' responses to the open-ended questionnaire was applied.

RESULTS

Comparison of Revision Quality and Post-test of Writing Proficiency Between Students Receiving Teacher Feedback and Automated Feedback

As the revision quality and the post-test of writing proficiency used two different writing tasks, the result of the within-subjects effect of the 2×2 mixed ANOVA was not relevant to the current study. The result of the interaction effect between writing occasion and feedback type was significant [$F(1, 68) = 10.93$,

TABLE 1 | Examples of teacher feedback by types of errors.

Types of errors		Errors	Teacher feedback
Mechanics	Spelling	There are six girls in my <u>dormatory</u> .	a spelling error
	Punctuations	In a word ^ my college life is busy.	an issue with punctuation
	Capitalization	All <u>My</u> classmates love playing table tennis.	capitalization problem
Vocabulary	Collocation	She always looks <u>for</u> me as a big sister.	a problem with collocation
	Word choice	Reading is a good way to <u>increase</u> my English proficiency.	a problem with word choice: "improve" should be used.
Grammar		<u>Receive</u> positive words from my friends is very important for me.	A noun is needed.

The underlined sections indicate errors in students' writing.

TABLE 2 | Summary of the data collection procedure.

Two classes	1st English session	2nd English session	3rd English session	4th English session
Teacher feedback	Students completed the first writing task (pre-test of English writing proficiency).	Students completed the second writing task and submitted their essays to the English teacher.	Students revised their essays using the teacher feedback. The revision was scored by Pigaiwang.	Students completed the third writing task (post-test of English writing proficiency).
Automated feedback	Students completed the first writing task (pre-test of English writing proficiency).	Students completed the second writing task and submitted their essays in the Pigaiwang.	Students revised their essays using the feedback from Pigaiwang. The revision was scored by Pigaiwang.	Students completed the third writing task (post-test of English writing proficiency).

$p < 0.01$, $\eta^2 = 0.19$], suggesting that the patterns of students' scores on revision quality and the post-test of writing proficiency were different by feedback type (see **Figure 1**).

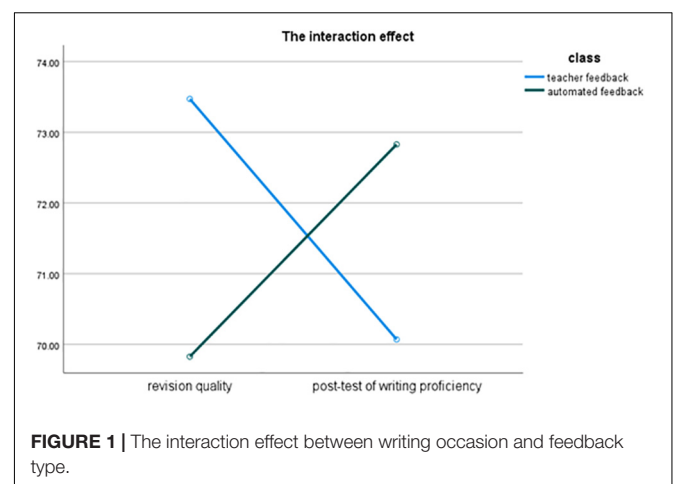
Separate one-way ANOVAs were conducted for revision quality and post-test of writing proficiency. For revision quality, Levene's test showed that the assumption of homogeneity of variances was met [$F(1,68) = 1.15$, $p = 0.29$]. The one-way ANOVA demonstrated that [$F(1, 68) = 8.76$, $p < 0.01$, $\eta^2 = 0.11$] students who received teacher feedback ($M = 73.47$, $SD = 5.59$) scored significantly higher than those who received automated feedback ($M = 69.83$, $SD = 4.67$). For the post-test of writing proficiency, Levene's test also confirmed that the assumption of homogeneity of variances was not violated [$F(1,68) = 0.24$, $p = 0.63$]. The results of the one-way ANOVA showed that the pattern was reversed [$F(1, 68) = 4.02$, $p < 0.05$, $\eta^2 = 0.06$]. Students in the automated feedback class ($M = 72.83$, $SD = 5.54$) received significantly higher scores than their peers in the teacher feedback class ($M = 70.07$, $SD = 5.59$).

As the students in the two classes did not differ in terms of their pre-test of English writing proficiency, the significantly better revision scores of the students who received teacher feedback suggested that teacher feedback was more effective in helping students revise their essays. In contrast, the significantly higher post-writing proficiency of the students who received automated feedback indicated that automated feedback might be more effective in developing FL learners' writing proficiency.

Comparison of Perceived Usefulness and Perceived Ease of Use Between Students Receiving Teacher Feedback and Automated Feedback

The results of the MANOVA found that the feedback type was significant [$F(2, 67) = 9.34$, $p < 0.01$; Wilk's $\Lambda = 0.78$, partial

$\eta^2 = 0.22$]. Levene's tests confirmed that the assumption of homogeneity of variances was met for both perceived usefulness scale [$F(1,68) = 0.29$, $p = 0.59$] and perceived ease of use scale [$F(1,68) = 0.20$, $p = 0.65$]. Univariate tests showed that students in the two different feedback conditions differed significantly on both perceived usefulness [$F(1, 68) = 18.42$, $p < 0.01$, $\eta^2 = 0.21$] and perceived ease [$F(1, 68) = 4.19$, $p < 0.05$, $\eta^2 = 0.06$] scales. Specifically, students who received teacher feedback ($M = 3.97$, $SD = 0.44$) had significantly higher ratings on the perceived usefulness scale than their peers who received automated feedback ($M = 3.46$, $SD = 0.54$). Students who received teacher feedback ($M = 4.05$, $SD = 0.58$) also had significantly higher ratings on the perceived ease of use scale than their counterparts who received automated feedback ($M = 3.75$, $SD = 0.62$).



Perceptions of the Strengths and Weaknesses of Teacher Feedback and Automated Feedback

The participants had mixed feelings toward both teacher feedback and automated feedback. **Table 3** summarizes the themes of the strengths and weaknesses of both types of feedback and the frequency of each theme. It should be noted that the total frequency of all the themes did not equate to the number of students, as some students only wrote about the strengths or the weaknesses, while some noted down more than one response to the strengths and the weaknesses.

The most frequently mentioned (mentioned by 20 students) strength of teacher feedback was that teacher feedback had balanced comments on both the positive and negative aspects of the students' writing. For instance, a student commented: "I like teacher feedback because it not only pointed out the problems and mistakes in my writing, but also included good comments on my essay. To me, this is really important, because these good words made me more confident about my English writing proficiency and encouraged me to make efforts to revise my essay." The second frequent strength of teacher feedback was the encouraging words used by the English teacher (mentioned by 17 students). These students believed that these encouraging words in teacher feedback enhanced their motivation and fostered their enthusiasm for English writing in the future, including the subsequent revisions. The third frequently mentioned strength of teacher feedback (mentioned by 12 students) was clarity and easiness of the language use, which, according to students, was easy to comprehend and hence can improve the efficiency of their revising process.

When looking at the positive comments on automated feedback, we found that the participants predominantly focused on automated feedback's ability to identify errors (mentioned by 32 students). An example response was: "One of the positive aspects of automated feedback is that it identifies the errors in my writing, by correcting these errors, my revised essay would be improved." This answer seems to suggest that the majority of students receiving automated feedback used such feedback as a mistake identification tool for them to fix errors in their English writing. The second most frequently mentioned strength (mentioned by 16 students) of automated feedback was that it provided suggestions for synonyms and detailed explanations of the differences between the synonyms. For instance, a student made such a comment: "Automated feedback is particularly good at offering multiple synonyms for me to choose. So I can use different words in my writing rather than always repeat the same word. I feel that this kind of feedback generated by Pigaiwang can enlarge my vocabulary size." The next most frequently mentioned strength (mentioned by 11 students) of automated feedback was its ability to provide a rank, which allowed the students to know their writing ability in relation to their fellow students.

In terms of the shortcomings, only a few students mentioned issues in teacher feedback. Four students believed that the feedback received from the English teacher sometimes lacked detailed explanations, as the teacher simply underlined the sentences or highlighted the words. Four students commented

that there were too many comments from the English teacher. In contrast, as many as 21 students pointed out that the comments generated by automated feedback were not always straightforward and comprehensible, which created barriers for them to revise their essays properly. Nine students mentioned that automated feedback emphasized too much on mechanical problems, such as punctuation and capitalization problems.

DISCUSSION AND CONCLUSION

The Effects of Teacher Feedback and Automated Feedback on the Cognitive Aspects of Foreign Language Writing (Revision Quality and Writing Proficiency Development)

In terms of the effects of teacher feedback and automated feedback on the cognitive aspects of FL writing, we found that students who received teacher feedback scored significantly higher on revision quality than those who received automated feedback, whereas students in the automated feedback class showed better performance on the post-test of their writing proficiency. The different effects of the two types of feedback on revision quality and on writing proficiency development suggest that teacher feedback and automated feedback may play different roles in helping FL writers revise and enhance their writing proficiency.

One of the possible reasons for the better effect of teacher feedback on the revision quality could be the low level of the English writing proficiency of our participants. Research has shown that students with low writing proficiency tend to overly rely on teacher feedback in the revision process (Zhang, 2020). The limited effect of teacher feedback on developing participants' writing proficiency could be that our participants might directly copy teacher feedback in the revision without knowing their problems, as shown in previous research (Lee, 2007; Zhao, 2010). Without knowing the sources of errors, participants would make the same mistakes again in their subsequent writing.

The reason for the better effect of automated feedback on developing our participants' writing proficiency may lie in its capacity to offer suggestions for synonyms, which may have enlarged our participants' vocabulary in the long run. Past research has reported that FL learners' receptive and productive vocabulary size is strongly associated with their writing proficiency, as a large vocabulary size allows students to express richer ideas in writing (Staehr, 2008; Shi and Qian, 2012; Lee, 2014; Lin, 2015; Miralpeix and Muñoz, 2018). For instance, Lee found that Korean university EFL learners' productive vocabulary size had significant effects on multiple aspects of their writing, including both content and language aspects. Similarly, among 67 Hong Kong university EFL learners, Lin reported that students' performance on the two vocabulary tests, namely Vocabulary Levels Test and Word Associates Test, could explain a quarter of their English writing performance. As we did not measure the vocabulary change of our participants in the automated feedback class, whether students who received

TABLE 3 | Themes of the strengths and weaknesses of teacher feedback and automated feedback.

	Strengths	Frequency	Weaknesses	Frequency
Teacher feedback	a balanced comments on both the positive and negative aspects of students' writing	20	lacking detailed explanations	4
	encouraging words	17	too many comments	4
	clarity and easiness of the language use	12		
Automated feedback	the ability to identify errors	32	ambiguous and incomprehensible	21
	the capacity to suggest for synonyms	16	too much emphasis on mechanical problems, such as punctuation and capitalization problems	9
	the function to rank essays to allow students to know their ranks in relation to their classmates	11		

automated feedback performed better on the post-test of the writing proficiency was related to their increased vocabulary needs further verification.

The Effects of Teacher Feedback and Automated Feedback on the Psychological Aspects of Foreign Language Writing (Students' Perceived Usefulness and Ease of Use, Their Perceptions of the Strengths and Weaknesses of the Feedback)

In terms of the effects of the two types of feedback on the psychological aspects of FL writing, we found that students in the teacher feedback class had significantly higher ratings on perceived usefulness and ease of use of the feedback than those in the automated feedback class. The qualitative responses from the open-ended questionnaire also reflected that, in general, the students hold more positive perceptions toward teacher feedback than toward automated feedback. One great barrier which prevents students from utilizing the comments generated by the Pigaiwang is the students' incapability of comprehending the comments. Even though the students have positive comments on Pigaiwang's feature of offering multiple synonyms, this does not mean that they know how to select the most appropriate word from these synonyms in the context of their writing. Students' lack of ability to use the feedback from Pigaiwang in proper ways may have inhibited them from effectively incorporating the feedback into the revision. This may also offer some explanations as to why the revision quality in the automated feedback condition was poorer than that in the teacher feedback condition.

As no previous research has compared students' ratings on perceived usefulness and ease of use between students receiving teacher feedback and automated feedback, it is unsure if the results found in our study represent a general pattern. It should be noted that the participants in our study are first-year university students who have just commenced their university learning. This means that our participants may lack learning autonomy due to the duck-feeding teaching style in Chinese high schools (Li and Baldauf, 2011). This may affect their perceptions of the usefulness and ease of use of the Pigaiwang as using

automated feedback requires learners to have some levels of learning autonomy so that they can sustainably interact with the machine (Warschauer and Ware, 2006; Lee, 2017). Future research should be conducted with more mature Chinese EFLs, to examine their perceptions of the usefulness and ease of use of the Pigaiwang.

Pedagogical Implications

The results of the study have some pedagogical implications for FL writing. In order to reduce teachers' workload, college English teachers may consider using a combination of teacher and automatic feedback in FL writing classes or using the two types of feedback in rotation. As suggested by Zhang (2020), teachers should use automated feedback as "a good supplement to writing instruction and a helpful writing assistance tool in the writing and revising process" (p. 12). They may use automatic feedback to check the language errors of students' drafts, such as spelling, punctuation, and grammar, and give students feedback on the contents and organization of their essays. Teachers should also make students fully aware of the advantages and disadvantages of automated feedback (Reynolds et al., 2021). For instance, our students commented positively about Pigaiwang's function of providing alternative lexical items. Teachers should instruct students how to use such a function to learn vocabulary, which has been shown to be positively associated with FL learners' writing proficiency (Staehr, 2008; Shi and Qian, 2012; Lee, 2014; Lin, 2015; Miralpeix and Muñoz, 2018).

Teachers should also consider organizing a workshop before asking students to use Pigaiwang in order for students to maximize the usefulness of the features in the Pigaiwang. In the workshop, teachers should explain all the useful functions and demonstrate the appropriate ways to use them through some concrete examples. Teachers may also need to explain different types of comments provided by Pigaiwang, such as what collocation problems are. Through this kind of workshop, students will become more prepared and more confident to navigate through the automated feedback platform, which will in turn encourage them to actively use the automated feedback during their writing and revising processes.

Limitations and the Directions for Future Research

When interpreting the results, some limitations of the study should be kept in mind. First, our study was a relatively small-scale study, which only involved 35 participants in each feedback condition. In addition, all the participants were recruited from a single university. These limitations in sampling limit the generalizability of the study. Future research should increase the number of participants and recruit participants from different universities in order for the sample to be more representative. Second, we did not include a control group in our study, as apart from comparing students' writing proficiency development, we also aimed to compare students' revision quality, perceived usefulness, and ease of use of the feedback, as well as their perceptions of the strengths and weaknesses of the feedback, all of which required students to receive some forms of feedback. However, without a control group, it was difficult to rule out the possibility that students' writing proficiency development is a result of their English learning rather than the feedback they received. The design of future studies will be significantly improved by including a control group.

Third, it should be noted that although the reliability coefficients of the perceived usefulness scale and perceived ease of use scale were all above 0.60, which was acceptable (Hair et al., 2010), they were slightly lower than the more commonly used 0.70, possibly due to the small sample size. Thus, cautions need to be taken to interpret the results related to the scales. Last but not least, while we asked the English teacher to give the feedback by covering the aspects similar to those provided in automated feedback, we did not in fact compare if the feedback provided by the teacher and Pigaiwang matched in terms of the aspects they covered. Therefore, it is unknown if the different effects of the two types of feedback on revision quality were influenced by different aspects of the two types of feedback covered. This limitation should be addressed in future research.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of the ethics requirements of Shaanxi Xueqian Normal

University. Requests to access the datasets should be directed to the Human Research Ethics Committee of Shaanxi Xueqian Normal University: kyc@snsy.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the School of Foreign Languages in Shaanxi Xueqian Normal University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ZW and FH contributed substantially to the conception of the work, acquisition, analysis, interpretation of the data, drafted the work and revised it critically for important intellectual content, approved the final version of the manuscript to be published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Chinese Ministry of Education Industry-Academia Cooperation Collaborative Education Project (Grant Number: 202102591010) and the Educational Reform Project of Shaanxi Xueqian Normal University (Grant Number: 21JG007Y).

ACKNOWLEDGMENTS

We wish to express their gratitude to the participants for their voluntary participation in the study.

REFERENCES

- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: is content feedback followed by form feedback the best method? *J. Sec. Lang. Writ.* 9, 227–257. doi: 10.1016/S1060-3743(00)00027-8
- Bai, L., and Hu, G. (2017). In the face of fallible automated feedback feedback: how do students respond? *Educ. Psychol.* 37, 67–81. doi: 10.1080/01443410.2016.1223275
- Biber, D., Nekrasova, T., and Horn, B. (2011). *The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis, TOEFL iBT RR-11-05*. Princeton: Educational Testing Service. doi: 10.1002/j.2333-8504.2011.tb02241.x
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *J. Sec. Lang. Writ.* 17, 102–118. doi: 10.1016/j.jslw.2007.11.004
- Bitchener, J., and Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. Milton Park: Routledge. doi: 10.4324/9780203832400
- Bitchener, J., and Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Lang. Teach. Res.* 12, 409–431. doi: 10.1177/1362168808089924
- Bitchener, J., and Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT J.* 63, 204–211. doi: 10.1093/elt/ccn043
- Bitchener, J., and Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *J. Second Lang. Writ.* 19, 207–217. doi: 10.1016/j.jslw.2010.10.002
- Boubekeur, S. (2015). Teaching and learning writing through providing teacher's feedback. *Internat. J. Appl. Ling. Engl. Literat.* 4, 16–22. doi: 10.7575/aiac.ijalel.v4n6p.16

- Calvo, R., and Ellis, R. (2010). Students' conceptions of tutor and automated feedback in professional writing. *J. Eng. Educ.* 99, 427–438. doi: 10.1002/j.2168-9830.2010.tb01072.x
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *J. Sec. Lang. Writ.* 12, 267–296. doi: 10.1016/S1060-3743(03)00038-9
- Chapelle, C., Cotos, E., and Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Lang. Test.* 32, 385–405. doi: 10.1177/0265532214565386
- Chen, C., and Cheng, W. (2008). Beyond the design of automated writing evaluation: pedagogical practices and perceived learning effectiveness in EFL writing classes. *Lang. Learn. Technol.* 12, 94–112.
- Chevillat, J. (2004). Automated scoring technologies and the rising influence of error. *Engl. J.* 93, 47–52. doi: 10.2307/4128980
- Conrad, S. M., and Goldstein, L. M. (1999). ESL student revision after teacher-written comments: text, contexts, and individuals. *J. Sec. Lang. Writ.* 8, 147–179. doi: 10.1016/S1060-3743(99)80126-X
- Cotos, E. (2011). Potential of automated writing evaluation feedback. *Calico J.* 28, 420–459. doi: 10.11139/cj.28.2.420-459
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 1989, 319–340. doi: 10.2307/249008
- Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Manag. Sci.* 35, 982–1003. doi: 10.1287/mnsc.35.8.982
- Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., et al. (2019). What makes for effective feedback: staff and student perspectives. *Assess. Eval. High. Educ.* 44, 25–36. doi: 10.1080/02602938.2018.1467877
- Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. *Assess. Writ.* 18, 7–24. doi: 10.1016/j.asw.2012.10.002
- Dikli, S. (2010). The nature of automated essay scoring feedback. *Calico J.* 28, 99–134. doi: 10.11139/cj.28.1.99-134
- Dikli, S., and Bley, S. (2014). Automated essay scoring feedback for second language writers: how does it compare to instructor feedback? *Assess. Writ.* 22, 1–17. doi: 10.1016/j.asw.2014.03.006
- Esfandiari, F., Yaqubi, B., and Marzban, A. (2014). Learning corrected target forms: Students' corrective feedback through teacher's written corrective feedback in Iranian EFL context. *Theory Pract. Lang. Stud.* 3, 504–512. doi: 10.4304/tpls.4.3.504-512
- Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *J. Educ. Technol. Soc.* 13, 246–256.
- Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assess. Eval. High. Educ.* 36, 51–62. doi: 10.1080/02602930903197883
- Ferris, D. (2011). *Treatment of error in second language student writing*. Ann Arbor: University of Michigan Press. doi: 10.3998/mpub.2173290
- Ferris, D., Liu, H., Sinha, A., and Senna, M. (2013). Written corrective feedback for individual L2 writers. *J. Sec. Lang. Writ.* 22, 307–329. doi: 10.1016/j.jslw.2012.09.009
- Ferris, D., and Roberts, B. (2001). Error feedback in L2 writing classes: how explicit does it need to be? *J. Sec. Lang. Writ.* 10, 161–184. doi: 10.1016/S1060-3743(01)00039-X
- Ferris, D. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Q.* 29, 33–53. doi: 10.2307/3587804
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Q.* 31, 315–339. doi: 10.2307/3588049
- Ferris, D. R. (2006). "Does error feedback help student writers?" in *Feedback in second language writing: Contexts and issues*, eds K. Hyland and F. Hyland (New York, NY: Cambridge University Press), 81–104. doi: 10.1017/CBO9781139524742.007
- Grimes, D., and Warschauer, M. (2010). Utility in a fallible tool: a multi-site case study of automated writing evaluation. *J. Technol. Learn. Assess.* 8:6.
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. London: Pearson Education.
- He, X. L. (2013). Reliability and validity of the assessment of the Pigaiwang on college students' writings. *Mod. Educ. Technol.* 23, 64–67.
- Hedgcock, J., and Lefkowitz, N. (1996). Some input on input: two analyses of student response to expert feedback in L2. *Mod. Lang. J.* 80, 287–308. doi: 10.1111/j.1540-4781.1996.tb01612.x
- Hu, X. W. (2015). The effect of self-modification on college English writing based on automated essay evaluation. *Comp.-Assis. Foreign. Lang. Educ.* 3, 45–49.
- Hyland, F. (1998). The impact of teacher written feedback on individual writers. *J. Sec. Lang. Writ.* 7, 255–286. doi: 10.1016/S1060-3743(98)90017-0
- Hyland, F. (2001). Providing effective support: Investigating feedback to distance language learners. *Open Learn.* 16, 233–247. doi: 10.1080/02680510120084959
- Hyland, K., and Hyland, F. (2006). Feedback on second language students' writing. *Lang. Teach.* 39, 83–101. doi: 10.1017/S0261444806003399
- Kim, P. H. (2015). *Feedback and eevision: A self-assessment intervention*. New York, NY: Columbia University.
- Lee, I. (2005). Error correction in the L2 classroom: what do students think? *TESL Can. J.* 22, 1–16. doi: 10.18806/tesl.v22i2.84
- Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: assessment for learning or assessment of learning? *Assess. Writ.* 12, 180–198. doi: 10.1016/j.asw.2008.02.003
- Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *J. Sec. Lang. Writ.* 17, 144–164. doi: 10.1016/j.jslw.2007.12.001
- Lee, I. (2013). Research into practice: written corrective feedback. *Lang. Teach.* 46, 108–119. doi: 10.1017/S0261444812000390
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer. doi: 10.1007/978-981-10-3924-9
- Lee, S. (2014). Korean EFL university students' English vocabulary size and their writing proficiency. *J. Korea Engl. Educ. Soc.* 13, 21–43. doi: 10.18649/jkees.2014.13.2.21
- Li, J. (2014). Examining genre effects on test takers' summary writing performance. *Assess. Writ.* 22, 75–90. doi: 10.1016/j.asw.2014.08.003
- Li, J., Link, S., and Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (automated feedback) feedback in ESL writing instruction. *J. Sec. Lang. Writ.* 27, 1–18. doi: 10.1016/j.jslw.2014.10.004
- Li, M., and Baldauf, R. (2011). Beyond the curriculum: a Chinese example of issues constraining effective English language teaching. *TESOL Q.* 45, 793–803. doi: 10.5054/tq.2011.268058
- Liao, H. (2016). Using automated writing evaluation to reduce grammar errors in writing. *ELT J.* 70, 308–319. doi: 10.1093/elt/ccv058
- Lin, H. F. (2015). *Roles of depth of vocabulary knowledge in EFL learner's writing proficiency*. Unpublished doctoral thesis. Hong Kong: Hong Kong Polytechnic University.
- Link, S., Mehrzad, M., and Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Comp. Assis. Lang. Learn.* 2020, 1–30.
- Luo, Y., and Liu, Y. (2017). Comparison between peer feedback and automated feedback in college English writing: a case study. *Open J. Mod. Ling.* 7, 197–215. doi: 10.4236/ojml.2017.74015
- Manap, M. R., Ramli, N. F., and Kassim, A. A. M. (2019). Web 2.0 automated essay scoring application and human ESL essay assessment: a comparison study. *Eur. J. Engl. Lang. Teach.* 5, 146–162.
- Miralpeix, I., and Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *Internat. Rev. Appl. Ling. Lang. Teach.* 56, 1–24. doi: 10.1515/iral-2017-0016
- Neo, M., Park, H., Lee, M. J., Soh, J. Y., and Oh, J. Y. (2015). Technology acceptance of healthcare E-learning modules: a study of Korean and Malaysian students' perceptions. *Turk. Online J. Educ. Tech. TOJET* 14, 181–194.
- Page, E. (2003). *Project essay grade: PEG. Automated essay scoring: A cross-disciplinary perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Radecki, P. M., and Swales, J. M. (1988). ESL student reaction to written comments on their written work. *System* 16, 355–365. doi: 10.1016/0346-251X(88)90078-4
- Ramineni, C. (2013). Validating automated essay scoring for online writing placement. *Assess. Writ.* 18, 40–61. doi: 10.1016/j.asw.2012.10.005
- Reynolds, B. L., Kao, C. W., and Huang, Y. Y. (2021). Investigating the effects of perceived feedback source on second language writing performance: a quasi-experimental study. *Asia-Pacific Educ. Res.* 30, 585–595. doi: 10.1007/s40299-021-00597-3

- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT J.* 59, 23–30. doi: 10.1093/elt/ccii003
- Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Ling. Educ.* 29, 73–82. doi: 10.1016/j.linged.2014.12.001
- Saito, H. (1994). Teachers' practices and students' preferences for feedback on second language writing: a case study of adult ESL learners. *TESL Can. J.* 11, 46–70. doi: 10.18806/tesl.v11i2.633
- Semke, H. D. (1984). Effects of the red pen. *For. Lang. Ann.* 17, 195–202. doi: 10.1111/j.1944-9720.1984.tb01727.x
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL Learners' acquisition of articles. *TESOL Q.* 41, 255–283. doi: 10.1002/j.1545-7249.2007.tb00059.x
- Shermis, M., and Burstein, J. (2003). *Automated essay scoring: A cross-disciplinary perspective*. London: Routledge. doi: 10.4324/9781410606860
- Shi, L., and Qian, D. (2012). How does vocabulary knowledge affect Chinese EFL learners' writing quality in web-based Settings?—Evaluating the relationships among three dimensions of vocabulary knowledge and writing quality. *Chin. J. Appl. Ling.* 35, 117–127. doi: 10.1515/cjal-2012-0009
- Silver, R., and Lee, S. (2007). What does it take to make a change? Teacher feedback and student revisions. *Engl. Teach.* 6, 25–49.
- Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Lang. Learn. J.* 36, 139–152. doi: 10.1080/09571730802389975
- Stevenson, M., and Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assess. Writ.* 19, 51–65. doi: 10.1016/j.asw.2013.11.007
- Tang, Y. M., Chen, P. C., Law, K. M., Wu, C. H., Lau, Y. Y., Guan, J., et al. (2021). Comparative analysis of Student's live online learning readiness during the coronavirus (COVID-19) pandemic in the higher education sector. *Comp. Educ.* 168:104211. doi: 10.1016/j.compedu.2021.104211
- Venkatesh, V., and Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Dec. Sci.* 39, 273–315. doi: 10.1111/j.1540-5915.2008.00192.x
- Venkatesh, V., and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field 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 a unified view. *MIS Q.* 2003, 425–478. doi: 10.2307/30036540
- Wang, F., and Wang, S. (2012). A comparative study on the influence of automated evaluation system and teacher grading on students' English writing. *Proc. Eng.* 29, 993–997. doi: 10.1016/j.proeng.2012.01.077
- Wang, H. J. (2016). An empirical research into scoring validity of AES. *J. Zhej. Univ. Technol. Soc. Sci.* 15, 89–93.
- Wang, P. (2013). Can automated writing evaluation programs help students improve their English writing? *Internat. J. Appl. Ling. Engl. Literat.* 2, 6–12. doi: 10.7575/ijalel.v.2n.1p.6
- Warden, C. (2000). EFL business writing behavior in differing feedback environments. *Lang. Learn.* 50, 573–616. doi: 10.1111/0023-8333.00141
- Warschauer, M., and Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogie Internat. J.* 3, 22–36. doi: 10.1080/15544800701771580
- Warschauer, M., and Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Lang. Teach. Res.* 10, 157–180. doi: 10.1191/1362168806lr190oa
- Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assess. Eval. High. Educ.* 31, 379–394. doi: 10.1080/02602930500353061
- Wilson, J., and Andrada, G. N. (2016). *Handbook of research on technology tools for real-world skill development*. Hershey, PA: IGI Global.
- Wilson, J., and Czik, A. (2016). Automated essay evaluation software in English Language Arts classrooms: effects on teacher feedback, student motivation, and writing quality. *Comp. Educ.* 100, 94–109. doi: 10.1016/j.compedu.2016.05.004
- Wilson, J., Olinghouse, N., and Andrada, G. (2014). Does automated feedback improve writing quality? *Learn. Disabil.* 12, 93–118.
- Yang, X. Q., and Dai, Y. C. (2015). Practice research on self-regulated writing model of college English teaching based on Pigai Network. *Comp. Assis. For. Lang. Educ.* 2, 17–23.
- Zamel, V. (1985). Responding to student writing. *TESOL Q.* 21, 697–715. doi: 10.2307/3586990
- Zhan, L. (2016). Written teacher feedback: Student perceptions, teacher perceptions, and actual teacher performance. *Engl. Lang. Teach.* 9, 73–84. doi: 10.5539/elt.v9n8p73
- Zhang, Z. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assess. Writ.* 43, 1–14. doi: 10.1016/j.asw.2019.100439
- Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: a comparative study in a Chinese English writing classroom. *Assess. Writ.* 15, 3–17. doi: 10.1016/j.asw.2010.01.002
- Zyad, H. (2016). Integrating Computers in the Classroom: Barriers and teachers' attitudes. *Internat. J. Instruct.* 9, 65–78. doi: 10.12973/iji.2016.916a

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 Wang and Han. 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
Jesus de la Fuente,
University of Navarra, Spain

REVIEWED BY
Alisa Belzer,
Rutgers, The State University
of New Jersey, United States
Jörg Schwarz,
Helmut Schmidt University, Germany

*CORRESPONDENCE
Aydin Yücesan Durgunoğlu
adurguno@d.umn.edu

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

RECEIVED 13 April 2022
ACCEPTED 27 July 2022
PUBLISHED 12 August 2022

CITATION
Durgunoğlu AY, Cantürk M, Kaya U,
Yazıcı B and Şahin K (2022) Moving
to remote learning in adult education:
Challenges and solutions of limited
technological resources
and capabilities.
Front. Educ. 7:919639.
doi: 10.3389/feduc.2022.919639

COPYRIGHT
© 2022 Durgunoğlu, Cantürk, Kaya,
Yazıcı and Şahin. 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.

Moving to remote learning in adult education: Challenges and solutions of limited technological resources and capabilities

Aydin Yücesan Durgunoğlu^{1*}, Meltem Cantürk², Uğur Kaya²,
Betül Yazıcı² and Kübranur Şahin²

¹Department of Psychology, University of Minnesota Duluth, Duluth, MN, United States, ²Anne Çocuk Eğitim Vakfı (Mother Child Education Foundation), Istanbul, Turkey

COVID-19 pandemic has caused disruptions in educational systems; hence, educators have developed innovative ways to address this new reality. Around the world, many adult education programs try to reach women who could not complete their education because of poverty and sociocultural issues, such as patriarchal barriers. COVID-19 was especially challenging for such lifelong learning programs because they operate with limited technological resources and work with learners with weaker digital proficiencies. In this paper, we describe how we rapidly adapted our face-to-face empowerment program for women to remote learning despite technological limitations. We also discuss the outcomes as evaluated by pre- and post-tests, participant observations and interviews. We finish by summarizing the successes and challenges of addressing this new reality in education.

KEYWORDS

literacy, numeracy, lifelong learning, adult education, women's empowerment, remote learning

Introduction

Around the world, a sizeable group of women could not complete their education because of poverty and sociocultural issues, especially patriarchal barriers. Many lifelong learning and adult education programs try to reach those women, because when women are more educated and participate more fully in their communities, there are significant impacts on the wellbeing of children and families, the economic vitality of a country and even on reducing conflict (Kabeer, 1999; LeVine et al., 2011; Weitzman, 2018).

COVID-19 pandemic was a serious setback for all students in formal education, but it was devastating for non-formal education programs because they operate with limited technological resources and usually work with learners with weaker digital proficiencies. In this paper, we describe our experiences in how we rapidly adapted a face-to-face

program to remote learning for a disadvantaged population. We share what we learned about moving to remote learning under technologically limited conditions and what the outcomes were. We also summarize the successes and challenges of addressing this new reality in education. Although there is a growing literature on standards for and the techniques to develop the digital literacy skills of adults (Vanek and Harris, 2020), the pandemic forced us to adapt the program with no time to teach learners about digital tools and instead to rely on the existing resources and capabilities of the women.

Under the umbrella of Mother Child Education Foundation (AÇEV), a non-governmental organization in Turkey, we have developed and implemented many literacy and numeracy programs for women (Durgunoğlu et al., 2003). A recent extension added a strong empowerment component to existing literacy and numeracy programs. The new program called Hayat Dolu Buluşmalar (POWER Program) targeted young women from low socio-economic backgrounds, with less than 8 years of schooling. The goal was to help women to get empowered to move themselves and their communities forward with the support of non-formal education. We assume that empowerment requires certain capabilities, such as literacy, numeracy, critical thinking and technology skills so that women can access societal resources (health, education, employment, etc.), express themselves in their communities, and participate in decision-making processes. Therefore, we embedded the empowerment topics (e.g., valuing self as a human being, women's rights, using technology effectively, preventing child-age marriage and childbearing) into reading, writing, critical thinking and math activities in the group. POWER Program consisted of 18 weekly face-to-face sessions, each lasting 3 h, across a 4-month period. We gave detailed pre- and post-tests to assess growth.

In each group, there was first a discussion of the topic of that session (e.g., violence against women), asking women their observations and thoughts; thus, activating women's background knowledge. Then, a text from the program textbook was read and critically discussed, followed by deep comprehension and application exercises (e.g., busting myths about why violence occurs and ways to address it). Next were math exercises and the discussion of the assignment for next week.

The first evaluation of the POWER Program (Durgunoğlu and AÇEV, 2018) indicated significant increases in the women's literacy and numeracy skills, self-confidence and participation in decision-making at home (Öztan et al., 2021).

However, with COVID-19, these in-person classes had to be discontinued. Because most women did not have computers or internet connection in their homes, but had smartphones; we rapidly pivoted to a WhatsApp (WA)-based version of the program using smartphones (Kavanaugh et al., 2013) and started implementing it. This new version had the following characteristics: (a) We had small groups of three to seven

women meeting *via* WA for 45 min every week because of the group size limit of WA (b) We had to reduce the number of sessions from 18 to 10; thus, had a program that lasted two and half months This was partly due to the constraints of the workplace and partly because the small group sizes required more teacher effort (c) A telecommunications grant enabled us to provide mobile data plans so that women could use WA without hurting their budgets (d) We curated existing YouTube videos on empowerment and math topics and asked women to watch them at home to supplement the work we did in the WA groups. (e) Before the program started, we mailed the textbook and other relevant paperwork to the participants. (f) During each online WA meeting, as before, women read the anchor text from the textbook, answered its comprehension questions and discussed the main points of the topic, followed by an applied exercise about the topic and finally reviewed the math topic of the week and reviewed the and the math exercises in the textbook. (g) The pre- and post-tests had to be shortened and were given by testers over the phone.

In this paper we first describe the qualitative and quantitative data from two POWER Program groups, as well as the observations of the WA meetings. In addition to short pre- and post-tests that were given to the participants, the first author joined the WA discussions every week and took extensive field notes. She also interviewed the participants from one of the groups 2–3 months after the education program had ended.

Methods

Participants

Participants were all women from two different cities in Turkey. Group 1 ($n = 6$) was from a city in the southeast and Group 2 ($n = 6$) was from a city in the northwest. Women in Group 2 were working full time in a textile factory, whereas the women in Group 1 were not employed outside of the home. Across both groups, only half of the women reported having a computer at home. However, all, except one, had smartphones and internet access (The one who did not have a smartphone borrowed a family member's phone during class time). Across both groups, the age of women ranged from 20–47 ($M = 34$). Except for one, all women had children ($M = 2.3$). All had completed between 5 and 8 years of education ($M = 6.75$ years). Of the 10 married women, the husbands had completed between 5 and 16 years of education ($M = 8.8$). When asked why their education was interrupted, only two reported being uninterested in school. The others described economic difficulties, families preventing them from going to school, and immigrating from another country. When asked why they were interested in attending this education program, nine said for improving themselves, learning new things, developing their self-confidence, doing something different. Three were

hoping to better support their children. At the end of the 2.5 months, three women in the second group did not complete the education program. One had a family emergency. The second left to get married (but later attended a second cohort). The third was moving to a different city. Both groups were taught by experienced POWER Program educators who had taught in-person POWER Program groups multiple times, so they only needed to learn about remote delivery as they were already familiar with the content and the approach of the program.

Materials and procedure

Before the education program started, we gave the participants a demographics questionnaire and a pre-test on the phone. We gave the pretest by including the paper copy of the test within the education program materials that were sent to the learners' homes. The learners were not given any indication that this was a test or that it should be examined ahead of time. We called the participants and asked them to take out the paper and give us the answers as we talked with them. The pre-test had two reading comprehension questions. In the first question, learners first read two sentences about a girl who wanted to fly, followed by a question about why she wanted to do that. The second question had three sentences describing the similarities between bird and reptile eggs and asked learners to summarize the two similarities. There were two general knowledge questions about societal support systems for women. One asked about what one can do upon hearing a neighbor experiencing domestic violence. The second question asked about how to help a 14-year-old who wanted to continue her education, but whose father wanted her to get married. On math, there were six questions, addition, subtraction, multiplication, and two problems: (a) calculating the new salary of a worker when given a 10% raise and (b) expressing as a fraction how much of the cake in the picture was eaten. The same tests were given at the end of the education program as post-tests.

Data analysis

Quantitative data

Reading, math and knowledge pre- and post-test scores were compared (see [Table 1](#)). Because the sample sizes are too small, the data were analyzed using non-parametric statistics. On the reading test, seven of the nine learners showed positive changes (Wilcoxon signed-rank test showed $Z = -1.807$, $p = 0.071$). The median scores were four for pre-test and five for post-test. Math and knowledge tests did not show significant changes. This is partly due to the fact that the pre-test scores were already at ceiling levels and there was no room for growth. Hence, shortening the tests to deliver over the phone led to truncated ranges.

Qualitative data: Discussions during the WhatsApp meetings

Most interesting discussions occurred when women shared their experiences and opinions, both before and after reading the empowerment texts. Brief examples are given below to illustrate that although they were in a WA meeting, women were able to discuss deeply some difficult topics. The examples below are from both groups of women and are organized around some of the units in the curriculum (women's own statements are in italics; all names are pseudonyms).

Effective communication

All women reported becoming aware of the communication mistakes they have noted in themselves and in others, such as not listening and getting angry quickly. Women in Group 2 brought examples from the workplace *"We don't have effective communication in the workplace. Nobody listens to another; lots of stress, no patience."* They also complained about how line supervisors bark orders at some workers but treat others more kindly. They articulated how so many problems in the workplace occurred because of lack of respect and consideration for each other.

The devastating effects of child marriage

Mercan gave the example of her own mother who got married at a very early age and today still played with her grandchildren's dolls as she never had such toys of her own. Nuran talked about how she got married very early and regretted it.

Uses of technology and cyber bullying

The women shared many observations of youth spending a lot of money on games, and cases of identity theft and swindling.

Gender equity

Sevcin gave the following example to gender discrimination: *"A man said that they were altogether five siblings. In fact, there were nine children in the family, but he did not count his four sisters."*

Division of labor at home

This topic generated a lot of discussion and immediate results. Participants highlighted how fathers can be role models for overcoming traditional gender roles in division of labor in the homes. Many women integrated the lessons from effective communication into requesting more help around the house from their family members. One favorite topic was calculating the cost of unpaid care and domestic work, invisible labor that women provide at home.

Valuing oneself as a person

When they documented what they in one day, Necla said *"I realized that I never set aside a time for myself."* In the

TABLE 1 Mean scores on literacy, math and general knowledge pre- and post-tests.

Name *did not complete	Reading pre max = 6	Reading post max = 6	Math pre max = 18	Math post max = 18	Knowledge pre max = 6	Knowledge post max = 6
Hayat*	5		18		4	
Nurten	3	5	18	18	2	6
Türker	5	5	18	18	4	4
Sevcan	3	4	18	18	5	6
Necla*	5		15		6	
Züleyha*	3		15		1	
Tuğçe	4	6	15	15	5	6
Özcan	3	5	9	18	6	5
Ayşe	5	6	18	18	6	6
Serap	5	6	15	15	4	6
Mercan	5	6	18	18	6	6
Nuran	4	6	12	18	6	6
Number showing positive change	7/9 significant		2/9 not significant		4/9 not significant	
Number showing no change	1/9		7/9		4/9	
Number showing negative change	1/9		0/9		1/9	

*The names are pseudonyms.

unit about women-friendly communities, they expressed the need for more green space, easier access to healthcare, feeling safe on the streets.

Children's rights

The unit on keeping children in school and children's rights resonated. Sevcan gave the example *"I started working at 14. When I came home, I would go out and play with my friends."* They all mentioned their strong wish to get their daughters to stay in school. Tuğçe gave the example of how she was taken from middle school when her breasts started appearing. There were many good examples of using effective communication when interacting with children.

Domestic violence

In this unit, women discussed what constitutes violence (physical, verbal, economic, sexual). However, the discussion was always based on the text, no personal examples were provided in this unit.

Qualitative data: Interviews conducted 2 or 3 months later: Growth and empowerment

One question in the demographic survey given at the beginning of the program asked women what their goals and plans for the future were. In Group 1 no one had specified a goal for the future. In Group 2, who were women working in the textile sector, five out of six had some specific goals such as to continue with their education, get their driver's license. This indicated that Group 2 was already working outside of the house and had a better sense of personal development opportunities. Therefore, it was more informative to see what changed in Group 1. Several months after the education program ended,

we contacted the six women from Group 1 by phone. We asked if they saw any changes in themselves, if others noticed any changes in them and if they had some future goals. Although empowerment is very difficult to define and measure (Kabeer, 1999), perceived changes and growth; making plans for one's future, thus implying agency; and awareness of societal constraints can be indicators of empowerment. Hence the following quotes illustrate these themes.

Serap said that she is making no effort to continue with her education because she has young children. However, she stated that her communication with her children was now different. She asks both her son and daughter to help with housework (*"not my husband, he cannot change"*). She started reading her education program books while her children were reading theirs. She also stated that her self-confidence increased tremendously. She has shared what she learned with a close relative and reported that this relative also showed some positive changes in how she treats her children. She also stated that doing something for herself was really important.

Tuğçe: She is expecting her fourth child, so although she was planning to get her driver's license, she had to post-pone it. She regrets not continuing with her schooling. She is using better communication techniques with her children. However, when she tried to suggest some techniques to her husband, he said *"Are you my teacher now?"* She also has a 10-year-old who is very angry. When discussed more, it became obvious that as the daughter is beginning to develop, the family started keeping her in the house and she became angry when she could not play outside any longer and act like the child she really is.

Ayşe: She stated *"My self-confidence improved. I want to open a bakery. My husband owns his business and sells dried fruit and spices. He also sells on the Internet. I started making a regional delicacy, a filled cookie. I designed the boxes with 12 slots to put these cookies, so they won't break. We started selling them on the Internet."* (Note: On the Internet page these cookies are featured

prominently, and receive good feedback, but their baker is not named or credited). Ayşe also showed a strong awareness and action goals when she said *“I got married at 17 because of family pressure. Luckily my husband turned out to be a good man and supported me. I am very strongly against girls marrying young.”*

Nuran: *“I went through vocational education to become a private security officer. I also drive a motorcycle. With two friends, we applied to complete the classes to get our high school degree. I am the main breadwinner, so I am working hard. I convinced a friend and she got her middle school diploma.”*

Özcan: She has three children and one with a chronic illness. *“I could not take any steps for education, yet I am better at convincing people and making decisions”* She decided and bought a piece of furniture by herself and was amazed that nobody objected; in fact, they liked it. *“This amazed me and made me feel like I can do this. I encouraged my aunt in the village not to get her daughter married at a young age for dowry money.”*

Mercan: (This was from discussion she had with the tester during the posttest, not a separate phone interview) *“This education program added so much to my life. I wish it were longer. I have applied for my driver’s license and to take classes to get my middle school diploma.”*

Discussion

The qualitative and quantitative data showed that despite considerable challenges, this program was effective in empowering the women who participated. The learners had considerable barriers as they lived in cultural contexts which seriously limited their choices, educational opportunities, gave them heavy caregiving responsibilities, all under difficult economic conditions. The education program was short and delivered remotely with very limited technological resources. Despite these challenges, we addressed serious topics and could see significant changes in women’s literacy skills, communication proficiencies, decision-making abilities, and overall self-confidence.

Observed challenges and what worked

Limited technology resources

WA and YouTube were the two resources that women could access and use easily, given their existing social practices. Women were already using the Internet to shop, to look for medical information, etc. and using WA to interact with family members. Therefore, the first consideration was to repurpose for education the tools that were used for social and recreational needs. Women sent their assignments as attachments in WA messages. They took pictures of their written work and submitted those. The cost was another barrier and we gave weekly data plans to address that, while letting women know that it was provided weekly and only if they attended

consistently. WA meetings could be noisy, so the educators encouraged taking turns in a deliberate manner.

Multiple responsibilities of women

All women had significant responsibilities at home, shouldering the biggest share of housekeeping and caregiving tasks. This also meant that there was a lot of commotion at home. In Group 2 women also worked 6 days a week and believed that Sunday was the only time they could rest. We accepted some interference from the household as women participated in the weekly face-to-face meeting. For example, we observed Necla discretely breastfeeding her baby while participating in the discussion. We also emphasized that this was a special time for women to set aside for themselves, to do something for themselves as a human being rather than a mother/sister/wife/daughter. One important factor was flexibility. At times when they had family responsibilities, they requested changing the meeting days/times. Educators asked for feedback from the group and changed if all agreed.

Difficulties of independent work

This program required women to do some work on their own. They did not have enough time and energy to do that. Therefore, we relied on short videos and encouraged them to spread the work throughout the whole week. We also kept written homework to one or two short exercises every week. We used multimedia to make women’s lives easier, for example some assignments could be sent as audio or video recordings, instead of written. As writing practice, we encouraged women to keep a diary and write their thoughts, but that was not received very well as Türker said *“I don’t want others to learn what I think and feel.”*

Encouraging persistence

When women are overwhelmed by their responsibilities, an educational program can be a burden and reduce attendance. Every class had 3–6 learners and this was a blessing because teachers got to know the learners very well and could keep in touch with them, encouraging them to persist. They had to find creative ways to encourage attendance. Once Serap’s husband came home and thought that Serap was neglecting the children while taking this education program on WA. He forbade her to attend any longer. The teacher talked with the husband’s sister and through her encouraged him to change his mind. Serap started attending. The second teacher working with women in the textile industry encouraged them to stretch and relax before starting the education program as these women were physically quite tired.

Content and approach

The content had to be relevant to people’s lives and interesting. Effective communication was always a favorite as they could see its impact very quickly. Math topics were received very well. They enjoyed refreshing what they had

learned in school (e.g., division) and the ability to use those skills in shopping and budget management. They also preferred interactive question-answer discussions rather than a teacher lecturing. Empowerment topics were couched within literacy and numeracy instruction and that worked better because developing basic skills was also important for the women.

Collecting evaluation data

Remote instruction makes it almost impossible to collect evaluation data. To solve this problem, we delivered the printed pre- and post-test to the learners in the package with their textbook, but did not give any indication that they were tests. Normally the tests are longer, but we had to shorten them considerably because they were asked over the phone. Experienced teachers (but not the learners' own teacher) called the learners and asked them to do the test over the phone without any help from family members. This limited how many items could be on the tests.

To summarize, although short and delivered under challenging technological conditions to women who are already under a heavy burden, the program was still very effective and well-received. To end with a quote from Sevcan: "I took some education programs in the workplace, so I thought this would be the same and was hesitant about joining. This was very different. I learned so much."

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 for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

AD and MC developed the curriculum. BY oversaw the revisions and implementation. UK and KŞ provided data collection and analysis support. AD conducted the observations. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank our volunteer instructors, field monitors, and most of all the learners, for their interest and contributions.

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

- Durgunoğlu, A., and AÇEV (2018). *The multidimensional effects of AÇEV Hayat Dolu Buluşmalar Program on women* (Unpublished report). İstanbul: AÇEV.
- Durgunoğlu, A. Y., Öney, B., and Kuşcul, H. (2003). Development and evaluation of an adult literacy program in Turkey. *Int. J. Educ. Dev.* 23, 17–36. doi: 10.1016/S0738-0593(01)00050-5
- Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Dev. Change* 30, 435–464. doi: 10.1111/1467-7660.00125
- Kavanaugh, A., Puckett, A., and Tatar, D. (2013). Scaffolding technology for low literacy groups: From mobile phone to desktop PC? *Int. J. Hum. Comput. Interact.* 29, 274–288. doi: 10.1080/10447318.2013.765766
- LeVine, R. A., LeVine, S., Schnell-Anzola, B., Rowe, M. L., and Dexter, E. (2011). *Literacy and mothering: How women's schooling changes the lives of the world's children*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195309829.001.0001
- Öztan, E., Şahinkaya, A., Durgunoğlu, A. Y., Aşık, D., Kaya, M. U., and Cantürk, M. (2021). "Women's empowerment through lifelong learning: The case of empowerment and literacy programme for young women in Turkey," in *Gender and education in politics, policy and practice*, eds M. Carlson, B. E. Halldórsdóttir, B. Baranović, A. S. Holm, S. Lappalainen, and A. Spehar (Cham: Springer), 91–108. doi: 10.1007/978-3-030-80902-7_6
- Vanek, J., and Harris, K. (2020). *Digital literacy and technology integration in adult basic skills education: A review of the research*. ProLiteracy. Available online at: https://www.proliteracy.org/Portals/0/pdf/Research/Briefs/ProLiteracy-Research-Brief-02_Technology-2020-06.pdf (accessed June 22, 2022).
- Weitzman, A. (2018). Does increasing women's education reduce their risk of intimate partner violence? Evidence from an education policy reform. *Criminology* 56, 574–607. doi: 10.1111/1745-9125.12181



OPEN ACCESS

EDITED BY

Jesús de la Fuente,
University of Navarra, Spain

REVIEWED BY

Guido Marco Cicchini,
National Research Council (CNR), Italy
Paloma González-Castro,
University of Oviedo, Spain

*CORRESPONDENCE

Cándida Delgado
candida.delgado@uca.es

†These authors have contributed
equally to this work and share first
authorship

SPECIALTY SECTION

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

RECEIVED 06 April 2022

ACCEPTED 26 July 2022

PUBLISHED 06 September 2022

CITATION

Mera C, Delgado C, Aragón E,
Menacho I, Canto MC and Navarro JI
(2022) Contributions of the
psychology of mathematical cognition
in early childhood education using
apps. *Front. Psychol.* 13:913970.
doi: 10.3389/fpsyg.2022.913970

COPYRIGHT

© 2022 Mera, Delgado, Aragón,
Menacho, Canto and Navarro. 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.

Contributions of the psychology of mathematical cognition in early childhood education using apps

Carlos Mera[†], Cándida Delgado^{*†}, Estíbaliz Aragón[†],
Inmaculada Menacho[†], María Del Carmen Canto[†] and
José I. Navarro[†]

Department of Psychology, University of Cádiz, Cádiz, Spain

Educational interventions are necessary to develop mathematical competence at early ages and prevent widespread mathematics learning failure in the education system as indicated by the results of European reports. Numerous studies agree that domain-specific predictors related to mathematics are symbolic and non-symbolic magnitude comparison, as well as, number line estimation. The goal of this study was to design 4 digital learning app games to train specific cognitive bases of mathematical learning in order to create resources and promote the use of these technologies in the educational community and to promote effective scientific transfer and increase the research visibility. This study involved 193 preschoolers aged 57–79 months. A quasi-experimental design was carried out with 3 groups created after scores were obtained in a standardised mathematical competence assessment test, i.e., low-performance group ($N = 49$), high-performance group ($N = 21$), and control group ($N = 123$). The results show that training with the 4 digital learning app games focusing on magnitude, subitizing, number facts, and estimation tasks improved the numerical skills of the experimental groups, compared to the control group. The implications of the study were, on the one hand, provided verified technological tools for teaching early mathematical competence. On the other hand, this study supports other studies on the importance of cognitive precursors in mathematics performance.

KEYWORDS

digital learning games, app, mathematical cognition, magnitudes comparison, numerical estimation, early education

Introduction

Currently, international standardised tests of mathematics achievement do not show encouraging data for Spanish students ([Ministerio de Educación y Formación Profesional, 2019](#); [OECD, 2019](#)). Student performance is below average compared to other OECD countries. Therefore, the widespread failure in mathematics learning in our education system has aroused the interest of the scientific community ([European Commission, 2020](#)).

The onset of formal learning in the logical-mathematical area occurs at ~5 years of age. According to some authors, this is the age when the first signs of risk learning difficulties appear (Wong and Chan, 2019; Zhang et al., 2020). Numeracy skills in early childhood are the building blocks for the later successful development of mathematical competence (Passolunghi and Lanfranchi, 2012; Hornung et al., 2014). In this sense, it is necessary to consider the predictive role of cognitive processes linked to mathematical achievement at an early age (Zhang et al., 2020). These processes may also represent a critical feature in the detection of students at risk of mathematics learning difficulties (MLD) in later grades (Barnes and Marks, 2020).

The basic numerical skills that support the development of mathematical skills between 5 and 8 years of age focus on the following four main factors (Aunio and Räsänen, 2016): (1) symbolic and non-symbolic number magnitude; (2) understanding of mathematical relationships (logical-relational principle, arithmetic principles, symbols of arithmetic operations, and place value system and base-ten); (3) counting skills (knowledge of number symbols, number word sequence, enumeration with concrete objects); and (4) basic skills in arithmetic (arithmetic combinations, addition, and subtraction skills with number symbols).

In mathematics learning, there are a number of variables that predict performance, and these are usually grouped into two categories, namely, general domain and specific domain (Passolunghi et al., 2015; Ramani et al., 2017). General-domain predictors refer to higher-order cognitive variables such as working memory and processing speed (Fritz et al., 2019), whereas specific-domain predictors contribute to the performance of a particular school skill. There are several domain-specific predictors related to mathematical competence (Aragón et al., 2016), including *magnitude comparison* (Matejko and Ansari, 2016; Xenidou-Dervou et al., 2017) and *numerical estimation* (Reynvoet et al., 2016; Friso-van den Bos et al., 2018).

Magnitude comparison is defined as the sensitivity to distinguish numerical quantities. This is an important skill during the early stages of education (De Smedt et al., 2013; Toll et al., 2015). Numerical quantities can be represented symbolically (Arabic digits) or as non-symbolic quantities (a set of dots). Either way, the discrimination of non-symbolic quantities is a predictor of early numeracy skills (Soto-Calvo et al., 2015; Cueli et al., 2019b). In symbolic representation, it is necessary to have the ability to make a correct and immediate identification of each of the numerical symbols represented. They must then contrast the quantities and decide if the number is larger or smaller (Merkley and Ansari, 2016).

Many authors place the representation of non-symbolic magnitude as the predecessor of symbolic magnitude. Some studies show that 5-year-olds perform better on non-symbolic comparisons than on symbolic comparisons (Matejko and Ansari, 2016; Canto-López et al., 2019). In contrast, this difference in performance seems to disappear at age 6. From

this age onwards, children tend to achieve similar results in symbolic and non-symbolic quantities. However, as they learn more about the symbolic representation system, the difference with respect to non-symbolic representation narrows (Li et al., 2018). Nevertheless, there are still many scientific questions in this regard, as the data are inconclusive.

Number line estimation is another domain-specific cognitive skill that is closely related to the approximate number system (ANS) (De Hevia, 2016; Reynvoet et al., 2016; Zhu et al., 2017). In a typical number line estimation task, children are asked to indicate the position of an Arabic numeral on an empty number line. The line indicates the number 0 on the far left and a larger number (usually 10, 100, or 1,000) on the far right. In the first test, a concrete number is given for the child to estimate the position corresponding to each requested number on the line (number-position). However, the same estimation task can be performed by providing a line with the same characteristics as the previous one but marking a specific position on the line with another small line in a perpendicular position. In this way, the participant has to indicate the number that corresponds to that particular position (position-number) (Siegler and Opfer, 2003).

Number line estimation seems to be important for learning mathematics. In fact, the basis of numerical cognition is considered an innate representation of numerical magnitude as a mental number line (Dehaene, 2003, 2011). Recent studies have shown that this domain-specific cognitive task has a high predictive value for mathematical achievement (Schneider et al., 2017; Cerda et al., 2018; Núñez-Peña et al., 2019).

Although results are not yet conclusive, training the ability to estimate magnitude on the number line may be useful in gaining adequate access to symbolic numbers and their relationship to magnitude. This suggests that the number line can be a powerful representational tool for strengthening connexions between symbols and the quantities they represent (Booth and Siegler, 2008). These cognitive processes underlying children's responses in the estimation task correlate with mathematics achievement even at later educational stages (Geary et al., 2013; Schneider et al., 2018).

Therefore, training these variables considered predictors with digital technology can bring academic benefits (Re et al., 2020). This study contributes to the research on mathematics app use with children aged 4 and 5 years and specifically touch-screen apps that contain digital learning games.

Digital learning games focus on the design of apps or videogames that incorporate learning models and educational content to enhance learning (Prensky, 2003). Digital technologies, such as a *tablet* or *computer applications*, support teachers and offer certain advantages for student learning. In this sense, it is necessary that the software is well-designed and its content based on the child's stage of development (Kucirkova et al., 2014; Hubber et al., 2016).

Several authors consider that technologies, compared to traditional resources of similar structure, can contribute to improving learning (Wouters et al., 2013; Fernández-Abella et al., 2019; Peralbo-Uzquiano et al., 2020). In other words, the tasks, and not the format in which they are presented, would be responsible for the progress.

Numerous benefits have been reported with the app for children aged 4–7 years in need of additional support with learning mathematics. In addition, apps have been shown to help children with poorer short-term memory make greater learning gains than those with higher memory skills (Outhwaite et al., 2017). For this reason, increased time learning mathematics with an educational storey app at home improved children's mathematical skills in primary school (Outhwaite et al., 2019). However, there are some disadvantages to consider. Abuse of apps can lead to the emergence of an addiction to this type of device, as well as promoting individual work and, consequently, social isolation (Bonilla-Barbosa, 2014). Furthermore, excessive use of technology can lead to a decrease in effort in some basic school skills for learning, such as writing (Graham et al., 2013).

There are educational technology games that fit the criteria of the early years' mathematics curriculum (Schacter et al., 2016; Sheehan et al., 2019; Schenke et al., 2020). However, few focus on the cognitive approach to mathematical learning (Aragón et al., 2017; Mera et al., 2019; Peralbo-Uzquiano et al., 2020).

Inside these multiple sceneries, in the last decade, educational psychology has proposed different lines of research related to the cognitive processes linked to mathematical learning at a young age, as well as the possibility of taking advantage of the knowledge available in this relationship to implement programmes with different characteristics. It is in this context that this study is located. In this research, we designed and implemented 4 digital learning app games to train specific cognitive predictors that should influence early mathematical competence in 5-year-old children. There are two objectives as follows: on the one hand, to demonstrate that the teaching of specific cognitive predictors through app games improves mathematical achievement; on the other hand, to empirically verify that the digital educational games designed to provide support in the teaching-learning process of mathematical competence, both in the classroom and at home.

Materials and methods

Participants

The total number of participants was 193 preschoolers from middle socio-economic and educational levels of families, whose ages ranged from 57 to 79 months ($M = 63.3$, $sd = 3.7$). Of them, 107 (55%) were girls, aged between 57 and 70 months ($M = 63.1$, $sd = 3.4$) and 86 (45%) were boys, aged between 57 and 79 months ($M = 63.4$, $sd = 4$). From the total sample of

students from 4 schools (two public schools and two subsidised schools), children with special educational needs were excluded, as judged by experts. Percentile scores obtained in TEMA-3 in the pre-intervention phase allowed the construction of the following three groups: high performance in mathematics (percentile higher than 80), low performance (percentile lower than 25), and average performance. Groups were distributed in the following way:

- Low-performance group (*LP group*) consisted of 31 girls (63.3%) and 18 boys (36.7%) aged 57–79 months ($M = 62.55$, $sd = 4$).
- High-performance group (*HP group*) consisted of 9 girls (42.9%) and 12 boys (57.1%) aged 58–70 months ($M = 64.6$, $sd = 3.5$).
- Average-performance group (*control group*) consisted of 67 girls (54.5%) and 56 boys (45.5%) aged 57–71 months ($M = 63.26$, $sd = 3.5$).

Assessment instruments

Test of early mathematics ability – third edition

This test is composed of two subtests that focus on the assessment of informal and formal thinking, both in concepts and skills. The informal subtest is composed of tasks aimed at assessing numeracy, quantity comparison, informal calculation, and basic informal concepts. The formal subtest assesses conventions related to the reading and writing of quantities, knowledge of numerical facts, formal calculation, and formal mathematical concepts. The test was administered individually and lasted around 30 min. This time varied according to the age of the pupils. This individually administered test for children between 3 and 8 years of age identifies students with mathematics learning difficulties or at-risk students. It consists of a total of 72 items, presented in order of increasing difficulty, which is administered until the student responds incorrectly to 4 items in a row. Cronbach's alpha was 0.91 (Ginsburg et al., 2007).

Symbolic and non-symbolic comparison test

To assess magnitude processing skills, participants were presented with a booklet of symbolic (Arabic digits) and non-symbolic (dots) number pairs and were asked to compare two numerical magnitudes and point (using a pencil) to the larger number within a given time (2 min per modality). The magnitudes vary from 1 to 9 and the side on which the larger magnitude is presented is counterbalanced in all items. For each format, 56 items were presented (56 symbolic pairs and 56 non-symbolic pairs). Cronbach's alpha was 0.86 (Nosworthy et al., 2013).

Numerical estimation task

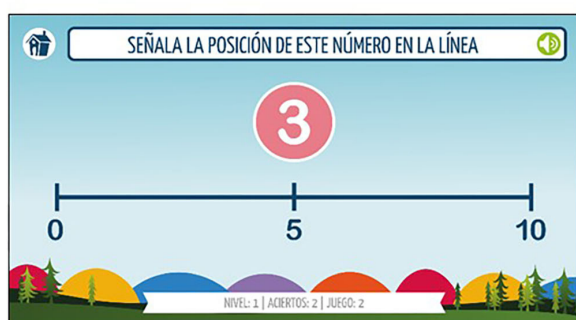
This pencil and paper test assesses number line estimation in its two modalities: *number-position* (a number is presented and the participant must know its position within a straight line), and *position-number*, where a sign is shown on the straight line and the participant must recognise which number it corresponds to. The test consists of 10 items for each modality, corresponding to the following numbers: 2, 4, 7, 8, 11, 13, 16, 17, 18, and 19, randomly presented. The comparison of means was made on the basis of the number of correct answers, with respect to the number requested vs. the number given by the student. For this purpose, the answer was assigned as correct if it did not have an error rate higher than $\pm 15\%$ of the number asked for. Cronbach's alpha was 0.80 (Siegler and Booth, 2004).

Intervention instruments

The intervention programme used consisted of 4 apps. These apps for use on touchscreen devices focus on stimulating and training the cognitive foundations associated with early mathematics learning through a simple game in which the participant only has to press a finger on the screen to answer (Mera et al., 2019) (Figure 1).

Tablet devices are mobile, lightweight, and do not rely on the motor skills needed to use other technologies, such as a computer keyboard and mouse (Kucirkova et al., 2014).

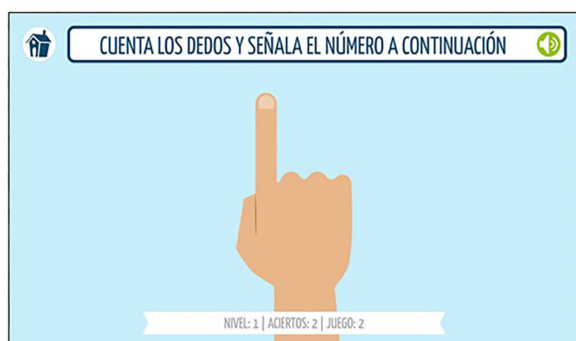
- *Compare amounts with Mon the dragon.* The game consists of discriminating quantities symbolically and non-symbolically represented, depending on the level of difficulty.
- *Quick counting with Mon the dragon.* The aim of the game is to develop the ability to discriminate small quantities; to count suddenly (without the need to point to each element); and to identify the position of the number within a number line.
- *Calculation with Mon the dragon.* The game aims to develop the ability to perform simple calculations that are stored in long-term memory (addition, subtraction, multiplication, or division), useful in the understanding and development of arithmetic concepts and facilitating problem-solving.
- *Find the hidden number with Mon the dragon.* Two estimation modes coexist in the app: (1) a number is shown and the child must place it in the appropriate position on a straight line (number-position mode); and (2), a mark is shown on the straight line and the player must determine



Find the hidden number with Mon the Dragon



Compare amounts with Mon the Dragon



Quick counting with Mon the Dragon



Calculation with Mon the Dragón

FIGURE 1
Illustrations of the App designed for the study.

<i>Task</i>	Pre-		Intervention				Post-	
	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May.	Jun.
Pre-intervention assessment and allocation of experimental groups : • It was attended by 70 participants ($LP=49$; $HP=21$) distributed in 7 training groups.								
35 training sessions of about 25 minutes each during school hours. The control group remained in class.								
Assessment of all participants.								

FIGURE 2
Procedural planning.

approximately which number would occupy that place (position-number mode).

Procedure

After the assignment of the intervention and control groups, 7 working-session groups were generated with 10 participants in each ($N = 70$). The empirical study was conducted using pre- and post-intervention measures. The sequence of the research had three phases, i.e., pre-, intervention and post-intervention. In the intervention phase, each working group carried out 35 sessions of ~25 min. The sessions were carried out in a separate classroom, in good working conditions, and during school hours, respecting break times or non-formal educational activities. During the training of the experimental groups, the control group remained in academic activities. In each working session, there were one or two evaluators with training and experience in dealing with preschoolers and using the app (Figure 2).

Each student was provided with a tablet to develop the intervention sessions. After the intervention period, in the post-intervention evaluation, the evaluators were maintained in order to try to minimise any extraneous variables that might appear due to the effect of the evaluator, leaving a minimum time margin of 6 months between the two evaluation periods, so that recall would not influence the results of the tests.

Data collection during the intervention phase

The apps run through the Internet and allowed immediately collect the user's general data, date, and task completion timing.

Likewise, the design of the app made it possible to generate a file with the user's answers: correct responses, errors, levels of difficulty reached, reaction, and response time for each session. All data were stored in a safe-encrypted database for subsequent statistical calculations.

Statistical analysis

To analyse differences between the three participating groups in early mathematical competence, a descriptive and inferential analysis of the results was carried out. For the descriptive analysis, measures of central distribution and dispersion were calculated, as well as the gains produced in each of the groups. For the inferential analysis, the *Kruskal-Wallis* test was used due to the size of the sample in the *HP* group ($N = 21$). This test is a non-parametric method with the groups replaced by categories. The *Wilcoxon* test was then used to analyse the differences found between each of the groups, which is based on the differences in the absolute value of the records obtained in the two evaluation phases. Finally, the effect size was calculated by applying Cohen's *d* test, which provides us with a practical significance of the results.

Results

Influence of app on mathematical competence

The descriptive statistics obtained in the pre-intervention and post-intervention evaluation phases were analysed, and the gains in the TEMA-3 test scores of each group were calculated (Table 1).

The *control* group recorded a baseline minimum score of 9 and a maximum score of 32 ($M = 21.52$; $sd = 3.74$) on the

TABLE 1 Means, typical deviations and gains from the TEMA-3 test between the pre- and post intervention phases.

TEMA-3	Pre-intervention M (sd)	Post-intervention M (sd)	Gains M
Control group	21.52 (3.74)	28.26 (5.31)	6.74
LP group	16.53 (2.61)	28.16 (4.99)	11.63
HP group	30.48 (5.52)	42.10 (5.01)	11.62

TABLE 2 Results of *post-hoc* tests in pairs on gains in the mathematical competence test.

	U of Mann-Whitney	Std. error	Desv. statistical test	Sig.
Control- HP	57.179	13.157	4.346	0.000
Control-LP	60.226	9.413	6.398	0.000
HP -LP	3.048	14.534	0.210	0.834

test (TEMA-3). The LP group obtained pre-intervention scores with a minimum of 7 and a maximum of 20 ($M = 16.53$; $sd = 2.61$). The HP group scored a minimum of 24 and a maximum of 39 ($M = 30.48$; $sd = 5.52$). After the intervention was carried out in the experimental groups, the descriptive statistics of the dependent variable during the post-intervention phase showed the following results: in the *control group* ($M = 28.26$ and $sd = 5.31$), in the *LP group* ($M = 28.16$ and $sd = 4.99$), and the *HP group* ($M = 42.10$ and $sd = 5.01$), respectively. It can be seen that the gains were higher for the groups that received instruction by training with the app.

To test the statistical significance of these gains, the *Kruskal-Wallis* test was used, being an extension of the *Mann-Whitney U*-test for 3 or more groups. The result confirmed that there was a significant difference at 95% ($H = 50.588$, $p < 0.05$), with the median between the groups considered to be different. Accordingly, *post-hoc* contrasts were performed using the *Mann-Whitney U*-test for the groups (Table 2).

The results of the pairwise comparisons showed, on the one hand, a significant contrast between the *control* and *HP* groups ($U_{(C-HP)} = 57.179$, $p < 0.001$) and between the *control* and *LP* groups ($U_{(C-LP)} = 60.226$, $p < 0.001$). On the other hand, no significant differences were found between the gains produced by the *LP* and *HP* groups. In view of the results obtained, the gains were significant at 95% between the *control group* and both experimental groups, with non-significant differences and very similar gains in the contrast between the experimental groups.

To analyse the differences found in each of the groups before and after the intervention, the Wilcoxon test showed significant differences between the evaluations of all the intervening groups as follows: the *control group* ($Z = -9.477$, $p < 0.05$), *LP group* (Z

TABLE 3 Effect size on the mathematical competence test (pre/post-intervention).

	Cohen's d test	r
Control group	1.47	0.59
LP Group	2.92	0.82
HP Group	2.20	0.74

TABLE 4 Descriptive statistics of domain-specific predictors (symbolic and non-symbolic comparison assessed in both assessment phases by groups).

		Control M (sd)	LP M (sd)	HP M (sd)
Symbolic comparison	Pre.	31.41 (9.3)	27.12 (8.9)	35.24 (9)
	Post.	41.00 (7.3)	36.96 (8.9)	44.67 (8.4)
Non-symbolic comparison	Pre.	34.19 (6)	32.49 (6.2)	34.38 (7.8)
	Post.	39.97 (6.1)	37.53 (8)	40.52 (5.8)
Estimate	Pre.	5.27 (1.9)	5.20 (1.8)	5.86 (2.1)
	Post.	5.15 (2.1)	7.90 (1.7)	9.1 (0.99)

LP, Low Performance; HP, High Performance.

$= -6.099$, $p < 0.05$), and *HP group* ($Z = -4.019$, $p < 0.05$). To check this increase, an effect size calculation was performed for the different groups (Table 3).

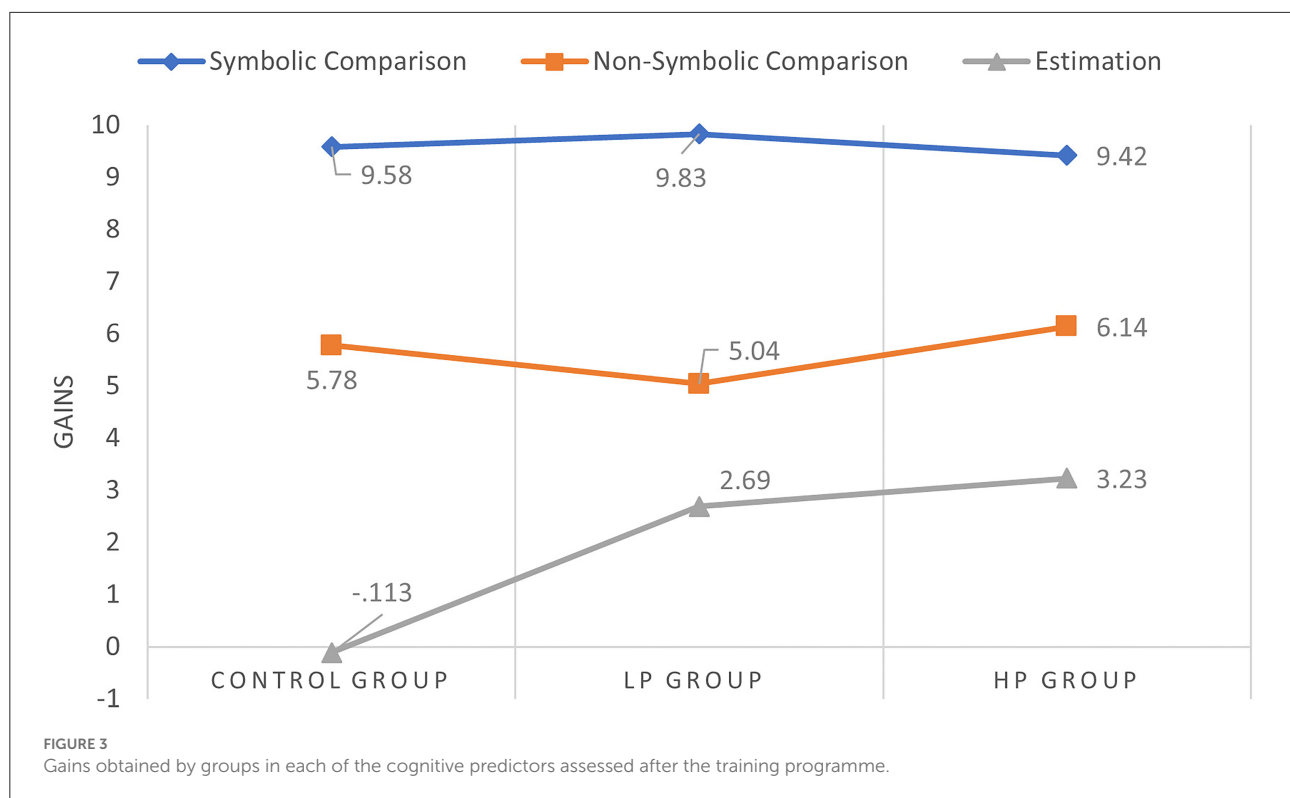
The result of the Cohen's *d* test showed that the effect size produced between the pre-intervention and post-intervention assessments in mathematical competence maintained high values in each of the assumptions, being higher in the experimental groups.

Influence of app on domain-specific cognitive predictors

The calculation was to check the influence of using apps over domain-specific cognitive predictors. For the hypothesis testing, descriptive statistics were calculated for the cognitive tests for the assessment of domain-specific cognitive predictors (Table 4) and the increase between groups (Figure 3).

Respecting the symbolic and non-symbolic comparison, all groups increased in the post-intervention evaluation phase compared to the pre-intervention assessment. In the numerical line estimation task, the *control group* mean in pre-intervention was 5.27 ($sd = 1.9$), while in post-intervention, the result was slightly lower with a mean of 5.15 ($sd = 2.01$). However, in both the *LP* and *HP groups*, scores were higher, increasing scores in both groups.

In the numerical line estimation task, the *control group* mean in pre-intervention was 5.27 ($sd = 1.9$) while in post-intervention the result was slightly lower with a mean of 5.15

TABLE 5 Result of *post-hoc* by couples testing in the estimation task.

	Mann-Whitney U	Std. error	Std. desv.	Sig.	Adjust sig.
Control-HP	59.505	13.120	4.535	0.000	0.000
Control-LP	50.199	9.387	5.348	0.000	0.000
LP-HP	-9.306	14.493	-0.642	0.521	1.00

LP, Low Performance; HP, High Performance.

($sd = 2.01$). However, in both the LP and HP groups, scores were higher, increasing scores in both groups.

In order to test the gaining means differences between the three groups, the *Kruskal-Wallis* test confirmed significant differences in the estimation task at 95% ($H = 40.983$, $p < 0.05$). In the symbolic and non-symbolic comparison tasks, the analysis of variance calculation indicated that there was no statistically significant difference.

Contrast tests confirmed significant gaining differences between the control and LP groups and between the control and HP groups. However, no statistically significant differences were found between the LP and HP groups. In contrast, the largest gains occurred between the control and HP groups (Table 5).

Considering the variances of each group between the pre- and post-assessment, the *Wilcoxon* test showed significant differences in the estimation task in the pre-intervention and

post-intervention evaluations were presented in the LP ($Z = -5.027$, $p < 0.05$) and HP ($Z = -3.670$, $p < 0.05$) groups. Regarding the groups that received instruction using the app, the control group ($Z = -0.035$, $p > 0.05$) did not show significant differences ($Z = -0.035$, $p > 0.05$) between the two assessment phases in the estimation task. In the symbolic comparison task, significant differences were observed for the different groups as follows: control ($Z = -8.268$, $p < 0.05$), LP ($Z = -5.385$, $p < 0.05$), and HP ($Z = -3.304$, $p < 0.05$). In relation to the non-symbolic comparison task, the range comparison between pre-intervention and post-intervention showed a significant difference in each of the groups as follows: control ($Z = -7.020$, $p < 0.05$), LP ($Z = 3.936$, $p < 0.05$), and HP ($Z = 3.287$, $p < 0.05$).

The effect size was high in all groups in both the symbolic and non-symbolic comparison tasks. As Table 6 shows, statistically significant differences were only observed in the estimation tasks in both experimental groups. In this sense, a larger effect size was found in the HP group followed by the LP group as a result of Cohen's d test. However, the control group showed no changes between the two evaluation phases.

Discussion

Currently, part of the scientific community is focusing on the use of technologies and their influence on the development

TABLE 6 Results of Cohen's *d* test to calculate the effect size between pre- and post-intervention for the symbolic and non-symbolic comparison tasks.

PRE/POS-intervention		Control Group	LP Group	HP Group
Comparación Simbólica	Cohen's <i>d</i> test	1.147	1.105	1.083
	R	0.497	0.483	0.476
Comparación No-Simbólica	Cohen's <i>d</i> test	0.955	0.704	0.893
	R	0.431	0.332	0.407
Estimación	Cohen's <i>d</i> test	−0.05	1.54	1.97
	R	−0.02	0.61	0.7

LP, Low Performance; HP, High Performance.

of young learners (Hatzigianni and Kalaitzidis, 2018). Easy-to-access devices are generally used as entertainment tools. Paediatricians, psychologists, and educators advise careful monitoring and limiting the use of digital devices among young children (Kabali et al., 2015).

The scientific literature focused on addressing the development of early mathematical cognition sustains the meaning of training different skills as a cognitive basis for learning math (Geary et al., 2013; Aragón et al., 2014, 2020; Soto-Calvo et al., 2015; Malone et al., 2019; Pace et al., 2019). These skills also provide the development of a solid background in early numeracy, which is considered critically important for later mathematical achievement (Ramani et al., 2018).

The mental number line is fundamental for learning mathematics. In fact, the basis of numerical cognition is considered an innate representation of numerical magnitude in the form of a numerical mental line (Dehaene, 1997, 2003). Recent studies consider that the specific-domain cognitive task of mathematics of estimating a numerical value and its positioning on a straight line has a high predictive value for mathematical achievement (Zhu et al., 2017; Cerda et al., 2018; Schneider et al., 2018; Núñez-Peña et al., 2019).

As a result of this study, the use of the app provided a statistically significant improvement in numeracy skills in both groups compared to the control group. Data suggest that these technological tools can be used for teaching math and helping early childhood educators to provide new experiences for their students (Mattoon et al., 2015).

Apps have also verified their effectiveness in mathematics learning, offering individualised instruction and using technological tools to promote this improvement (Schacter and Jo, 2017; Miller, 2018; Outhwaite et al., 2019; Schenke et al., 2020). However, some studies have found no significant intervention effects with the use of programmes that can be purchased from the current digital platforms (Hellstrand et al., 2020). The critical issue is that it is necessary to experimentally

prove its effectiveness in educational settings. Such an evaluation has been possible in this study. App training contributed to improving the mathematical performance of 5-year-old students. The app had educational usefulness.

In relation to the influence of the app on the mathematical learning cognitive background, the number line estimation task showed a higher gain in groups that had received the instruction program.

Positive results were achieved with both types of estimation tested, i.e., number-position and position-number. Estimation is considered critical for learning mathematics. Numerical cognition could be attributed to the representation of numerical magnitude in the form of a mental line and spatial-numerical associations that are already active in early childhood (McCrink and De Hevia, 2018).

Second, both low and high achievers experienced a statistically significant increase in performance in estimation. These results suggest that the instructional system is adaptable and usable by any new mathematics learner, regardless of starting level. These data are more relevant considering that other studies have suggested the causal role of these skills in mathematics (Obersteiner et al., 2013; Moeller et al., 2015; Rugani et al., 2015).

Third, the results contrast with the control group performance, whose average scores in the post-assessment decreased. This suggests that number line estimation tasks could be included in the educational itinerary of early education students. These data are consistent with several contributions. In fact, a meta-analysis (Schneider et al., 2018) with a total of 10,576 participants (aged 4–14 years) showed that the ability to estimate on the number line was a robust tool for diagnosing and predicting numeracy, increasing with age.

As for the magnitude comparison, the learning gainings in the different instructional groups were similar. No significant differences were observed between the groups that had received the instructional programme and the control group. These results could be explained by the accessibility of this type of task for students in the second cycle of early education. The official educational curriculum in Spain and other countries establishes the need to accurately estimate collections or quantities of continuous subjects. Despite not finding statistically significant differences in the use of the app, in terms of magnitude comparison, this task should be relevant to be taught through this instructional programme due to its potential importance in improving mathematics skills (Laski and Siegler, 2007; De Smedt et al., 2009; Matejko and Ansari, 2016; Xenidou-Dervou et al., 2017; Cueli et al., 2019a). Comparing quantities development becomes relevant for improving mathematics learning as this task is part of the ANS for non-symbolic comparison and the PNS for symbolic comparison. Both are closely related to estimation on the number line (Siegler and Booth, 2004; De Hevia, 2016; Reynvoet et al., 2016; Zhu et al., 2017).

One example of this is that symbolic comparison and, to a lesser extent, number line estimation skills of early childhood students are predictive of numeracy and longitudinal predictors of overall mathematics performance at these ages (Toll et al., 2015; Lourenco and Bonny, 2017; Mera et al., 2017; Hawes et al., 2019).

Symbolic and non-symbolic magnitude processing skills show different developmental trajectories, with symbolic skills showing greater gains than non-symbolic skills during the early educational stages (Matejko and Ansari, 2016). Similar results were found in this study. A possible explanation for these results may also be related to the type of test used for the assessment of ANS, as results in favour of training are found in the estimation test (van 't Noordende et al., 2021), but not in the non-symbolic comparison test. This fact does not call into question the involvement of the ANS in the development of mathematical competence, but these results may be attributable to the measures used.

This is a research study sustaining the trajectory of the last decade of research in educational psychology. It emphasises how important it is to analyse cognitive variables linked to mathematical learning from the first schooling years. The research focuses on the use of technological contributions to the field of education as contrasted tools. These tools can be used rationally in both formal and informal mathematics education. The study has the limitations of quasi-experimental studies. In choosing the experimental sample, we do not have truly random matched groups (which are always difficult to select in educational and clinical settings). For future research, a similar design could be carried out but within an experimental study analysing the influence of the independent variables without the bias of learning difficulty or optimal performance, as has been the case in this study. Considering another concern about the study, we refer to the potential influence that the general predictors could have on the specific ones. That is, there are no purely specific predictors since to carry out any specific mathematical skill, general skills are also required (perceptive, executive, attentional, and processing speed).

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 reviewed and approved by the Ethics committee of the Research in Cádiz. The legal protocols of

ethical guarantees for the development of the study were followed at all times, both in relation to authorizations and data protection. This work followed the International Code of Ethics in the Humanities and Social Sciences of the Centre for Research Ethics and Bioethics. Written informed consent was obtained from the parents or guardians of the participants, as well as from the schools.

Author contributions

CM: application of assessment tools and intervention programme in educational settings, statistical analysis of results, and theoretical discussion of early mathematical competencies. CD: literature review, application of assessment tools in educational settings, and writing the final version of the manuscript. EA: students assessment, statistical analysis, theoretical discussion of mathematical cognition, and contextualisation of statistical results according to the scientific literature. IM: references updating, translation, and reviewing the different manuscript versions. MC: application of assessment tools and intervention programme in educational settings. JN: students assessment, statistical analysis, theoretical discussion of psychometric instruments applied, analysis of comparable results in literature and review of the different manuscript versions, and in charge of the ethical consent of participants. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the MINECO/FEDER Spanish Government grants (PID2019-105584GB-I00 MINECO/FEDER and PID2020-119561RB-I00).

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

- Aragón, E., Aguilar, M., and Navarro, J. I. (2017). Improving number sense in kindergarten children with low achievement in mathematics. *Ann. Psychol.* 33, 311–318. doi: 10.6018/analesps.33.2.239391
- Aragón, E., Cerda, G., Aguilar, M., Mera, C., and Navarro, J. I. (2020). Modulation of general and specific cognitive precursors to early mathematical competencies in preschool children. *Eur. J. Psychol. Educ.* 36, 405. doi: 10.1007/s10212-020-00483-4
- Aragón, E., and Navarro, J. I., and Aguilar, M. (2016). Domain-specific predictors for fluency calculation at the beginning of primary school education. *Electron. J. Res. Educ. Psychol.* 14, 482–499. doi: 10.14204/ejrep.40.15107
- Aragón, E., Navarro, J. I., Aguilar, M., and Cerda, G. (2014). Cognitive predictors of 5-year-old students' early number sense. *J. Psychodidactics* 20, 83–97. doi: 10.1387/RevPsicodidact.11088
- Aunio, P., and Räsänen, P. (2016). Core numerical skills for learning mathematics in children aged five to eight years – a working model for educators. *Eur. Early Child. Educ. Res. J.* 24, 684–704. doi: 10.1080/1350293X.2014.996424
- Barnes, A., and Marks, R. (2020). “Formative assessment in mathematics education: Key debates,” in *Debates in Mathematics Education* (London: Routledge), 243–255.
- Bonilla-Barbosa, J. (2014). Ventajas y desventajas de las TIC en el aula. *#ashtag*. 4, 124–131. Available online at: <http://investigaciones.cun.edu.co/investigaciones/ojs/index.php/hashtag/article/view/46>
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Canto-López, M. C., Aguilar, M., García-Sedeño, M. A., Navarro, J. I., Aragón, E., Delgado, C., et al. (2019). Numerical estimation and mathematical learning methodology in preschoolers. *Psychol. Rep.* 124, 438–458. doi: 10.1177/0033294119892880
- Cerda, G., Aragón, E., Pérez, C., Navarro, J. I., and Aguilar, M. (2018). The open algorithm based on numbers (ABN) method: an effective instructional Approach to domain-specific precursors of arithmetic development. *Front. Psychol.* 9:1811. doi: 10.3389/fpsyg.2018.01811
- Cueli, M., Areces, D., García, T., Rodríguez, C., Vallejo, G., and González-Castro, P. (2019a). Influence of initial mathematical competencies on the effectiveness of an intervention. *Br. J. Educ. Psychol.* 89, 288–306. doi: 10.1111/bjep.12239
- Cueli, M., Areces, D., McCaskey, U., Álvarez-García, D., and González-Castro, P. (2019b). Mathematics competence level: the contribution of non-symbolic and spatial magnitude comparison skills. *Front. Psychol.* 10:465. doi: 10.3389/fpsyg.2019.00465
- De Hevia, M. D. (2016). “Link between numbers and spatial extent from birth to adulthood,” in *Continuous Issues in Numerical Cognition. How Many or How Much*, ed A. Henik (London; San Diego, CA: Cambridge, Oxford: Elsevier), 37–58.
- De Smedt, B., Noël, M. P., Gilmore, C., and Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? a review of evidence from brain and behavior. *Trends Neurosci. Educ.* 2, 48–55. doi: 10.1016/j.tine.2013.06.001
- De Smedt, B., Verschaffel, L., and Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *J. Exp. Child Psychol.* 103, 469–479. doi: 10.1016/j.jecp.2009.01.010
- Dehaene, S. (1997). *The Number Sense : How the Mind Creates Mathematics*. Oxford University Press. Available online at: <https://search.proquest.com/openview/eb503d0651163a6743e1b49d523d978a/1?pq-origsite=gscholar&cbl=41299>
- Dehaene, S. (2003). The neural basis of the weber-fechner law: a logarithmic mental number line. *Trends Cogn. Sci.* 7, 145–147. doi: 10.1016/S1364-6613(03)00055-X
- Dehaene, S. (2011). *The Number Sense: How The Mind Creates Mathematics*. OUP USA.
- European Commission (2020). *Education and Training Monitor 2020 - Publications Office of the EU (europa.eu)*. Available online at: <https://op.europa.eu/en/publication-detail/-/publication/92c621ce-2494-11eb-9d7e-01aa75ed71a1/language-en> (accessed March 1, 2022).
- Fernández-Abella, R., Peralbo-Uzquiano, M., Durán-Bouza, M., Brenlla-Blanco, J. C., and García-Fernández, M. (2019). Virtual intervention programme to improve the working memory and basic mathematical skills in early childhood education. *Revista de Psicodidáctica* 24, 17–23. doi: 10.1016/j.psicod.2018.09.002
- Friso-van den Bos, I., Kroesbergen, E. H., and Van Luit, J. E. H. (2018). Counting and number line trainings in kindergarten: effects on arithmetic performance and number sense. *Front. Psychol.* 9, 975. doi: 10.3389/fpsyg.2018.00975
- Fritz, A., Haase, V., and Räsänen, P. (Eds.) (2019). *International Handbook of Mathematical Learning Difficulties. From the Laboratory to the Classroom*. Switzerland: Springer
- Geary, D. C., Hoard, M. K., Nugent, L., and Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLoS ONE* 8, e54651. doi: 10.1371/journal.pone.0054651
- Ginsburg, H., Baroody, A. J., del Río, M. C. N., and Guerra, I. L. (2007). *TEMA 3: Test de competencia matemática básica*. Madrid: Tea.
- Graham, S., Gillespie, A., and McKeown, D. (2013). Writing: importance, development, and instruction. *Read. Writ.* 26, 1–15. doi: 10.1007/s11145-012-9395-2
- Hatzigianni, M., and Kalaitzidis, I. (2018). Early childhood educators' attitudes and beliefs around the use of touchscreen technologies by children under three years of age. *Br. J. Educ. Technol.* 49, 883–895. doi: 10.1111/bjet.12649
- Hawes, Z., Nosworthy, N., Archibald, L., and Ansari, D. (2019). Kindergarten children's symbolic number comparison skills predict 1st grade mathematics achievement: evidence from a two-minute paper-and-pencil test. *Learn Instr.* 59, 21–33. doi: 10.1016/j.learninstruc.2018.09.004
- Hellstrand, H., Korhonen, J., Linnanmäki, K., and Aunio, P. (2020). The number race—computer-assisted intervention for mathematically low-performing first graders. *Eur. J. Spec. Needs Educ.* 35, 85–99. doi: 10.1080/13488678.2019.1615792
- Hornung, C., Schiltz, C., Brunner, M., and Martin, R. (2014). Predicting first-grade mathematics achievement: the contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Front. Psychol.* 5, 272. doi: 10.3389/fpsyg.2014.00272
- Hubber, P. J., Outhwaite, L. A., Chigeda, A., McGrath, S., Hodgen, J., and Pitchford, N. J. (2016). Should touch screen tablets be used to improve educational outcomes in primary school children in developing countries? *Front. Psychol.* 7:839. doi: 10.3389/fpsyg.2016.00839
- Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., et al. (2015). Exposure and use of mobile media devices by young children. *Pediatrics* 136, 1044–1050. doi: 10.1542/peds.2015-2151
- Kucirkova, N., Messer, D., Sheehy, K., and Panadero, C. F. (2014). Children's engagement with educational iPad App: insights from a Spanish classroom. *Comput. Educ.* 71, 175–184. doi: 10.1016/j.compedu.2013.10.003
- Laski, E. V., and Siegler, R. S. (2007). Is 27 a big number? correlational and causal connections among numerical categorization number line estimation and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Li, Y., Zhang, M., Chen, Y., Deng, Z., Zhu, X., and Yan, S. (2018). Children's non-symbolic and symbolic numerical representations and their associations with mathematical ability. *Front. Psychol.* 9, 1035. doi: 10.3389/fpsyg.2018.01035
- Lourenco, S. F., and Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Dev. Sci.* 20, e12418. doi: 10.1111/desc.12418
- Malone, S. A., Heron-Delaney, M., Burgoyne, K., and Hulme, C. (2019). Learning correspondences between magnitudes, symbols and words: evidence for a triple code model of arithmetic development. *Cognition* 187, 1–9. doi: 10.1016/j.cognition.2018.11.016
- Matejko, A. A., and Ansari, D. (2016). Trajectories of symbolic and nonsymbolic magnitude processing in the first year of formal schooling. *PLoS ONE* 11, e0149863. doi: 10.1371/journal.pone.0149863
- Matton, C., Bates, A., Shifflet, R., Latham, N., and Ennis, S. (2015). Examining computational skills in prekindergartners: the effects of traditional and digital manipulatives in a prekindergarten classroom. *Early Child. Res. Pract.* 17, 1–9. Available online at: <https://files.eric.ed.gov/fulltext/EJ1072360.pdf>
- McCrink, K., and De Hevia, M. D. (2018). From innate spatial biases to enculturated spatial cognition: the case of spatial associations in number and other sequences. *Front. Psychol.* 9, 415. doi: 10.3389/fpsyg.2018.00415
- Mera, C., Ruiz, G., Aguilar, M., Aragón, E., Delgado, C., Menacho, I., et al. (2019). Coming together: R&D and children's entertainment company in designing apps for learning early math. *Front. Psychol.* 9:2751. doi: 10.3389/fpsyg.2018.02751

- Mera, C., Ruiz, G., Navarro, J. I., Aragón, E., Delgado, C., and Aguilar, M. (2017). APP designed for early math training. "Magnitudes comparison," in *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality – TEEM* (Cádiz).
- Merkley, R., and Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: Evidence from brain and behavior. *Curr. Opin. Behav. Sci.* 10, 14–20. doi: 10.1016/j.cobeha.2016.04.006
- Miller, T. (2018). Developing numeracy skills using interactive technology in a play-based learning environment. *Int. J. STEM Educ.* 5, 1–11. doi: 10.1186/s40594-018-0135-2
- Moeller, K., Fischer, U., Nuerk, H. C., and Cress, U. (2015). Computers in mathematics education - training the mental number line. *Comput. Human Behav.* 48, 597–607. doi: 10.1016/j.chb.2015.01.048
- Nosworthy, N., Bugden, S., Archibald, L., Evans, B., and Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS ONE* 8, e67918. doi: 10.1371/journal.pone.0067918
- Núñez-Peña, M. I., Colomé, A., and Aguilar-Lleyda, D. (2019). Number line estimation in highly math-anxious individuals. *Br. J. Psychol.* 110, 40–59. doi: 10.1111/bjop.12335
- Obersteiner, A., Reiss, K., and Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learn. Instr.* 23, 125–135. doi: 10.1016/j.learninstruc.2012.08.004
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. Paris: OECD. Available online at: https://read.oecd-ilibrary.org/education/pisa-2018-assessment-and-analytical-framework_b25efab8-en#page1
- Outhwaite, L. A., Faulder, M., Gulliford, A., and Pitchford, N. J. (2019). Raising early achievement in math with interactive app: a randomized control trial. *J. Educ. Psychol.* 111, 284–298. doi: 10.1037/edu0000286
- Outhwaite, L. A., Gulliford, A., and Pitchford, N. J. (2017). Closing the gap: Efficacy of a tablet intervention to support the development of early mathematical skills in UK primary school children. *Comput. Educ.* 108, 43–58. doi: 10.1016/j.compedu.2017.01.011
- Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., and Hirsh-Pasek, K. (2019). Measuring success: within and cross-domain predictors of academic and social trajectories in elementary school. *Early Child. Res. Q.* 46, 112–125. doi: 10.1016/j.ecresq.2018.04.001
- Passolunghi, M. C., and Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: a longitudinal study from kindergarten to first grade. *Br. J. Educ. Psychol.* 82, 42–63. doi: 10.1111/j.2044-8279.2011.02039.x
- Passolunghi, M. C., Lanfranchi, S., Altoé, G., and Sollazzo, N. (2015). Early numerical abilities and cognitive skills in kindergarten children. *J. Exp. Child Psychol.* 135, 25–42. doi: 10.1016/j.jecp.2015.02.001
- Peralbo-Uzquiano, M., Fernández-Abella, R., Durán-Bouza, M., Brenlla-Blanco, J. C., and Cotos-Yáñez, J. M. (2020). Evaluation of the effects of a virtual intervention programme on cognitive flexibility, inhibitory control and basic math skills in childhood education. *Comput. Educ.* 159, 104006. doi: 10.1016/j.compedu.2020.104006
- Premsky, M. (2003). Digital game-based learning. *Comput. Entertain.* 1, 21–21. doi: 10.1145/950566.950596
- Ramani, G. B., Jaeggi, S. M., Daubert, E. N., and Buschkuhl, M. (2017). Domain-specific and domain-general training to improve kindergarten children's mathematics. *J. Numer. Cogn.* 3, 468–495. doi: 10.5964/jnc.v3i2.31
- Re, A. M., Benavides-Varela, S., Pedron, M., De Gennaro, M. A., and Lucangeli, D. (2020). Response to a specific and digitally supported training at home for students with mathematical difficulties. *Front. Psychol.* 11:2039. doi: 10.3389/fpsyg.2020.02039
- Reynvoet, B., Smets, K., and Sasanguie, D. (2016). "Number sense: what's in a name and why should we bother?" in *Continuous Issues in Numerical Cognition. How Many or How Much*, ed A. Henik (London; San Diego, CA; Cambridge, Oxford: Elsevier), 195–214.
- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mApping in the newborn chick resembles humans' mental number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379
- Schacter, J., and Jo, B. (2017). Improving preschoolers' mathematics achievement with tablets: a randomized controlled trial. *Math. Educ. Res. J.* 29, 313–327. doi: 10.1007/s13394-017-0203-9
- Schacter, J., Shih, J., Allen, C. M., DeVaul, L., Adkins, A. B., Ito, T., et al. (2016). Math shelf: a randomized trial of a prekindergarten tablet number sense curriculum. *Early Educ. Dev.* 27, 74–88. doi: 10.1080/10409289.2015.1057462
- Schenke, K., Redman, E. J., Chung, G. K., Chang, S. M., Feng, T., Parks, C. B., et al. (2020). Does "measure up!" measure up? evaluation of an iPad App to teach preschoolers measurement concepts. *Comput. Educ.* 146, 103749. doi: 10.1016/j.compedu.2019.103749
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., et al. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev. Sci.* 20, e12372. doi: 10.1111/desc.12372
- Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., et al. (2018). Associations of number line estimation with mathematical competence: a meta-analysis. *Child Dev.* 89, 1467–1484. doi: 10.1111/cdev.13068
- Sheehan, K. J., Pila, S., Lauricella, A. R., and Wartella, E. A. (2019). Parent-child interaction and children's learning from a coding application. *Comput. Educ.* 140, 103601. doi: 10.1016/j.compedu.2019.103601
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–250. doi: 10.1111/1467-9280.02438
- Soto-Calvo, E., Simmons, F. R., Willis, C., and Adams, A. M. (2015). Identifying the cognitive predictors of early counting and calculation skills: evidence from a longitudinal study. *J. Exp. Child Psychol.* 140, 16–37. doi: 10.1016/j.jecp.2015.06.011
- TIMSS (2019). Ministerio de Educación y Formación Profesional. (2020). TIMSS 2019. Estudio Internacional de Tendencias en Matemáticas y Ciencias. Informe español. Available online at: <https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/timss/timss-2019.html>
- Toll, S. W. M., Van Viersen, S., Kroesbergen, E. H., and Van Luit, J. E. (2015). The development of (non-)symbolic comparison skills throughout kindergarten and their relations with basic mathematical skills. *Learn. Individ. Differ.* 38, 10–17. doi: 10.1016/j.lindif.2014.12.006
- van 't Noordende, J. E., Kroesbergen, E. H., Leseman, P. P., and Volman, M. C. J. (2021). The role of non-symbolic and symbolic skills in the development of early numerical cognition from preschool to kindergarten age. *J. Cogn. Dev.* 22, 68–83. doi: 10.1080/15248372.2020.1858835
- Wong, T. T. Y., and Chan, W. W. L. (2019). Identifying children with persistent low math achievement: the role of number-magnitude mapping and symbolic numerical processing. *Learn. Instr.* 60, 29–40. doi: 10.1016/j.learninstruc.2018.11.006
- Wouters, P., Van Nimwegen, C. V., Van Oostendorp, H., and Van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *J. Educ. Psychol.* 105, 249–265. doi: 10.1037/a0031311
- Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., and van Lieshout, E. C. D. M. D. M. (2017). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learn. Instr.* 50, 1–13. doi: 10.1016/j.learninstruc.2016.11.001
- Zhang, X., Räsänen, P., Koponen, T., Aunola, K., Lerkkanen, M.-K. K. M., and Nurmi, J. (2020). Early cognitive precursors of children's mathematics learning disability and persistent low achievement: a 5-year longitudinal study. *Child Dev.* 91, 7–27. doi: 10.1111/cdev.13123
- Zhu, M., Cai, D., and Leung, A. W. S. (2017). Number line estimation predicts mathematical skills: difference in grades 2 and 4. *Front. Psychol.* 8, 1576. doi: 10.3389/fpsyg.2017.01576



OPEN ACCESS

EDITED BY

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

REVIEWED BY

Hao-Chiang Koong Lin,
National University of Tainan, Taiwan
Watcharee Ketpichainarong,
Mahidol University, Thailand

*CORRESPONDENCE

Huei-Tse Hou
hthou@mail.ntust.edu.tw

SPECIALTY SECTION

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

RECEIVED 29 April 2022

ACCEPTED 17 August 2022

PUBLISHED 12 September 2022

CITATION

Hou H-T (2022) Augmented reality
board game with multidimensional
scaffolding mechanism: A potential
new trend for effective organizational
strategic planning training.
Front. Psychol. 13:932328.
doi: 10.3389/fpsyg.2022.932328

COPYRIGHT

© 2022 Hou. 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.

Augmented reality board game with multidimensional scaffolding mechanism: A potential new trend for effective organizational strategic planning training

Huei-Tse Hou*

Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei City, Taiwan

KEYWORDS

scaffolding, strategic planning, game-based learning (GBL), board game, augmented reality, situated learning, formative assessment

Introduction

The global trend in the field of employee training and educational innovation emphasizes the training of key competencies (Halász and Michel, 2011; Tiana et al., 2011; Valle and Manso, 2013). The new training trend emphasizes training programs that develop key competencies for organizational members and use these competencies to appropriately solve complex situational problems. Among which, strategic planning is an important ability for governments, businesses, and organizations facing dynamic situations (Bryson et al., 2010; Poister, 2010). Especially in the current complex international environment, with rapid changes in politics, epidemics, wars, economies, and natural disasters, it is crucial to cultivate collaborative and cross-disciplinary talents with sensitivity. Strategic planning can be applied to all areas of strategy development practice.

In the teaching of strategic planning, the provision of contextualized examples is particularly crucial because of the need to analyze specific case contexts. Situated learning (Brown et al., 1989) emphasizes that learners learn in real or simulated contexts, using real work as the main axis of learning, and using context, clues, scaffolding, and diagnostics to promote the depth of learning (Zydneva et al., 2012; Hou and Keng, 2021). If the contextual context and strategic planning tasks are provided and the external environment and internal constraints of the organization are simulated, it is expected that learners can develop the ability to analyze and organize information to make strategic planning and further achieve learning transfer. However, the design of situated learning requires clear goals, feedback, and interaction as well as the ability to motivate learners and provide

appropriate support tools to achieve better learning outcomes (Norman, 2014). To promote students' motivation and cognitive engagement in contextual learning, the use of game-based learning is a promising approach. Currently, studies on game-based learning have found positive benefits of games for teaching and learning (Annetta et al., 2009; Hou and Li, 2014; Hou, 2015; Hou and Keng, 2021).

Business simulation games emphasize that players evaluate the benefits, costs, and outcomes of their decisions and learn from analyzing the contextual information of the game (Doyle and Brown, 2000). To simulate real face-to-face interactions in organizations, board games that emphasize physical interactions are suitable for teaching organizational strategic planning. The game mechanics of board games often emphasize interpersonal interaction and mutual collaboration (LeBlanc and Bearison, 2004), and many board games emphasize that players must think strategically to win the game. Several scholars have conducted research on educational board games and found that board games are a positive aid to teaching and learning (von Wangenheim et al., 2012; Hou and Keng, 2021). However, in the context of strategy planning with board games, players often need to analyze more information about the changing context of the external context of the organization. In this case, the use of augmented reality (AR) should help to provide a combination of physical cards and online information.

The strategy-planning game should not only provide participants with a contextual experience but also focus on reflection after the experience. The reflective and adaptive behavior of the participants in the game is crucial, as it can stimulate new strategy ideas (Kiili, 2007) or lead to reflective thinking and behavior (Hou, 2015). A multidimensional scaffold can help learners become more contextualized or engage in cognitive thinking (e.g., Hou et al., 2022). Scaffolds include various types, including cognitive concept scaffolds (guidance on knowledge content concepts), meta-cognitive scaffolds (guidance on how to think), procedural scaffolds (guidance on how to use resources or tools), and strategic scaffolds (guidance on problem-solving strategies) (Hannafin et al., 1999).

In strategy planning, the provision of both meta-cognitive scaffolding and strategy scaffolding is critical. These two types of scaffolds focus on providing learners with guidance on self-planning, monitoring, and evaluation of in-game strategies or lead to thinking about additional problem-solving strategies. In this regard, a board game alone would not be able to provide timely dynamic simulations of a large amount of business data and graphs that follow the implementation of in-game strategy planning and provide learners with immediate meta-cognitive guidance based on this data (e.g., guidance on how to adjust their thinking about strategies). In addition, these scaffolds are useful for diagnosing the learning experiences as formative assessments, which provide feedback and enhance learning performance, and for improving learning motivation and effectiveness. The use of AR technology to provide these

scaffolds and diagnostic assessments can provide additional features to stimulate reflection and strategic thinking.

Currently, there is a lack of board games dedicated to strategic planning ability development and a lack of strategic planning games that provide multiple dimensions of scaffolding at the same time. Therefore, this study proposes to use AR technology, situated learning, and multi-dimensional scaffolding theory to plan a multi-scaffolding-oriented AR educational board game framework for strategic planning ability training.

A multidimensional scaffold-oriented AR board game framework

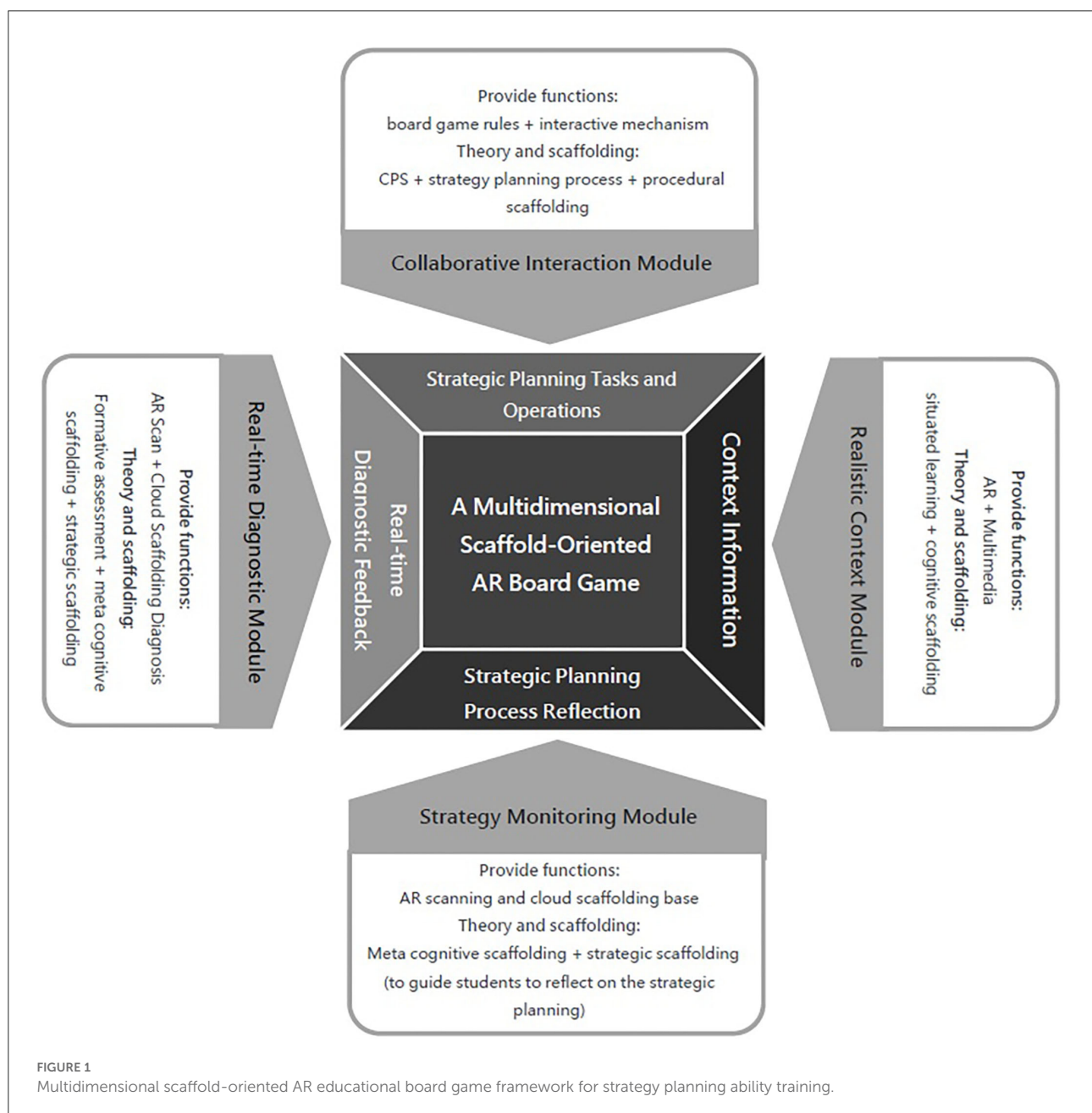
This study proposes a multidimensional scaffold-oriented AR educational board game framework for strategy planning ability training. This framework integrates multi-dimensional scaffolding (including cognitive scaffolding, metacognitive scaffolding, strategic scaffolding, peer scaffolding, and procedural scaffolding), AR technology, and board game interaction mechanism to provide an integrated game-based interactive environment for players to learn strategic planning in the game.

As shown in the Figure 1 below, the entire framework is divided into four modules, which provide four functions: realistic situations, strategy monitoring, real-time diagnosis, and collaborative interaction. Among them, the *realistic context module* will apply the design principles of situated learning (Brown et al., 1989), provide a realistic contextual context through multimedia and AR technology, and provide a cognitive scaffold (Hannafin et al., 1999; Saye and Brush, 2002; Hou and Keng, 2021) to provide a variety of external and internal contextual contexts and scenarios (e.g., news, environmental economic events, competitor news, internal organizational temporal events, or data models) for managing the context as a contextual context for tasks and clues. These clues can be used as cognitive scaffolds to facilitate learning and thinking. In addition, a physical board game is used as a *collaborative interaction module* to design collaborative interaction game rules based on the collaborative problem solving theory (CPS) (Nelson, 1999) to guide groups of learners in the same camp to collaborative problem solving, in which the game card content, game rules, and goal tasks need to be integrated into each stage of the strategy planning process (Hill et al., 2016) to facilitate the alignment of game goals with learning objectives. Players need to conduct resource planning or process planning within the organization to achieve the goals of the strategic planning task. The guide to the strategic planning process can be used as a procedural scaffold to assist players in the complete process of strategic planning.

The *strategy monitoring module* provides a meta-cognitive scaffold that allows learners to scan or enter decision data on specific cards using AR. The module can provide corresponding cues before, during, and after the execution of a strategy based on the data. Finally, the *real-time diagnostic module* provides a formative assessment of the diagnostic and feedback after reviewing the implementation of the player's proposed strategy. Through the scanning of specific combinations of different board cards and the cloud-based scaffolding diagnostics (e.g., diagnostics of strategies and outcomes formed by players by combining related action cards), the app will display various

triggered outcomes after the execution of strategies (e.g., diagnostic feedback on the latest value changes and real-time conditions resulting from the implementation of strategy combinations). These formative evaluations can also be used as meta cognitive scaffolds and strategy scaffolds to guide learners to make more in-depth strategy adjustments.

Instructors can use this framework to design board game cards and game interaction rules (e.g., matching, combination, sorting, and other game mechanisms) according to learning objectives. With the common AR picture recognition editing tools on the Internet, various multimedia scaffold designs can be



made, and a scaffold-oriented AR educational board game can be completed.

Discussion

Previous research has found that to promote organizational learning and innovation, there should be a shift to teaching models that promote critical thinking and analytical skills (Dirkx et al., 2006). Previous research has found that strategic planning training lacks experience and feedback in realistic contexts (Ganesh and Sun, 2015), and gaming activities using AR technology in multimedia contexts with multidimensional scaffolding (e.g., Hou et al., 2022) are expected to help address this issue. There is a lack of educational game design framework that combine face-to-face collaborative interactions with multimedia contextual realistic organizational strategy planning activities. The proposed multi-dimensional scaffolding combined with strategy planning theory, situated learning theory, and augmented reality game guidance model should be theoretically innovative and important.

Author contributions

H-TH: conceptualization, investigation, and writing—original draft.

References

- Annetta, L. A., Minogu, J., Holmes, S. Y., and Cheng, M. T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Comput. Educ.* 53, 74–85. doi: 10.1016/j.compedu.2008.12.020
- Brown, J. S., Collins, A., and Duguid, P. (1989). Situated cognition and the culture of learning. *Educ. Res.* 18, 32–42. doi: 10.3102/0013189X018001032
- Bryson, J. M., Frances Stokes, B., and Kaifeng, Y. (2010). The state of public strategic management research: a selective literature review and set of future directions. *Am. Rev. Public Adm.* 40, 495–521. doi: 10.1177/0275074010370361
- Dirkx, J. M., Mezirow, J., and Cranton, P. (2006). Musings and reflections on the meaning, context, and process of transformative learning: a dialogue between John M. Dirkx and Jack Mezirow. *J. Transform. Educ.* 4, 123–139. doi: 10.1177/1541344606287503
- Doyle, D., and Brown, F. W. (2000). Using a business simulation to teach applied skills—the benefits and the challenges of using student teams from multiple countries. *J. Eur. Ind. Train.* 24, 330. doi: 10.1108/03090590010373316
- Ganesh, G., and Sun, Q. (2015). Using simulations in the undergraduate marketing capstone case course. *Mark. Educ. Rev.* 19, 7–16. doi: 10.1080/10528008.2009.11489054
- Halász, G., and Michel, A. (2011). Key Competences in Europe: interpretation, policy formulation and implementation. *European Journal of Education* 46, 289–306. doi: 10.1111/j.1465-3435.2011.01491.x
- Hannafin, M., Land, S., and Oliver, K. (1999). “Open learning environments: foundations, methods, and models,” in: *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*, Vol. II, ed. C. M. Reigeluth (Mahwah, NJ: Lawrence Erlbaum Associates), 115–40.
- Hill, C. W. L., Jones, G., and Schilling, R. M. A. (2016). *Strategic Management: Theory*. Nashville, TN: South-Western Pub.
- Hou, H. T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: a video-based process exploration. *Comput. Hum. Behav.* 48, 424–435. doi: 10.1016/j.chb.2015.02.010
- Hou, H. T., Fang, Y. S., and Tang, J. T. (2022). Designing an alternate reality board game with augmented reality and multi-dimensional scaffolding for promoting spatial and logical ability. *Interact. Learn. Environ.* doi: 10.1080/10494820.2021.1961810
- Hou, H. T., and Keng, S. H. (2021). A dual-scaffolding framework integrating peer-scaffolding and cognitive-scaffolding for an augmented reality-based educational board game: an analysis of learners' collective flow state and collaborative learning behavioral patterns. *J. Educ. Comput. Res.* 59, 547–573. doi: 10.1177/0735633120969409
- Hou, H. T., and Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Comput. Hum. Behav.* 30, 29–38. doi: 10.1016/j.chb.2013.07.052
- Kiili, K. (2007). Foundation for problem-based gaming. *Br. J. Educ. Technol.* 38, 394–404. doi: 10.1111/j.1467-8535.2007.00704.x
- LeBlanc, G., and Bearison, D. J. (2004). Teaching and learning as a bi-directional activity: investigating dyadic interactions between child teachers and child learners. *Cogn. Dev.* 19, 499–515. doi: 10.1016/j.cogdev.2004.09.004
- Nelson, L. M. (1999). “Collaborative problem solving,” in: *Instructional Design Theories and Models: A New Paradigm of Instructional Theory*, ed. C. M. Reigeluth (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 241–267.
- Norman, D. (2014). *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Diversion Books.

Funding

This research was supported by the projects from the US Air Force Office of Scientific Research (AFOSR) project (20IOA038) and the Ministry of Science and Technology, under contract numbers MOST-108-2511-H-011-003-MY3 and MOST-110-2511-H-011-004-MY3.

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.

Poister, T. H. (2010). The future of strategic planning in the public sector: linking strategic management and performance. *Public Adm. Rev.* 70, S246–S254. doi: 10.1111/j.1540-6210.2010.02284.x

Saye, J. W., and Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educ. Technol. Res. Dev.* 50, 77–96. doi: 10.1007/BF02505026

Tiana, A., Moya, J., and Luengo, F. (2011). Implementing key competences in basic education: reflections on curriculum design and development in Spain. *Eur. J. Educ.* 46, 307–322. doi: 10.1111/j.1465-3435.2011.01482.x

Valle, J., and Manso, J. (2013). Key competences as a trend in the supranational educational policies of the European Union. *Rev. Educ.* 12–33. doi: 10.4438/1988-592X-RE-2013-EXT-255

von Wangenheim, C. G., and Savi, R., and Borgatto, A. F. (2012). DELIVER! – an educational game for teaching earned value management in computing courses. *Inf. Softw. Technol.* 54, 286–298. doi: 10.1016/j.infsof.2011.10.005

Zydneya, J. M., Bathkeb, A., and Hasselbringc, T. S. (2012). Finding the optimal guidance for enhancing anchored instruction. *Interact. Learn. Environ.*, 22, 668–683. doi: 10.1080/10494820.2012.745436



OPEN ACCESS

EDITED BY

Claudio Longobardi,
University of Turin, Italy

REVIEWED BY

Musa Nushi,
Shahid Beheshti University, Iran
Lawrence Jun Zhang,
The University of Auckland,
New Zealand
Fengjuan Zhang,
Jilin University, China

*CORRESPONDENCE

Luxin Yang
yangluxin@bfsu.edu.cn

SPECIALTY SECTION

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

RECEIVED 30 April 2022

ACCEPTED 08 August 2022

PUBLISHED 22 September 2022

CITATION

Kang Y and Yang L (2022) Examining
EFL teachers' changing conceptions
of research: A case study of a
continuing professional development
program in mainland China.
Front. Psychol. 13:933061.
doi: 10.3389/fpsyg.2022.933061

COPYRIGHT

© 2022 Kang and Yang. 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.

Examining EFL teachers' changing conceptions of research: A case study of a continuing professional development program in mainland China

Yan Kang¹ and Luxin Yang^{2*}

¹English Education Department, College of Foreign Languages, Capital Normal University, Beijing, China, ²School of English and International Studies, The National Research Centre for Foreign Language Teaching Materials, Beijing Foreign Studies University, Beijing, China

Drawing on teacher interviews, written teacher reflections, teacher research proposals, and research papers, this study explored the outcome and process of teacher learning during their participation in a continuing professional development program. It has been found that the teachers changed their conceptions of research over the course of learning concerning the nature, purpose, and process of research and the relationship between teaching and research. Dialogic exchanges and reading research papers, along with the scaffolding of the teacher educator, enabled the teachers to validate their practices, link others' perspectives up to their own, and re-situate research in light of their current practices. The findings provide insights into the nature of teachers' conceptual change and how learning opportunities can be better built into continuing professional development programs.

KEYWORDS

continuing professional development, conceptions of research, teacher change, EFL teachers, teaching and research

Introduction

Continuing professional development has been a focus for curriculum development and teacher education research since the 1970s. There has been an extensive body of literature that explores the issue in terms of both theory and policy. In recent years, an increasing number of studies have sought to examine empirically teachers' experiences of learning in such activities. However, few studies evaluate actual programs in specific contexts (Kiely and Davis, 2010; Borg, 2018). Researchers argued that high-quality continuing professional development programs should be designed based on a thorough understanding of what enables teachers to change and how teachers change as a result

of learning in such programs (Borko et al., 2010; Sahin and Yildirim, 2015; Labone and Long, 2016; Tabatabaee-Yazdi et al., 2018; Hayes, 2019). Therefore, more studies need to probe into the program procedures that support teacher learning.

To address this need, the present study followed up on a group of Chinese school EFL teachers in a continuing professional development (CPD) program organized by the local EFL Teaching and Research Office (TRO) in a district of Beijing, China, with the aim of introducing reading-writing-integrated instructional principles (Hirvela, 2016) and improving the participating teachers' teaching effectiveness through classroom research. Designed and implemented by the second author, an experienced teacher educator who specializes in EFL teaching and teacher education, the CPD program consisted of 4 monthly workshops, totaling 24 hours. The design of the program was based on two central assumptions about teacher learning. First, learning is effective when teachers share experiences and jointly explore their practices in collaborative processes. The collaborative dialog constitutes the social context that supports the construction of professional knowledge. Second, learning is furthered by reading research articles relevant to teachers' teaching and research contexts. Therefore, this study in particular examined whether changes took place in teachers' conceptions of research as a result of their participation in the CPD program and how the teachers utilized the learning opportunities afforded by the program to (re)construct their conceptions.

Literature review

Teachers' conceptions of research

Since the 1960s, there has been an emergent drive to engage teachers more fully in educational research to improve classroom teaching and learning and, to a broader extent, promote their autonomous professional development. The underlying rationale is that when teachers engage with (through reading) and in (by doing) research, they will be able to not only make pedagogical decisions informed by sound research evidence but also play an ever more active role in curriculum development (Stenhouse, 1975; Hargreaves, 2001; Borg, 2010). Borg (2010) defined teacher research as "systematic inquiry" conducted by teachers in their professional contexts and pointed out that engagement in research can enhance teachers' understanding of their work and, consequently, has the potential to promote quality teaching and learning in individual classrooms and inform institutional improvement and educational policy (p.395).

Stimulated by this interest in encouraging teachers to become research-engaged, researchers in the field of general education began to examine teachers' conceptions of research such as what they think of research, the specific meanings they

attach to research, and how such understanding influences their professional life (Shkedi, 1998; Everton et al., 2000, 2002). In the past decade, the same strand of inquiry began to emerge in the literature on English language teaching. One underpinning argument has been that efforts to promote teacher research engagement are more likely to be rewarded if they are based on an understanding of teachers' conceptions of research and the role played by research in their work (Borg, 2009).

These studies, echoing those conducted outside ELT, revealed that English teachers worldwide held different conceptions of research, with some adhering to conventional positivist notions of scientific inquiry and others equating research with a teaching-oriented process; they often read or wrote less formal research articles than teaching-oriented articles that share innovative classroom activities and teaching approaches; their research engagement was largely driven by institutional requirements or considerations of promotion, rather than their own need for professional development (e.g., Allison and Carey, 2007; Borg, 2009; Gao et al., 2011; Gao and Chow, 2012; Vu, 2021; Gironzetti and Muñoz-Basols, 2022). Current research also revealed some barriers to teacher research, including lack of time, lack of awareness about the importance of research, lack of research skills, lack of mentorship, limited funding and resources, non-collaborative school culture, leadership attributes, and political issues (Barkuizen, 2009; Gao and Chow, 2012; Borg, 2013, 2010; Borg and Liu, 2013; Rahimi and Weisi, 2018; Farsani and Babii, 2019; Sato and Loewen, 2019; Alhassan and Ali, 2020; Li and Zhang, 2022). It has been suggested that teachers would be more inclined to change their conceptions about research, be better able to cope with the different constraints in their contexts, and consequently become more autonomous in their professional development when they were supported in their research engagement either by their colleagues or university teacher educators (Banegas et al., 2013; Wang and Zhang, 2014; Yuan and Lee, 2015; Dimmock, 2016; Cornelissen et al., 2017).

Continuing professional development

It has been widely acknowledged that continuing professional development (CPD) plays an important role in improving or changing teachers' cognition and practices. Kelchtermans (2004) defined CPD as a learning process that leads to "changes in teachers' professional practice (actions) and in their thinking about that practice" (p.220). The term is adopted in the present study as it distinguishes the development of teachers throughout their career from their professional development during the teacher education and induction phases and connotes "a broader range of developmental possibilities" than one-shot in-service training that targets the transfer of certain content to teachers (Wermke, 2011, p. 666). Current research on continuing professional development tends to follow two thematic lines: one that focuses on the process of

learning and the other that focuses on the product. The former targets the dynamics of knowledge construction, whereas the latter examines the implications of CPD activities for knowledge building, for example, the conditions that afford changes.

Teacher change in CPD programs is widely reported in general education research (Park and So, 2014; Ottley et al., 2015; Turner et al., 2017; Zimmerman et al., 2017). However, the volume of such research is limited in the ELT context (Borg, 2011). Within that small number of studies, the results however were quite mixed, with some reporting stability (Lamb, 1995; Kubanyiova, 2006; Sim, 2011; Ming, 2019), some providing evidence of significant change (Lamie, 2004; Kiely and Davis, 2010; Teng, 2016; Karimi and Zade, 2018; Gorter and Arocena, 2020; He and Bagwell, 2022), and others identifying different levels of change across individuals (Tillema and van der Westhuizen, 2006; Borg, 2011; So, 2013; Chaaban, 2016; Coburn and Borg, 2022). Borg (2011) pointed out that the impact of language teacher education on teachers deserves much more empirical attention as our understanding of the issue remains incipient. Therefore, additional research is needed to identify the conditions under which professional development opportunities are connected to teacher change.

Differences exist in these studies regarding program characteristics, research methodology, and how change is operationalized. Nevertheless, existing studies do reveal some conditions that lead to effective professional development. In a review of 38 professional development studies, Kalinowski et al. (2019) highlighted three broad areas of effective program characteristics: (1) structural features such as multiple delivery formats, expert involvement, and consideration of teachers' needs, interests, and existing knowledge; (2) content-related features such as application-oriented knowledge; and (3) didactic features such as cooperation and collaboration, input, application and reflection, active learning, and supplementary materials. These features were consistent with previous research findings that argue for practice- and inquiry-based collaborative programs that characterize such components as peer observations, experience sharing, and meaningful discussions (e.g., Garet et al., 2001; van Veen et al., 2012; Park and So, 2014).

Drawing on data from interviews and written reflections, Kiely and Davis (2010) found that the collaborative component of the CPD program worked well with the teachers as they understood the context of collaboration as supportive and reassuring and based on shared values. However, the researchers found that the use of reading was not as successful as it had been expected, which they attributed to a lack of space for reading in teachers' working lives and limited relevance of the reading to teachers' daily practices. Likewise, Park and So (2014) proved the effectiveness of such activities as peer observations, individual reflections, and peer interactions in a school-initiated collaborative professional development program. In a study with Irian EFL teachers, Tabatabaee-Yazdi et al. (2018) proved

that CPD components that lead to teachers' success include the opportunities to update their knowledge, collaborate with others, reflect on teaching, and make important decisions in their classrooms.

Regarding the research on the change process, there are a relatively small number of studies that capture the dynamics of knowledge construction during teacher learning (e.g., So, 2013; Selvi and Martin-Beltrán, 2016; Choi, 2015, 2017; Ming, 2019). Adopting a sociocultural perspective, the researchers postulated that teacher change is a process of development or learning achieved by teachers as learners working in a learning community, and an outcome of complex interactions between the individuals and the various agents in that community, including supervisors, colleagues, teacher educators, and students (Clarke and Hollingsworth, 2002). Therefore, rather than equating "change" with a radical shift, the current study acknowledges the cumulative nature of teacher change and defines the teacher change in conceptions as a gradual process that allows varying degrees of intensity on a developmental continuum.

In a study on student teachers' cognition development, Cabaroglu and Roberts (2000) identified 11 categories of change processes, namely, awareness, confirmation, elaboration, addition, re-ordering, relabeling, linking up, disagreement, reversal, pseudo-change, and no change. Drawing on several related studies, Tillema and van der Westhuizen (2006) proposed three criteria to promote teachers' collaborative knowledge construction: (1) raising awareness and gaining further understanding and insights; (2) recognizing others' viewpoints as relevant and valid to achieve perspective; and (3) involving in the group process and showing interest in using the practical results. In a longitudinal study examining the process of teacher learning, Bakkenes et al. (2010) identified six kinds of thinking processes that led to varying learning outcomes, among which three were associated with the positive aspects of learning, namely, experimenting, considering one's practice, and getting ideas from others. Focusing on teachers engaged in a post-graduate-level teacher education course, Selvi and Martin-Beltrán (2016) found that teachers engaged in three kinds of active learning processes: externalizing their stance and challenging theory, struggling with contrasting perspectives, and re-situating theory in light of past, present, and future teaching-learning contexts. The researchers emphasized that the praxis building of teachers is dynamic, multifaceted, evolving, and negotiated in context. In a comparative study on a short-term certification program in South Korea and a comparable long-term program in the United States, Choi (2017) summarized three interrelated patterns of cognitive change, namely, capturing and repositioning their assumptions, gaining and seeking pedagogical implications, and inner conflicts and reconciliation.

It can be concluded that our understanding of teachers' sense-making processes remains inadequate.

Tillema and van der Westhuizen (2006) warned that a lack of agreement regarding the nature, process, and outcome of teacher learning would lead to less precise prospects for knowledge building in the profession of teaching. Therefore, more research is needed to capture “how teachers come to know what they know” and “how certain concepts in teachers’ consciousness develop over time” (Johnson, 2009, p.17), particularly in the context of continuing professional development. To contribute to this domain, the current study adopts a qualitative case study approach (Merriam, 2001) to explore both the outcome and the process of learning during teachers’ participation in a CPD program. We specifically focus on teachers’ changes in their conceptions of research throughout the program. The study was guided by the following research questions:

1. Will the participating teachers’ conceptions of research change over the course of the CPD program? If yes, what kind of change has taken place?
2. What are the program characteristics that contribute to the change?
3. How do the teachers reconstruct their conceptions of research during their participation in the program?

The study

Participants

The participating teachers in the program were 33 senior high school EFL “gugan” (literally “backbone” in Chinese) teachers, that is, outstanding teachers selected by their schools and accredited by the local Education Commission. Their participation in the program was mandatory for the renewal of the title. A total of five demographically diverse teachers were purposefully chosen as the “focal” cases (Merriam, 2001) in the current study, who exhibited a high degree of involvement in the program since the commitment to collaboration has been considered a crucial factor that may affect the results of learning (Tillema and van der Westhuizen, 2006). They were pinpointed according to the quantity and quality of their verbal contributions to the group activities after the first author calculated and reviewed the valuable insights (i.e., ideas that were related to the input of the workshop and the connections they made to their contexts) generated by each participating teacher in the first workshop, based on the teachers’ oral debriefings after group discussions and their first written reflections. The focal teachers were approached by the first researcher before the second workshop and voluntarily agreed to take part in the study. Table 1 presents the profiles of the participants.

All five participants were women, whom we called Lucy, Jane, Wanda, Frances, and Sunny. They had a variety of research experiences, such as writing or publishing research papers and/or engaging in research projects. The more experienced they were, the more extensively and intensively they had been involved in research. This is understandable as engagement in research is one of the major conditions to qualify for the title of “gugan” teacher in China.

Data collection

Data in this study included teachers’ written reflections, research proposals and manuscripts of research papers submitted in fulfillment of the requirements of the program, and interviews. After each workshop, every participating teacher was required to submit a written reflection on their learning experiences following a KWLH (what I know, what I want to know, what I have learned, and how I put it into practice) protocol. At the end of the program, the participating teachers were expected to submit a proposal on researching reading-writing-integrated EFL instruction, based on which a full-length research paper was developed within 2 months after the completion of the program. The preparations for the final research paper permeated the whole program. In each workshop, the teachers discussed the selection of research topics and the choices of research design with their peers and the teacher educator based on the workshop input (e.g., lectures and reading). By the end of the program, most of the teachers have finished their data collection. The tasks that remained after they exited the program were to analyze the data and draft the paper.

Upon the completion of the program, the first author held semi-structured interviews individually with each focal teacher. The interviews covered the following topics: (1) learning experiences in the program, (2) past experiences of research and future research plan, (3) difficulties in conducting research, (4) understanding of research, and (5) understanding of EFL reading and writing. The interview data were transcribed by a research assistant.

The data sets constructed included (1) 20 written reflections (four for each teacher), (2) five audio-recorded and transcribed individual teacher interviews, and (3) five research proposals and research papers. The multiple sources of data enabled the researchers to triangulate the data and guarantee the trustworthiness of the study (Miles and Huberman, 1994).

The researchers were aware that the use of data from interviews and program assignments would introduce accommodation bias or the possibility that the participants would make an effort to generate examples of how the program had influenced them simply to be helpful (Kennedy, 2008). Therefore, the focal teachers were informed that the purpose of the study was to help the teacher educator improve the future design of the program, and they were encouraged to

objectively comment on their learning experiences. As the program was not supposed to give teachers grades, there was no conflict of interest.

Data analysis

Multiple sources of data were first analyzed individually and then combined and reanalyzed in a qualitative manner (Miles and Huberman, 1994). The process involves three steps. First, we read through all the data to reveal salient themes and recurrent patterns, and then the data from each teacher were reanalyzed based on these themes. The coding was conducted both deductively by using *a priori* categories evolved out of the literature and the research questions, and inductively by identifying the concepts that emerged from the data. Second, we conducted a cross-analysis of data, in which meaningful themes emerging from the interview data were compared with themes emerging from teacher reflections. The themes were then triangulated with teachers' research proposals and research papers. The cross-check produced 10 sub-themes, namely, "conceptions about the nature of research," "conceptions about the purpose of research," "conceptions about the relationship between teaching and research," "conceptions about the process of research," "peer interaction," "reading," "scaffolding," "validating," "linking up," and "re-situating," which were grouped into three broad categories: teachers' conceptions of research, factors that contribute to the change, and the process of change. Based on the final themes and sub-themes, data from the five focal teachers were scrutinized to identify changes in teachers' conceptions, which led to the creation of [Appendix Table 1](#). Third, the final results from the multiple sources of data were cross-compared numerous times to validate the findings.

All the data were analyzed collaboratively by the two researchers. Each author first analyzed the data independently and then compared the results to enhance the credibility. The comparisons and discussions allowed the research team to agree on all the themes and quotes. Meanwhile, member checking was conducted to verify an accurate interpretation of the data, and where necessary, clarification was obtained from the teachers.

Findings and discussion

Changes in teachers' conceptions of research

Data analysis has revealed that the five focal teachers reconstructed their conceptions of research in four major areas: (1) the nature of research, (2) the purpose of research, (3) the relationship between teaching and research, and (4) the process of research. [Appendix Table 1](#) presents the teachers'

conceptions before and after the program with abridged quotes from teacher reflections and/or interviews.

First, the teachers developed a broader awareness of the different forms that research can take and the varying approaches to research. On entry into the program, the teachers viewed research as something "grand," or academic, which should build on "sound theories and rich data," employ "complicated statistical tools," and produce "generalizable results." The view aligns with a positivist notion of scientific research and resonates with the findings of previous studies conducted both in China and in other countries (Borg, 2009; Gao et al., 2011; Gao and Chow, 2012).

At the end of the program, the teachers' understanding of the nature of research became more comprehensive. They realized that research could be "small," or teaching-oriented; it should "stand on the solid ground of teaching" and include "studies about any aspects of teaching with various methods and tools available," for example, action research, lesson study, and narrative inquiry. Sunny's comments reflected the emerging conceptions:

"All that we have experienced in our everyday practice can become the subject of research. Research is to transform the phenomena that we observe into something that can better guide our teaching through rational thinking and systematic organization." (Sunny, interview)

Meanwhile, the learning experience in the program resolved the teachers' uncertainties about the nature of the research. For example, in her first written reflection, Jane reviewed her previous research experiences and questioned the "research" that she has done before:

"...we wrote 'papers' summarizing teaching almost every semester and shared them with our colleagues. I'm not sure whether these can be called research." (Jane, reflection 1)

When she was asked the same question in the follow-up interview, Jane responded with more certainty:

"Now, I think what we have done IS research. I understand 'research' from two perspectives. On the one hand, research helps us find feasible ways to teach. On the other hand, it helps us solve practical problems in the classroom. In this sense, I think we do well in research because we have done a lot to improve teaching. What we lack is the ability to turn everyday practice into research." (Jane, interview)

Second, the teachers' conceptions about the purpose of research shifted from "instrumental" to "integral." In line with the previous studies (e.g., Gao et al., 2011), all the focal teachers agreed that research was beneficial to their professional growth. But they engaged in research mainly to meet the administrative

requirements for appraisal and promotion. As a result, they felt overburdened by all kinds of “research assignments.”

In the program, the teachers realized that research should not become an ‘extra burden.’ Both reading and doing research are to solve problems in their teaching practices. Jane’s written reflection illustrated the point:

“We seldom think about what to write until it’s almost the deadline. . . But now, I think, teaching comes first. When you have done something new or meaningful, you will be eager to share it with other teachers. Then you start to write. On the other hand, if you’re faced with problems, go and read the literature. You can find the solutions from other teachers’ experiences.” (Jane, reflection 3)

Sunny agreed with the idea that the purpose of doing research is to “publish,” or to “share your experiences with other people so that the whole community of teaching can reap the benefits” (interview). She argued that research should not become something that is imposed on teachers:

“Research embodies your personal feelings and emotions because you put your time and energy into it. It’s something personal, not institutional.” (Sunny, interview)

Third, along with the change in the conceptions about the purpose of research, the teachers’ views about the relationship between research and teaching have also been elaborated. Before the program, their understanding was unidirectional, mainly focusing on what research could do for teaching such as recording, sharing, and improving teaching. They thought that research was independent of teaching, so they had to “balance” teaching and research in their busy schedules. After the program, all the focal teachers agreed that “research is interwoven with the actual teaching process.” Jane reflected,

“Research is a synthesis of teaching. Teaching feeds research and research nourishes teaching. They are nurturing each other.” (Jane, reflection 2)

The idea was supported by Frances who wrote in her final reflection that “teaching is research and vice versa” and further explained the idea in the interview:

“They (teaching and research) are closely related and can never be separated. When a teacher takes a full record of her practice and analyzes it systematically, she is doing research. To research is to record, reflect and refine teaching.” (Frances, interview)

Sunny described her understanding of the relationship between research and teaching in her final reflection and used a metaphor in the follow-up interview to illustrate the idea:

“...our research is grounded in teaching and teaching enriches research. We naturally turn into a teacher-researcher.” (Sunny, reflection 4)

“Research is salt, without which teaching would be tasteless. An appropriate amount of research would make your teaching career tasty.” (Sunny, interview)

Wanda revealed the changes in her understanding of the relationship between teaching, research, and theory in her third reflection:

“I had thought that research means to apply theories to practice. . . I have learned that theories also come from teaching. Research helps generalize theories from teaching. It is a bridge between theory and practice.” (Wanda, reflection 3)

Fourth, by giving the teachers a thorough grounding in basic research concepts that they could draw on consistently, such as reading for relevant theories, selecting research paradigms and methods, and systematic gathering and interrogation of evidence from a range of sources, the CPD program consolidated the teachers’ view that research should be conducted systematically.

The focal teachers admitted that their previous approach to research was rather “unsystematic” and the papers they have written were “like diaries.” By the end of the program, most of the teachers were able to articulate the whole research process.

“When we teach with a purpose, or when we have a vision for our students, we begin to design specific plans. Then we implement the plan step by step. This is the process of research.” (Sunny, reflection 4)

In both written reflections and follow-up interviews, three of the focal teachers mentioned the need to “reverse” how they used to conduct research. Lucy wrote in her second reflection that research means “organization and classification of materials,” and she elaborated on the point in the interview:

“I usually put my teaching materials in different folders on my computer and label them according to the contents. For example, one folder contains my teaching plans. The other contains my students’ work. When I want to write a paper, I have to search through these folders for all the necessary materials. But now I realize that the process should be reversed. I should decide on the research topic first and develop a research plan. In this way, I will be able to determine the types of data I need for researching the

topic and keep a record with a specific focus. The folders will then be organized according to research topics and data types. . . In our busy work, if we can spend some time sorting out and categorizing what we have done, we'll become good researchers." (Lucy, interview)

In the interview, Wanda described the features of the research process according to her understanding of action research cycles:

"Research is cyclical. When we discover a practical problem in our classrooms, we try to find a way to solve it. Then we identify new problems and continue the cycle. In this research process, we improve our teaching practices and deepen our understanding of teaching." (Wanda, interview)

Program characteristics that contribute to the change

A total of three favorable conditions have been found to facilitate teacher change, namely, peer interaction, literature reading, and teacher educator's scaffolding. The following section illustrates how the teachers understand the three program characteristics and how they function to trigger the change.

Peer interaction

Both the written reflections and follow-up interviews revealed that the teachers highly valued the experiences of peer interaction because the collective discussions engaged them in reflective dialogs about their previous teaching and research experiences and their understanding of the program contents. The interaction helped externalize the teachers' private practices and implicit personal theories, functioning as a mediator for collaborative knowledge construction. Through peer interaction, the teachers discovered alternatives to their work and gained new insights into the teaching and research of EFL reading and writing. For example, Lucy wrote in her second reflection that the discussions enriched her repertoire of teaching and enlightened her from time to time. Frances shared a similar idea:

We share our experience and practice. The discussions free our thinking which tends to be suppressed by our daily routine. It's a collision of ideas, from which we discover similarities, look for differences, and draw inspiration. (Frances, interview)

The findings lend support to the argument developed by previous studies that the insights into the topics,

generated through interaction with their peers, enable teachers to enhance or fine-tune the personal knowledge base of their work (Cochran-Smith and Lytle, 1999; Orland-Barak, 2006; Tillema and van der Westhuizen, 2006). In the present study, peer interactions allow the teachers to share different teaching styles and research possibilities, express divergent beliefs and perspectives, and search for alternatives.

In addition, the present study also found that the dialogs were particularly useful to less experienced teachers, who sought to network with their more experienced counterparts and thereby expand their social and professional connections. Jane, the youngest focal teacher in the study, reflected how the collaborative dialogs acquainted her with other teachers:

"Through the discussions, I got to know a lot of outstanding teachers in our group, who came from different schools in our district. They will become my resources. In the future, I know whom I can turn to for help if I encounter any problems." (Jane, interview)

Borg (2014) pointed out that the opportunity to "meet, talk to, exchange ideas with, and learn from other ELT professionals" would make an important contribution to teachers' professional development (p.39). This is especially true in China where "guanxi," a form of social connectivity deeply entrenched in culture involving the continuous exchange of favors among individuals, exerts an intangible influence over individuals' lives and social practices (Bian, 2018). Having a good "guanxi" with these outstanding teachers of a higher level or ranking (e.g., principals or head teachers) means easier access to highly competitive prizes, honors, and more chances for professional development. By connecting "gugan" teachers in possession of both professional expertise and professional resources, the CPD program helped the younger teachers weave and expand their "guanxi" network in the professional community, resulting in more possibilities to achieve greater success in their future careers.

Reading of research studies

This study also found that reading research papers created a positive effect on teacher professional development. In the current CPD program, reading materials were distributed in electronic format several days before each workshop so that the teachers could read by themselves in their private time. During the workshops, the teachers read the materials again in groups with the help of the teacher educator, resulting in three apparent benefits. First, reading enhanced the teachers' research literacy. In the following excerpt, Jane discussed how the reading of research studies acquainted her with the basics of good research:

“They (the papers) showed how research can be done, how we can select the participants and collect the data. . . I got more familiar with the structure of an academic paper. When I’m to write a paper in the future, I will probably read these papers again.” (Jane, interview)

Second, reading encouraged the teachers to analyze their research practices based on the new meanings derived from other teachers’ research perspectives. Sunny described how the reading restored her professional confidence and reduced the feeling of uneasiness that surfaced in her classroom experimentations:

“We’re now incorporating the reading of original English novels in our class. But we’re not sure whether the new method would work. Sometimes we feel uneasy because we have to bear the pressure coming from the students, the parents, and the school administration. When I read about the same practice researched by other teachers in their published papers, I feel reassured. I think we’re probably walking down the right path and simply need more persistence.” (Sunny, interview)

Third, reading prompted the teachers to reflect on their classroom realities and apply the research findings to their contexts. The teachers thought that most of the research studies that they have read in the workshops were practically oriented and therefore “attractive” to them. After reading these papers, they were “eager to apply the results” to their classrooms. Wanda expressed her interest in experimenting with continuation writing tasks in her reflection:

“I think the continuation writing tasks used after reading could be a very useful way to integrate writing with reading. My students are always frustrated with English writing. If I ask them to continue a story after they have finished reading it, they will probably be interested.” (Wanda, reflection 3)

Gao and Chow (2012) argued that teachers need to learn from reading other teachers’ research narratives, the research procedures that were adopted, and the data that were collected in the process. The rich understanding derived from reading inspired the teachers to reflect on their contexts and determine whether and how they can utilize the new information. This study proves the positive effect of reading on helping teachers reconstruct their conceptions of research, as well as their teaching and research practices.

Teacher educator’s scaffolding

This study found that the teacher change is related to the teacher educator’s scaffolding in both peer interaction and reading. On the one hand, the focal teachers highly appraised the teacher educator’s participation in their collaborative discussions. During the discussions, the teacher educator employed mediational strategies such as asking confrontational questions to push the teachers to articulate their ingrained assumptions, using direct questions to uncover what might have been concealed from the teachers and offering comments to expose the teachers to various perspectives. The dialogic interactions triggered the teachers to have in-depth reflection and consequently support their collaborative knowledge (re)construction.

The focal teachers thought the feedback offered by the teacher educator during their group discussions was “timely,” “engaging,” and “stimulating,” helping them detect the connection between practices and theories and rethink the relationship between teaching practices and doing research. Jane reflected on the feedback that she received after group discussions:

“At the beginning, it seemed like casual sharing of teaching practices in groups. But when we finished our oral reports, she (the teacher educator) asked several questions about our ideas and gave some comments. She helped us examine our classroom practices from various researchable perspectives.

TABLE 1 Participant profiles.

Name	Gender	Age	Teaching experience (years)	Educational background	Previous research experience (self-reported number of research papers written)
Lucy	F	30	5	MA in English literature	8
Jane	F	27	5	BA in English education	5
Wanda	F	33	10	BA in English literature	More than 10
Frances	F	39	18	BA in English education	more than 30
Sunny	F	50	33	MA in English education	more than 60

Then we got a lot of new ideas about research and suddenly realized that there could be so much to write about.” (Jane, reflection 2)

In the follow-up interview, Sunny commented on the timing of the teacher educator’s participation:

“Most of us are experienced teachers. Some of the ideas (raised by the teacher educator) are not strange to us. But the timing was perfect. They were offered during or after our discussions when we were actively engaged. We could immediately understand how they can be practically applied to our teaching and research. The suggestions and comments made the whole discussion more fruitful and enlightening.” (Sunny, interview)

It is therefore evident that rather than imposing a preferred new practice on the teachers, the teacher educator paved a new way of seeing, understanding, or interpreting the familiar ideas, rendering them more applicable to the teachers’ contexts.

On the other hand, scaffolding was provided when the teacher educator guided the teachers through their reading of the selected research studies. Instead of plunging the teachers into a sea of literature, the CPD program adopted a guided reading procedure, in which the teachers read and discussed the literature in groups and the teacher educator provided feedback tailored to the needs of the ongoing discussions in the forms of summarizing, questioning, clarifying, and critiquing. With the support of the teacher educator, reading became easier and more rewarding. For example, Lucy discussed her experience of the guided reading:

“When I read the papers at home, I found it very difficult to understand, particularly the theoretical concepts and the internal logic of writing. But with the help of the teacher educator, I came to understand them.” (Lucy, interview)

The effect of the guided reading can be triangulated with Lucy’s description of her mentor during the interview, in which she attributed the master teacher’s skillful lesson design to her “PCK”—a term Lucy has read about in one of the papers:

“My mentor plans her lesson in such a structured way that her class is always enjoyable to students. . . I think it is the PCK that enables her to do so.” (Lucy, interview)

The previous excerpt proved that Lucy not only understood the academic term in the paper but also was able to reflect on her own professional life in the professional discourse. The results indicated that with proper guidance and support, reading can become less cumbersome and more relevant to the teachers.

In the present CPD program, the teacher educator acted as a co-thinker, mediator, and critical friend to scaffold the collaborative efforts of the teachers. With the support and the prompt, the teachers were aided in reflective thinking to articulate their implicit assumptions about teaching and research, scrutinize the different facets of their practices, and explore the perspectives that they might otherwise not be able to see.

Effective participation of teacher educators has long been recognized as a vital requirement of professional programs that aim to transform teachers in both conceptions and practices (Hayes, 2019). Ingvarson et al. (2005) highlighted that timely and insightful feedback on what one is doing or has done is crucial to the reflection on one’s practices and the development of understanding. The current findings confirmed the argument developed by many researchers that continuing professional development characterizing teacher collaboration, reading, and mentoring can function more successfully than those one-shot courses targeting the transfer of certain content to teachers (Mann, 2005; Borg, 2015). The results also highlighted the value of professional support frameworks designed around a collaborative discussion component in CPD programs and accentuated the mediational role played by teacher educators in assisting teachers to develop their practical wisdom (Orland-Barak, 2006; Luneberg and Korthagen, 2009; Sedova, 2017).

Processes of change in teachers’ conceptions

Concerning how teachers reconstructed their conceptions of research, this study found that the focal teachers engaged in three active processes of knowledge (re)construction throughout the CPD program: (1) validating one’s practices, (2) linking others’ perspectives up to one’s own, and (3) re-situating research in light of one’s current practice. The three types of mental activities, varying in the depth and breadth of thinking, share a common denominator, that is, the teachers’ reflection on their past and present experiences in connection with the program input.

Validating my practices: I’m doing right!

The knowledge (re)construction at the lowest level is to validate the practices according to the information obtained in the program, a process in which all the focal teachers have taken to confirm their stance. Belcher (2007) pointed out that teachers are sometimes faced with making pedagogical decisions without complete confidence in their efficacy. The CPD program provided a framework for the teachers to reexamine their experience-based intuitive teaching and research.

In the interviews, the teachers claimed that they “became more confident” as they found other teachers doing “similar things” in different contexts and “explored more possibilities for

future research.” For example, Lucy discussed how the reading of research studies prompted her reflection on practice and triggered her interest in conducting research:

“I have incorporated the reading of original novels in my class for a whole year and published 2 articles. In this program, we have read several articles about extra curriculum reading. Some were written from perspectives similar to my own. Others were from different perspectives... I’ve got many ideas for further study.” (Lucy, interview)

This pattern of thinking has also emerged from the case of Sunny. For many years, Sunny has implemented inquiry-based learning in her teaching or reading. For example, after finishing a textbook unit about the topic of success, she asked her students to read some famous people’s success stories and write group reports on the elements that lead to their success. In this learning process, she acted as a facilitator to help her students select the reading materials and write the reports. During the discussion of potential research topics, Sunny shared her practice with other teachers and was encouraged by the teacher educator to research it. The following excerpt is her reflection on the learning experience:

“I’ve met many objections these years. They said it was too time-consuming. But I continued because my students liked it. When I shared this practice in our discussion, the teacher educator suggested that I do some research on it. She recommended some books about inquiry-based learning. I feel more confident now for I’m doing right!” (Sunny, interview)

Paran (2017) contended that research-oriented or evidenced-based CPD programs can correct unreliable intuition and experience and break the vicious circle of received wisdom. Our study aligns with this argument and further proposes that such a CPD program could connect teachers’ intuitive teaching to research evidence and provide the necessary professional and emotional support to validate their professional wisdom.

Linking others’ perspectives up to my own: I can also do that!

At a higher stage of knowledge (re)construction, the focal teachers utilized the social context provided by the CPD program to locate their positions not only in their own experiences but also in the experiences of others. With the diversified perspectives afforded by the peers, the teacher educator, and the authors of the selected reading, the teachers engaged in an active synthesis of information, sorting out the relevant viewpoints and connecting those with their contexts, based on which the teachers reconstructed their conceptions and practices. For example, both Jane and Frances

reported that they modified their teaching practices according to what they had read and/or discussed. Frances conducted a study on the effectiveness of the new practice in her final research paper.

Belcher (2007) proposed that when teachers read published research articles, they bring not only a fund of knowledge based on prior teaching and research experience but also the current awareness of their local situations, all of which place them in a strong position to judge the relevance and transferability of the pedagogical implications of those studies. The current findings further testified that teachers are inclined to experiment with the relevant ideas in their classrooms and transform their teaching and research practices accordingly.

Moreover, the linkup process also resulted in the incorporation of new research tools into the teachers’ repertoire of research practices. In the interviews, some of the teachers explicitly expressed their interest in the newly acquainted research methods. For example, after reading some research studies conducted with action research and listening to the teacher educator’s lecture about the method, Lucy wrote in her reflection,

“I think it’s a very useful method for teachers. When I’m faced with a problem, I can use action research to solve it.” (Lucy, reflection 4)

In the discussion of research proposals when Sunny shared her experience of the inquiry-based reading and writing project, the teacher educator introduced the method of narrative inquiry and suggested that in addition to drafting lesson study reports and evaluating the effect of the project on students with experiments, the teacher could use narrative inquiry to record her experiences in the implementation of the project. Sunny reflected on the episode in the interview,

“Telling stories is also research. This is new to me. I have learned a lot from this experience (inquiry-based learning). I think I can write my own story and share my experiences with others.” (Sunny, reflection 3)

The current findings confirm the argument that the teachers are interested in concrete ideas, rather than abstract or theoretical terms, and they value the research works that can be applied to their professional contexts (Guskey, 1986; Belcher, 2007; Macalister, 2018; Avidov-Ungar and Herscu, 2020). Although the teachers’ emerging interest in the new perspectives might not necessarily lead to their actual “experimentation,” they at least learned to take ownership of these new ideas when they made conscious efforts to connect them with their contexts. As Guskey (1986) has posited, when teachers are intrigued enough to try the new practices and leave the program with a ‘Well, let’s see’ attitude, the change process starts (p.9).

Re-situating research in light of my current practices: I can do more!

As new understanding about research emerged, the teachers took a further step to reassess the research value of their daily work and direct their research efforts in light of what they are currently doing in the classrooms. It is in this process that the teachers recognized that research should be hinged on teaching, and the everyday practices that they have taken for granted deserve more thorough investigations.

All the focal teachers reported their experiences of reflecting on alternative research perspectives based on their current practices. Some were poised to “aim higher” to “plan the teaching from the perspective of research.” In the following excerpt, Frances discussed how she decided to design a research study based on what she was doing in her classrooms:

“I tried extra-curriculum reading at the beginning of this semester. My students have read three books by now. They did all kinds of after-reading activities, for example, rewriting the ending of a story, or changing the story into a stage play. Now I think I can do more!. Maybe I can design a questionnaire to examine my students’ attitudes, and collect their assignments to evaluate the effect of such reading on writing. There are many research perspectives.” (Frances, interview)

In the following excerpt, Lucy discussed how the reading of an article about “classroom silence” triggered her interest to further probe the issue in her classrooms:

“I have experienced “classroom silence” many times. The article analyzed some of the reasons for silence but didn’t talk about how to solve the problem. I think I will review my cases of classroom silence and think about the solutions. Maybe I can write an article on this.” (Lucy, interview)

Kirkwood and Christie (2006) argued that teacher research in the context of CPD should seek to capture the uniqueness of the practitioners’ particular professional context. The current CPD program provided a chance for the teachers to critically analyze their work with the available research tools and derive research ideas from their classrooms. As a result, the teachers became more “research-minded.”

The current findings support the recent trends in teacher education toward practice- and inquiry-oriented collaborative models of CPD and prove the positive role of the CPD program in (re)shaping teachers’ conceptions of research. However, it would be too soon to conclude that the CPD program has led or will lead to changes in teachers’ conceptions. What can be certain is that the learning opportunities created by the program activities, in particular, discussions and the reading papers guided by the teacher educator, constitute an important facet of “the change environment” that affords teacher growth

(Clarke and Hollingsworth, 2002, p.954). The catalysts that act upon the environment and accelerate the transformation are teacher reflection and their willingness to learn.

On the one hand, teacher reflection mediates the teachers’ change process. The conversations with peers, the reading of professional publications, and the input from the teacher educator provided the teachers with a new stimulus to articulate, negotiate, and develop a new understanding of teaching and research. As Tillema and van der Westhuizen (2006) have pointed out, teacher collaboration, together with reflection, is a central feature of collaborative knowledge construction.

On the other hand, teachers’ willingness to learn, that is, a psychological state that “involves a desire to learn, experiment, and see or do something that has not been seen or done before” (Van Eekelen et al., 2006, p. 411), triggers teacher reflection and sets the wheels of teacher change in motion. Researchers have proved that teachers with strong wills to learn would be more likely to challenge their existing knowledge, think up and try out new actions to explore the newly observed phenomena, test the tentative understanding, or affirm the moves that they have invented to change things for the better (Oosterheert and Vermunt, 2001; Van Eekelen et al., 2006). Hayes (2019) pointed out that CPD is likely to succeed when teachers are willing to engage in CPD activities that are mandated and actively seek out opportunities to improve their practice in personally meaningful ways. The three ways of knowledge (re)construction identified in this section are indicative of the teachers’ self-regulated reflective thinking, which is driven by a strong will to learn and eventually leads to optimal learning results.

Conclusion

The present study explored both the outcome and process of teacher learning in a CPD program locally developed as part of the routine EFL teacher development initiatives. The findings contribute to the body of work that has called for more coherent programs of teacher learning in design and implementation and demonstrate how opportunities for teacher learning can be built into CPD programs to achieve beneficial effects. The present research offers two implications for continuing professional development.

First, CPD programs that characterize features such as peer interaction and literature reading call for deep involvement of teacher educators. The guidance and feedback offered by a teacher educator or mentor could facilitate teachers’ information processing and support their collaborative knowledge construction. Second, teachers should be motivated to learn by program contents that are closely related to their professional lives. Once teachers get interested in the new ideas, they will assume an active role in the program activities and reap their benefits.

Despite the insights offered, the current study has limitations. The issue of sustainability is crucial to any CPD program if it targets teachers' long-term development. Therefore, more longitudinal studies are needed to follow up with the teachers in their actual research engagement to capture the trajectories of learning after they have exited the program and offer more insights into how they utilize what they have obtained from the program to inform their teaching and research. Moreover, the "idiosyncrasies" of participant selection cannot be neglected when the results are to be applied to other contexts since all the focal cases are outstanding teachers with extensive research experiences, high reflexivity, and a strong desire to learn, which enabled them to fully utilize the learning opportunities afforded by the CPD program. Future CPD programs might consider revising the current program framework to satisfy the needs of target teachers. It is also unfortunate that the study was not able to involve male teachers in the focal cases to provide a more diversified perspective of teacher learning and teacher change.

Data availability statement

The raw data supporting the conclusions of this article will be available upon request.

Ethics statement

This study was reviewed and approved by Beijing Foreign Studies University. The participants provided their written informed consent to participate in this study.

References

- Alhassan, A., and Ali, H. I. H. (2020). EFL teacher research engagement: Towards a research-pedagogy nexus. *Cogent Arts Humanit.* 7:1840732. doi: 10.1080/23311983.2020.1840732
- Allison, D., and Carey, J. (2007). What do university language teachers say about language teaching research? *TESL Canada J.* 24, 61–81. doi: 10.18806/tesl.v24i2.139
- Avidov-Ungar, O., and Herscu, O. (2020). Formal professional development as perceived by teachers in different professional life periods. *Prof. Dev. Educ.* 46, 833–844. doi: 10.1080/19415257.2019.1647271
- Bakkenes, I., Vermunt, J. D., and Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learn. Instruc.* 20, 533–548. doi: 10.1016/j.learninstruc.2009.09.001
- Banegas, D., Pavese, A., Velázquez, A., and Vélez, S. M. (2013). Teacher professional development through collaborative action research: Impact on foreign English-language teaching and learning. *Educ. Action Res.* 21, 185–201. doi: 10.1080/09650792.2013.789717
- Barkuizen, G. (2009). Topics, aims and constraints in English teacher research: A Chinese case study. *TESOL Q.* 43, 113–125. doi: 10.1002/j.1545-7249.2009.tb00231.x
- Belcher, D. (2007). A bridge too far? *TESOL Q.* 41, 396–399. doi: 10.1002/j.1545-7249.2007.tb00066.x
- Bian, Y. (2018). The prevalence and the increasing impact of Guanxi. *China Q.* 235, 597–621. doi: 10.1017/S0305741018000541
- Borg, S. (2009). English language teachers' conceptions of research. *Appl. Linguist.* 30, 358–388. doi: 10.1093/applin/amp007
- Borg, S. (2010). Language teacher research engagement. *Lang. Teach.* 43, 391–429. doi: 10.1017/S0261444810000170
- Borg, S. (2011). The impact of in-service teacher education on language teachers' beliefs. *System* 39, 370–380. doi: 10.1016/j.system.2011.07.009
- Borg, S. (2013). *Teacher Research in Language Teaching: A Critical Analysis*. Cambridge, MA: Cambridge University Press.
- Borg, S. (2014). The benefits of attending ELT conferences. *ELT J.* 69, 35–46. doi: 10.1093/elt/ccu045
- Borg, S. (2015). "Beyond the workshop: CPD for English language teachers," in *Professional Development for English Language Teachers: Perspectives from Higher Education in Turkey*, ed. S. Borg (Ankara: British Council).
- Borg, S. (2018). Evaluating the impact of professional development. *RELC J.* 49, 195–216. doi: 10.1177/0033688218784371

Author contributions

YK did the initial data analysis and drafted the literature review and findings. LY created the context for data collection and finalized the manuscript. Both authors contributed to the article and approved the submitted version.

Funding

This study was part of the project of Research on Educating Excellent Foreign Language Teachers in the New Era supported by China National Social Science Fund (#19BYY220).

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.

- Borg, S., and Liu, Y. (2013). Chinese college English teachers' research engagement. *TESOL Q.* 47, 270–299. doi: 10.1002/tesq.56
- Borko, H., Jacobs, J., and Koellner, K. (2010). "Contemporary approaches to teacher professional development," in *International Encyclopedia of Education*, Vol. 7, eds P. Peterson, E. Baker, and B. McGaw (Oxford: Elsevier), 548–556.
- Cabaroglu, N., and Roberts, J. (2000). Development in student teachers' pre-existing beliefs during a 1-Year PGCE programme. *System* 28, 387–402. doi: 10.1016/S0346-251X(00)00019-1
- Chaaban, Y. (2016). Examining changes in beliefs and practices: English language teachers' participation in the school-based support program. *Prof. Dev. Educ.* 43, 592–611. doi: 10.1080/19415257.2016.1233508
- Choi, T. H. (2015). The impact of the 'Teaching English through English' policy on teachers and teaching in South Korea. *Curr. Issues Lang. Plan.* 16, 201–220. doi: 10.1080/14664208.2015.970727
- Choi, T. H. (2017). Teacher change from a short-term certification programme and its comparison with that of a comparable long-term programme: Cases from South Korea and the USA. *Int. J. Comp. Educ. Dev.* 19, 207–218. doi: 10.1108/IJCED-06-2017-0009
- Clarke, D., and Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18, 947–967. doi: 10.1016/S0742-051X(02)00053-7
- Coburn, J., and Borg, S. (2022). The impact of an in-service course on primary school English teachers: Case studies of change. *Eur. J. Teach. Educ.* 1–17. doi: 10.1080/02619768.2022.2096002
- Cochran-Smith, M., and Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in community. *Rev. Res. Educ.* 24, 249–305. doi: 10.3102/0091732X024001249
- Cornelissen, F., McLellan, R. W., and Schofield, J. (2017). Fostering research engagement in partnership schools: Networking and value creation. *Oxford Rev. Educ.* 43, 695–717. doi: 10.1080/03054985.2017.1316251
- Dimmock, C. (2016). Conceptualising the research-practice-professional development nexus: Mobilising schools as 'research-engaged' professional learning communities. *Prof. Dev. Educ.* 42, 36–53. doi: 10.1080/19415257.2014.963884
- Everton, T., Galton, M., and Pell, T. (2000). Teachers' perspectives on educational research: Knowledge and context. *J. Educ. Teach.* 26, 167–182. doi: 10.1080/02607470050127081
- Everton, T., Galton, M., and Pell, T. (2002). Educational research and the teacher. *Res. Papers Educ.* 17, 373–401. doi: 10.1080/0267152022000031388
- Farsani, M., and Babii, E. (2019). EFL faculty members' research engagement: Voices from Iran. *Res. Post Compulsory Educ.* 24, 37–59. doi: 10.1080/13596748.2019.1584437
- Gao, X., and Chow, A. (2012). Primary school English teachers' research engagement. *ELT J.* 66, 224–232. doi: 10.1093/elt/ccr046
- Gao, X., Barkuizen, G., and Chow, A. (2011). 'Nowadays, teachers are relatively obedient': Understanding primary school English teachers' conceptions of and drives for research in China. *Lang. Teach. Res.* 15, 61–81. doi: 10.1177/1362168810383344
- Garet, M., Porter, A., Desimone, L., Birman, B., and Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* 38, 915–945. doi: 10.3102/00028312038004915
- Gironzetti, E., and Muñoz-Basols, J. (2022). Research engagement and research culture in Spanish language teaching (SLT): Empowering the profession. *Appl. Linguist.* 1–28. doi: 10.1093/applin/amac016
- Gorter, D., and Arocena, E. (2020). Teachers' beliefs about multilingualism in a course on translanguaging. *System* 92:102272. doi: 10.1016/j.system.2020.102272
- Guskey, T. (1986). Staff development and the process of teacher change. *Educ. Res.* 15, 5–12. doi: 10.3102/0013189X015005005
- Hargreaves, D. (2001). "Revitalising educational research: past lessons and future prospects," in *Taking Education Really Seriously: Four Years' Hard Labor*, ed. M. Fielding (London: Routledge Falmer).
- Hayes, D. (2019). "Continuing professional development/continuous professional learning for English language teachers," in *The Routledge Handbook of English Language Teacher Education*, eds S. Walsh and S. Mann (New York, NY: Routledge), 155–168.
- He, Y., and Bagwell, D. (2022). Supporting teachers working with English learners: Engagement and impact of a professional development program. *TESOL J.* 13:e632. doi: 10.1002/tesj.632
- Hirvela, A. R. (2016). *Connecting Reading and Writing in Second Language Writing Instruction*, 2nd Edn. Ann Arbor, MI: University of Michigan Press.
- Ingvanson, L., Meiers, M., and Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Educ. Policy Anal. Arch.* 13, 1–28. doi: 10.1186/s12913-016-1423-5
- Johnson, K. E. (2009). *Second Language Teacher Education: A Sociocultural Perspective*. New York, NY: Routledge.
- Kalinowski, E., Gronostaj, A., and Vock, M. (2019). Effective professional development for teachers to foster students' academic language proficiency across the curriculum: A systematic review. *AERA Open* 5, 1–23. doi: 10.1177/2332858419828691
- Karimi, M. N., and Zade, S. S. H. (2018). Teachers' use of motivational strategies: Effects of a motivation-oriented professional development course. *Innov. Lang. Learn. Teach.* 13, 194–204. doi: 10.1080/17501229.2017.1422255
- Kelchtermans, G. (2004). "CPD for professional renewal: moving beyond knowledge for practice," in *International Handbook on the Continuing Professional Development of Teachers*, eds C. Day and J. Sachs (Maidenhead: Open University Press), 217–237.
- Kennedy, M. M. (2008). Contributions of qualitative research to research on teacher qualifications. *TESOL Q.* 30, 344–367. doi: 10.3102/0162373708326031
- Kiely, R., and Davis, M. (2010). From transmission to transformation: Teacher learning in English for speakers of other languages. *Lang. Teach. Res.* 14, 277–295. doi: 10.1177/1362168810365241
- Kirkwood, M., and Christie, D. (2006). The role of teacher research in continuing professional development. *Br. J. Educ. Stud.* 54, 429–448. doi: 10.1111/j.1467-8527.2006.00355.x
- Kubanyiova, M. (2006). Developing a motivational teaching practice in EFL teachers in Slovakia: Challenges of promoting teacher change in EFL contexts. *TESL EJ* 10, 1–17.
- Labone, E., and Long, J. (2016). Features of effective professional learning: A case study of the implementation of a system-based professional learning model. *Prof. Dev. Educ.* 42, 54–77. doi: 10.1080/19415257.2014.948689
- Lamb, M. (1995). The consequences of INSET. *ELT J.* 49, 72–80. doi: 10.1093/elt/49.1.72
- Lamie, J. M. (2004). Presenting a model of change. *Lang. Teach. Res.* 8, 115–142. doi: 10.1191/1362168804lr137oa
- Li, Y., and Zhang, L. J. (2022). Influence of mentorship and the working environment on English as a foreign language teachers' research productivity: The mediation role of research motivation and self-efficacy. *Front. Psychol.* 13:906932. doi: 10.3389/fpsyg.2022.906932
- Luneberg, M., and Korthagen, F. (2009). Experience, theory and practical wisdom in teaching and teacher education. *Teach. Teach. Theory Pract.* 15, 225–240. doi: 10.1080/13540600902875316
- Macalister, J. (2018). Professional development and the place of journals in ELT. *RELC J.* 49, 238–256. doi: 10.1177/0033688218771385
- Mann, S. (2005). The language teacher's development. *Lang. Teach.* 38, 103–118. doi: 10.1017/S0261444805002867
- Merriam, S. (2001). *Qualitative Research and Case Study Applications in Education*. San Francisco, CA: Jossey-Bass.
- Miles, M. B., and Huberman, A. M. (1994). *Qualitative Data Analysis*. Beverly Hills, CA: Sage.
- Ming, L. (2019). *Understanding the Impact of INSET on Teacher Change in China*. Singapore: Palgrave Pivot.
- Oosterheert, I. E., and Vermunt, J. D. (2001). Individual differences in learning to teach: Relating cognition, regulation and affect. *Learn. Instruc.* 11, 133–156. doi: 10.1016/S0959-4752(00)00019-0
- Orland-Barak, L. (2006). Convergent, divergent and parallel dialogues: Knowledge construction in professional conversations. *Teach. Teach. Theory Pract.* 12, 13–31. doi: 10.1080/13450600500364547
- Ottley, J. R., Piasta, S. B., Mauck, S. A., O'Connell, A., Weber-Mayr, M., and Justice, L. M. (2015). The nature and extent of change in early childhood educators' language and literacy knowledge and beliefs. *Teach. Teach. Educ.* 52, 47–55. doi: 10.1016/j.tate.2015.08.005
- Paran, A. (2017). 'Only connect': Researchers and teachers in dialogue. *ELT J.* 71, 499–508. doi: 10.1093/elt/ccx033
- Park, M., and So, K. (2014). Opportunities and challenges for teacher professional development: A case of collaborative learning community in South Korea. *Int. Educ. Stud.* 7, 96–108. doi: 10.5539/ies.v7n7p96

- Rahimi, M., and Weisi, H. (2018). The impact of research practice on professional teaching practice: Exploring EFL teachers' perception. *Cogent Educ.* 5:1480340. doi: 10.1080/2331186X.2018.1480340
- Sahin, I., and Yildirim, A. (2015). Transforming professional learning into practice. *ELT J.* 70, 241–252. doi: 10.1093/elt/ccv070
- Sato, M., and Loewen, S. (2019). Do teachers care about research? The research-pedagogy dialogue. *ELT J.* 73, 1–10. doi: 10.1093/elt/ccy048
- Sedova, K. (2017). Transforming teacher behaviour to increase student participation in classroom discourse. *Teach. Dev.* 21, 225–242. doi: 10.1080/13664530.2016.1224775
- Selvi, A. F., and Martin-Beltrán, M. (2016). Teacher-learners' engagement in the reconceptualization of second language acquisition knowledge through inquiry. *System* 63, 28–39. doi: 10.1016/j.system.2016.08.006
- Shkedi, A. (1998). Teachers' attitudes towards research: A challenge for qualitative researchers. *Int. J. Qual. Stud. Educ.* 11, 559–577. doi: 10.1080/095183998236467
- Sim, J. Y. (2011). *The Impact of in-Service Teacher Training: A Case Study of Teachers' Classroom Practice and Perception Change*. Ph. D Thesis. Coventry: University of Warwick.
- So, K. (2013). Knowledge construction among teachers within a community based on inquiry as stance. *Teach. Teach. Educ.* 29, 188–196. doi: 10.1016/j.tate.2012.10.005
- Stenhouse, L. (1975). *An Introduction to Curriculum Research and Development*. London: Heinemann.
- Tabatabaee-Yazdi, M., Motallebzadeh, K., Ashraf, H., and Baghaei, P. (2018). Continuing professional development strategies: A model for the Iranian EFL teachers' success. *Sage Open* 8, 1–14. doi: 10.1177/2158244018764234
- Teng, L. S. (2016). Changes in teachers' beliefs after a professional development project for teaching writing: Two Chinese cases. *J. Educ. Teach.* 42, 106–109. doi: 10.1080/02607476.2015.1135228
- Tillema, H., and van der Westhuizen, G. J. (2006). Knowledge construction in collaborative enquiry among teachers. *Teach. Teach.* 12, 51–67. doi: 10.1080/13450600500365403
- Turner, F., Brownhill, S., and Wilson, E. (2017). The transfer of content knowledge in a cascade model of professional development. *Teach. Dev.* 21, 175–191. doi: 10.1080/13664530.2016.1205508
- Van Eekelen, I. M., Vermunt, J. D., and Boshuizen, H. P. A. (2006). Exploring teachers' will to learn. *Teach. Teach. Educ.* 22, 408–423. doi: 10.1016/j.tate.2005.12.001
- van Veen, K., Zwart, R., and Meirink, J. (2012). "What makes teacher professional development effective? A literature review," in *Teacher Learning that Matters: International Perspectives*, eds M. Kooy and K. van Veen (New York, NY: Routledge), 3–21.
- Vu, M. T. (2021). Between two worlds? Research engagement dilemmas of university English language teachers in Vietnam. *RELC J.* 52, 574–587. doi: 10.1177/0033688219884782
- Wang, Q., and Zhang, H. (2014). Promoting teacher autonomy through university-school collaborative action research. *Lang. Teach. Res.* 18, 222–241. doi: 10.1177/1362168813505942
- Wermke, W. (2011). Continuing professional development in context teachers continuing professional development culture in Germany and Sweden. *Prof. Dev. Educ.* 37, 665–683. doi: 10.1080/19415257.2010.533573
- Yuan, R., and Lee, I. (2015). Action research facilitated by university-school collaboration. *ELT J.* 69, 1–10. doi: 10.1093/elt/ccu031
- Zimmerman, W. A., Knight, S. L., Favre, D. E., and Ikhlef, A. (2017). Effect of professional development on teaching behaviors and efficacy in Qatari educational reforms. *Teach. Dev.* 21, 324–345. doi: 10.1080/13664530.2016.1243571

Appendix

Table 1 Summary of changes in teachers' conceptions of research.

Name	Sub-categories of conceptions of research	Before the program	After the program
Lucy	Nature of research	Research should be scientific.	Research does not need to be very theoretical. Research should be grounded in teaching.
	Purpose of research	I wrote papers to record the new practices in my classroom.	We can study both new practices and classroom routines to improve teaching.
	Relationship between teaching and research	Research helps record my experiments in teaching.	Research is about every aspect of teaching. Research and teaching are interwoven.
	Process of research	The papers I have written were not very professional. I look for materials when I'm to write a research paper.	My old way should be reversed. Research starts from a purposeful collection of materials
Jane	Nature of research	Not sure what is research. What I have done before was probably not real research. Research topics should be grand.	What I have done IS research. It helps find feasible ways to teach and solve practical problems. Research topics can be small.
	Purpose of research	We wrote papers to meet all kinds of requirements.	We write to share new practices with others and find solutions from reading research.
	Relationship between teaching and research	Research is for recording and reflecting on teaching.	Research and teaching are nurturing each other.
	Process of research	Research should be systematic. My research is only a summary of teaching.	We can choose a topic, select a method, and collect data in the process of teaching.
Wanda	Nature of research	Good research needs (statistical) data. Research should produce generalizable results.	To do research is to study our practices. There are many research methods.
	Purpose of research	As a teacher, I am supposed to do research. Doing research will improve teaching.	To do research is to solve our problems in teaching.
	Relationship between teaching and research	Research means to apply theories to practice.	Theories also come from teaching. Research helps generalize theories from teaching.
	Process of research	We have a lot of data but do not know how to develop them into real research.	Research is a process of discovering problems, solving problems, and rediscovering new problems.
Frances	Nature of research	Research needs theories and data. We use statistics to analyze data.	Research means the studies about teaching with various methods and tools available.
	Purpose of research	To do research is to share my experiences with others.	To do research is to record, reflect, and refine teaching.
	Relationship between teaching and research	Research can improve teaching. I have to strike a balance between teaching and research.	Teaching and research are closely related and inseparable.
	Process of research	Research should be carried out in a step-by-step manner.	Research requires systematic planning and data collection. We plan the research before we teach.
Sunny	Nature of research	To do research is to have experiments. Research is data-driven.	Everything about teaching can be researched. Research takes many forms. It can be narrative.
	Purpose of research	The school wants us to do research. It is part of our job. We do experiments in classrooms to improve teaching.	To research is to publish and to share. Research is personal, not institutional.
	Relationship between teaching and research	Research helps improve teaching.	Research is grounded in teaching and teaching enriches research.
	Process of research	Research needs systematic planning.	Research requires rational thinking and systematic organization. Research is to implement the plan step by step.



OPEN ACCESS

EDITED BY

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

REVIEWED BY

Evan Andrew Wilhelms,
College of Wooster, United States
Soraj Hongladarom,
Chulalongkorn University, Thailand

*CORRESPONDENCE

Weijane Lin
vjlin@ntu.edu.tw

SPECIALTY SECTION

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

RECEIVED 27 May 2022

ACCEPTED 17 August 2022

PUBLISHED 27 September 2022

CITATION

Yueh H-P, Huang C-Y and Lin W (2022)
Examining the differences between
information professional groups
in perceiving information ethics: An
analytic hierarchy process study.
Front. Psychol. 13:954827.
doi: 10.3389/fpsyg.2022.954827

COPYRIGHT

© 2022 Yueh, Huang 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.

Examining the differences between information professional groups in perceiving information ethics: An analytic hierarchy process study

Hsiu-Ping Yueh^{1,2}, Ching-Yin Huang³ and Weijane Lin^{3*}

¹Department of Psychology, National Taiwan University, Taipei City, Taiwan, ²Department of Bio-Industry Communication and Development, National Taiwan University, Taipei City, Taiwan,

³Department of Library and Information Science, National Taiwan University, Taipei City, Taiwan

Information and communication technology (ICT) has a great impact on contemporary society and people's lives. Especially with the pervasive access to rapidly developing technology, the impact of ICT on society and human values, the norms of ICT use, and the ethical issues derived from them are beyond the past ethical framework and deserve more research attention. The purpose of this study was to explore the key factors that influence the decision-making behaviors of information professionals when they are faced with information ethics issues. The study adopted the analytic hierarchical process method to develop the evaluation framework and criteria for information professional ethics and employed the professional fields of library and information science and information technology as examples to compare whether information professionals in different fields make different judgments on the aforementioned decision-making criteria. The results of the study validated the professional information ethics hierarchy and criteria and contributed to the field of information ethics research by providing information on the aspects that need attention in the cultivation of professionals in different fields.

KEYWORDS

information ethics, professional ethics, analytical hierarchy process, library information science, information technology - IT

Introduction

The rapid development of information technology has led to dramatic changes in the economies, politics, and cultures of countries around the world. The Internet now plays a crucial role in a globalized society, but while the Internet provides considerable convenience, it also generates a number of ethical issues. For example, the Internet allows easy access to various materials under copyright protection, which is largely disregarded by many users. Reports on information systems security and control show that this unethical Internet activity has resulted in significant losses of profit for publishing companies (Anthes, 1996). As Mason (1986) stressed, information ethics is critical in the information era and should be emphasized, identified, and classified as a new and specific area of knowledge to be explored and studied (Floridi and Sanders, 2001).

In the general approach to ethical dilemmas, decision makers must choose one of two different ethical values in preference to the other (MacLagan, 2003). Information ethics is concerned with the moral dilemmas and ethical conflicts that arise when human beings interact with information, information and communication technologies (ICTs), and information systems (Carbo, 2008). Critical issues, including privacy, property, accuracy, and accessibility of information, are often encountered by people during their interaction with information (Mason, 1986), and throughout their careers, students who major in information-related fields will face a wide range of ethical dilemmas related to issues such as trust (Kelton et al., 2008) and transparency (Fleischmann and Wallace, 2009). Employees' lack of information ethics could compromise institutional information security, damage an institution's reputation, or even cause financial losses. Therefore, for information majors, information ethics is not simply part of their literacy but also directly associated with their profession.

With reference to the importance of information ethics for professional development, the Association for Computing Machinery (ACM, 1992) proposed the ACM Code of Ethics and Professional Conduct, which identifies the personal responsibility and commitment of information professionals. Information ethics also gained attention from the research field of information science education. Chang (2011) utilized Kohlberg's cognitive moral development model to measure improvements in students' information ethics values through technology-mediated learning models. The results showed that e-learning materials that involved multiple media representations improved students' comprehension as well as respect toward the norms, privacy, accessibility, and intellectual property issues.

However, current resources for learning and training of information ethics remain quite limited. According to a local investigation of the undergraduate curriculum by Lin and Chou (2014), only 37% of all the 199 departments or institutes of

information science and engineering in Taiwan had courses related to information ethics. For this limited number of courses, their further review of 53 syllabi suggested that most of these courses were taught with didactic approaches in classroom teaching, which failed to convey the dynamic and situated nature of information ethics. Pedagogically, previous studies (Liu and Yang, 2012; Fiesler et al., 2020) supported that the ability to assess ethical dilemmas involved high-level thinking, and therefore pedagogies that encourage critical thinking, such as case-based discussions, video tutorials, debates, and role-playing, would be more suitable for information ethics curriculum. Additionally, in terms of the nature of the learning resources, given the many theoretical frameworks of information ethics proposed by previous research (Hunt and Vitell, 1986; Thong and Yap, 1998), the transformation of these conceptual norms into practical guidelines of actions and curriculum calls for further efforts to clearly structure the elements of information ethics. Furthermore, information professional organizations in different fields follow different codes of ethics due to different backgrounds, field knowledge, and even the content of their pre-service training. For example, the professional training objectives of the library and information science, computer science, and engineering fields are aligned with those of information technology, but the focus of education for library and information science personnel is on the professional knowledge and information application skills required to provide library and information services (ASIS&T, 1992; IFLA, 2012; ALA, 2021), with an emphasis on people and services. Those in the information technology field are more oriented toward learning how to apply various information technology techniques (ACM, 2018; IEEE, 2020), and they thus focus on the data and technology itself. In addition to the differences in the professional training received in school, the actual practices and professional ethics of the workplace are also different in the field of library and information science and the field of information technology. Therefore, the scope of information professionals should be expanded to include other professionals in the field of library and information science and information technology, teachers and students, in addition to the previous focus on management information system engineers or professionals, in order to truly meet the contemporary social context.

Motivated by the aforementioned issues, the purpose of this study was to explore the key factors that influence the decision-making behaviors of information professionals when they are faced with information ethics issues. In addition, this study employed the professional field of library and information science and the field of information technology as examples to compare whether information professionals in different fields make different judgments on the aforementioned decision-making criteria, and then it explored the possible influencing factors and the planning of information ethics education for professionals. Past empirical studies of ethics were often framed

around dilemmas, mostly in the form of situational questions for which examinees would have to make the choice they thought more appropriate (often one out of two). However, the purpose of this study was not to directly assess the subjects' ethical decision making choices regarding the stated situational issues by means of situational questions, but rather to collate the guidelines suggested by professional organizations and then to allow professionals to assess the relative importance of these guidelines.

Therefore, this study focused on decision analysis as a theoretical and methodological approach which can be mainly applied to decision making problems under uncertainty and with several evaluation criteria, with reference to the Analytic Hierarchy Process (AHP) method developed to systematize complex problems, decompose them at different levels, and evaluate them in a comprehensive manner by quantifying them and finding their context (Saaty, 1990). This study first compiled the professional ethical standards developed by the major information professional societies and summarized the important elements of information ethics. The results of the study are intended to validate the professional information ethics standards compiled in the preceding section, and then to provide a reference contribution to the field of information ethics research by providing information on the aspects that need attention in the cultivation of professionals in different fields.

Information ethics for information professionals

Ethics are a set of principles and concepts that judge whether a behavior is right or not. Compared to morals, ethics focus on the complex relationships and interactions between people at the social level. Various forms of ethical theories have been discussed and developed for people to explore how to interact properly with others (Arnold and Bowie, 2019). Among them, four types of theoretical approaches, including consequence-based theories, duty-based theories, rights-based theories, and virtue-based theories, have been regarded as critical and fundamental in information ethics education (Fallis, 2007). These four approaches complement one another by their alternative emphases on different aspects of information ethics, and mastering them facilitates learners' further understanding of the importance of the ethical codes or principles. Furthermore, to adopt these approaches in the ethical decision-making process, Thong and Yap (1998) in their empirical investigation of 243 entry-level information system professionals suggested consequence-based approaches that reveal the consequences of ethical/unethical behaviors, which could facilitate learners' comprehension of the codes and were more likely to lead to ethical decisions.

As information ethics involves field practices and interaction with others for common well-being, practical concerns have placed great emphasis on issues related to information ethics in the contexts of different professionals, types of information, and rapidly changing technologies. Since the 1980s, several professional societies of informaticists, such as the Association for Computing Machinery (Gotterbarn et al., 2018), Institute of Electrical and Electronics Engineers (IEEE, 2020), American Library Association (Witt, 2017; ALA, 2021), and International Federation of Library Associations (IFLA, 2012; Trepanier et al., 2019) have developed their own codes of ethics to guide practices of handling information ethics issues. These codes of ethics demonstrate the discipline-specific values and responsibilities of information professionals and institutions to society, as library-related associations accentuated ethical issues in providing information services (LAROC, 2002; IFLA, 2012; ALA, 2021) while computer societies highlighted concerns about producing information (ASIS&T, 1992; ACM, 2018; IEEE, 2020). However, as shown in Table 1, they also shared a common focus on issues raised by Mason (1986), whose systematic arguments about the four ethical issues of privacy, accuracy, property, and accessibility have attracted informaticists' attention.

Our review in Table 1 shows that Mason's four issues have been included in the professional ethics in both fields, suggesting a common emphasis on privacy, accuracy, property, and accessibility across information professionals' practices. Nevertheless, due to the highly contextual nature of information ethics (Nissenbaum, 2019), even when the code of ethics is fully understood, the actual decisions and practices in the field could be much more complex and dilemmatic (Fallis, 2007; McMenemy et al., 2014). Advancing from the above analysis, Table 2 provides a further comparison and discussion of the three main criteria included in the two domains using Mason's (1986) PAPA framework: general moral imperatives, specific professional responsibilities, and organizational/workplace responsibilities. These were categorized by considering the professional organizations' information management norms and personnel roles in relation to their cognitive and behavioral performance. The four aforementioned issues are included in the following behavioral levels, which are covered in the content of the evaluation items.

Based on the above analysis, the content of ethics for information professionals includes moral responsibility, professional ethics, and workplace ethics, while the content of the professional code of conduct in the field of library information implies the concept of a limited field in the library and is therefore closely integrated with the mission and values of the organization. The professional codes of conduct in the field of information technology tend to provide sets of principle-based guidelines. Among them, ACM's Code of Ethics, released in 1992 (ACM, 2018), covers a wide range of topics and is regarded as a solid basis for considering informaticists'

TABLE 1 PAPA-related statements in professional codes of informaticists.

Library and information science		Computer science	
Institute	Related articles	Institute	Related articles
Privacy			
ALA	3. We protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted.	ACM	1.6 Respect privacy 1.7 Honor confidentiality
IFLA	3. Privacy, secrecy and transparency: Librarians and other information workers respect personal privacy, and the protection of personal data, necessarily shared between individuals and institutions. . .	IEEE	1. . . to strive to comply with ethical design and sustainable development practices, to protect the privacy of others, and to disclose promptly factors that might endanger the public or the environment.
LAROC	10. Librarians shall strictly observe business secrets, protect the privacy of readers, and not seek to benefit themselves or harm others.	ASIS&T	- To resist all forms of censorship, inappropriate selection and acquisitions policies, and biases in information selection, provision and dissemination
Accuracy			
ALA	1. We provide the highest level of service to all library users through appropriate and usefully organized resources; equitable service policies; equitable access; and accurate, unbiased, and courteous responses to all requests	ACM	1.3 Be honest and trustworthy. 2.9 Design and implement systems that are robustly and useably secure.
IFLA	5. Neutrality, personal integrity and professional skills: Librarians and other information workers are strictly committed to neutrality and an unbiased stance regarding collection, access and service. . .	IEEE	5. to seek, accept, and offer honest criticism of technical work, to acknowledge and correct errors, to be honest and realistic in stating claims or estimates based on available data. . .
LAROC	3. Based on the principle of neutrality, librarians should collect all kinds of information and protect the readers' rights.	ASIS&T	- not knowingly making false statements or providing erroneous or misleading information - to improve the information systems. . . to the best of their means and abilities by providing the most reliable and accurate information and acknowledging the credibility of the sources as known or unknown
Property			
ALA	4. We respect intellectual property rights and advocate balance between the interests of information users and rights holders.	ACM	1.5 Respect the work required to produce new ideas, inventions, creative works, and computing artifacts.
IFLA	4. Open access and intellectual property: . . . Librarians and other information workers have a professional duty to advocate for exceptions and limitations to copyright restrictions for libraries. . .	IEEE	4. to avoid unlawful conduct in professional activities. 5. . . And to credit properly the contributions of others
LAROC	4. Librarians should strive to preserve various kinds of library information and promote cultural exchange	ASIS&T	- not using their position beyond their authorized limits or not using their credentials to misrepresent themselves
Accessibility			
ALA	1. We provide the highest level of service to all library users through appropriate and usefully organized resources; equitable service policies; equitable access; and accurate, unbiased, and courteous responses to all requests. 7. We distinguish between our personal convictions and professional duties and do not allow our personal beliefs to interfere with fair representation of the aims of our institutions or the provision of access to their information resources.	ACM	1.4 Be fair and take action not to discriminate. 2.7 Foster public awareness and understanding of computing, related technologies, and their consequences.
IFLA	2. Responsibilities toward individuals and society: . . . librarians and other information workers ensure that the right of accessing information is not denied and that equitable services are provided for everyone. . . 4. Open access and intellectual property: Librarians and other information workers' interest is to provide the best possible access for library users to information and ideas in any media or format. . .	IEEE	2. to improve the understanding by individuals and society of the capabilities and societal implications of conventional and emerging technologies, including intelligent systems. 7. to treat all persons fairly and with respect, and to not engage in discrimination based on characteristics such as race, religion, gender, disability, age, national origin. . .
LAROC	2. Librarians should provide services based on the principle of equal access and not discriminate. 6. Librarians should continue to improve readers' access and ability to use library resources.	ASIS&T	- adhering to principles of due process and equality of opportunity.

practices. Therefore, it was used as the main framework of the hierarchical analysis and the development of assessment criteria in this study in an attempt to understand the importance

of the different target constructs and assessment criteria for information professionals in the fields of both library and information science and information technology.

Methodology

The analytical hierarchy process method

This study adopted the Analytical Hierarchy Process (AHP), which is a Multiple Attribute Decision Making (MADM) method (Saaty, 1987). The behaviors of individuals and groups in making decisions are complex and often include many different dimensions of consideration (Lin and Huang, 2013). Because of the different characteristics of each decision-maker, the importance of each aspect in the decision-making process could also vary. Based on a review of the professional and information ethics literature in general, and specifically the ACM Codes of Ethics, this study first developed an analytical hierarchy of information ethics.

The concept of information ethics was composed of three major elements: general moral imperatives, specific professional responsibilities, and organizational/workplace responsibilities. With the three major elements as the main criteria for decision-making, sub-criteria were further developed as specific combinations of the characteristics of each main criterion. After a pilot study with 5 experts from academia and practice sectors of information science, several constructs under each major element were merged or deleted according to the rules of the AHP method by the value of consistency index and consistency

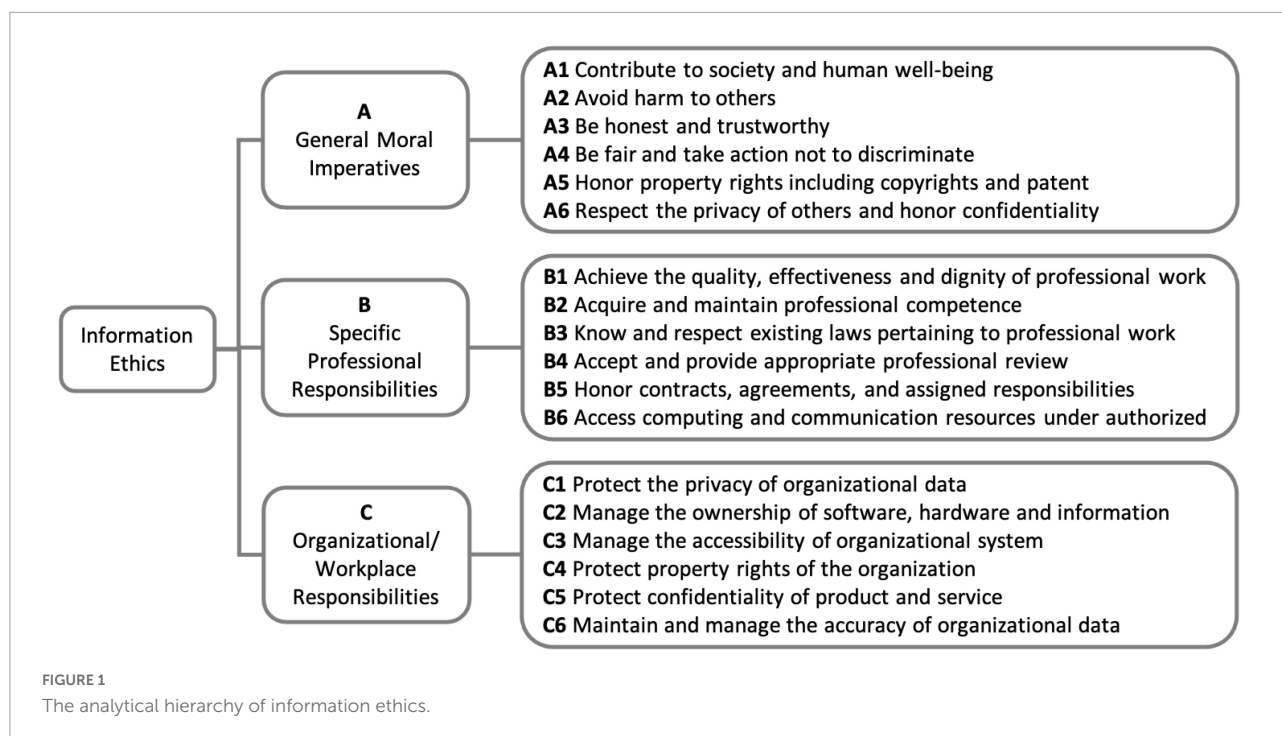
ratio (Saaty, 1987). The final analytical hierarchy of information ethics is shown as Figure 1.

Data collection and analysis

Based on the structure of the decision criteria, the final AHP questionnaire was compiled as a scale by Saaty (1987). This scale consists of all the criteria in Figure 1 and contains pairwise comparisons of criteria at each hierarchical level. The ranking system is on a scale from 1 to 9, where 1 means that two criteria are of equal importance and 9 indicates the highest importance of one criterion over the other. With the instrument, data were collected from experts with experience as practitioners and researchers in information science. They were drawn from departments or divisions of information science, information management, and information engineering in universities and in the public and private sectors that provide training, products, or services. These experts were invited to examine and weigh the critical factors regarding information ethics in the context of their practices with this AHP questionnaire. In this study, the data were analyzed in the Expert Choice 11.5 software package. After the model of the ethical hierarchy of information professionals was established, the judgment results of each expert's answers were entered into a relative comparison matrix, after which

TABLE 2 Comparison of the professional ethics criteria.

Library and information science	Computer science
Definition/Discussion	Definition/Discussion
<p>Main Characteristic 1: General moral imperatives</p> <p><i>The code of ethics in the field of library and information science is closely related to its purpose of promoting social, cultural and economic well-being, and is based on the Universal Declaration of Human Rights, which states that all human beings have the freedom of opinion, expression and access to information (IFLA, 2012).</i></p>	<p><i>The Code of Information Ethics, developed by professional associations in the field of information technology, takes into account the impact of information technology on society and aims to promote human and social well-being and avoid harming others. When the ACM updated its code of ethics in 2018, it further emphasized the importance of the public nature of IT, that all people are stakeholders in IT, that transparency should be promoted and that stable and trustworthy systems should be established, and that it is prudent to modify or discontinue infrastructure and functions that the public relies on (ACM, 1992, 2018).</i></p>
<p>Main Characteristic 2: Specific professional responsibilities</p> <p><i>Professional ethical judgments regarding the development of professional work and competencies, in the fields of both library and information science and information technology, are referred to in order to maintain the quality, effectiveness, and dignity of professional work, as well as to further and maintain one's professional competencies, including the study of relevant legal issues.</i></p>	<p><i>The ACM Code of Conduct ACM (1992, 2018) more comprehensively states that one should understand the laws related to professional work, accept and provide appropriate professional commentary, respect contractual obligations and agreed liabilities, and be authorized to access computer and communication resources. In 2018, the association added details on how one should measure one's professional competence.</i></p>
<p>Main Characteristic 3: Organizational/workplace responsibilities</p> <p><i>The Code of Ethics of the two professional associations covers fewer levels of professional ethical judgment on organizational responsibility and only includes the need to maintain the confidentiality of the organization's mission.</i></p>	<p><i>In addition to the aforementioned role in maintaining the confidentiality of an organization's mission, the field of information technology encompasses the use of computers and communication software that should be used under authorization. However, because of the importance of professional ethics for organizations, how to manage information is also gaining attention in information management issues (Buchanan et al., 2009).</i></p>
<p><i>Depending on the format of the information, the career stage, and the four issues of PAPA, it can be expanded as follows: protecting the privacy of the organization, protecting the confidentiality of products and services, and managing the accuracy of the data.</i></p>	



the weights were calculated and the consistency of each questionnaire was checked.

hierarchical structure was appropriate for comparative analysis of subsequent levels of decision-making elements.

Results and discussions

Analytic hierarchy process consistency testing

A total of 52 valid questionnaires were collected. The respondents included 25 IT experts (10 academic experts and 15 practitioner experts) and 27 library and information science experts (7 academic experts and 20 practitioner experts). Because consistency is essential for pairwise comparison matrices to ensure that the results are meaningful and it is also required by the AHP method (Xu and Xu, 2020), in order to ensure the validity of the pairwise comparisons and achieve uniformity of responses, each questionnaire was checked for consistency (Saaty, 1987). Methodologically, the consistency index (CI) of the derived weights should be less than 0.1 to ensure that the set of judgments is reliable. The analysis of the overall inconsistency index (O.I.I.) suggested that 10 responses out of 25 IT experts and 7 responses out of 27 LIS experts should be removed due to insufficient consistency (O.I.I. > 0.1).

As shown in Table 3, the final number of respondents included in the AHP analysis was 35, including 15 experts in the IT field (8 academic experts and 7 practitioner experts) and 20 experts in the library and information science field (5 academic experts and 15 practitioner experts). The overall inconsistency index (O.I.I.) for both fields was less than 0.1, indicating that the

The main criteria analysis of two expert groups

The main criteria of “general moral imperatives,” “specific professional responsibilities,” and “organizational/workplace responsibilities” valued by different informaticist groups of information technology (IT) and library and information science (LIS) are listed in Tables 4, 5, respectively. The results suggested that, although both the academic and practitioner IT experts were identical in their emphasis on the three criteria, academic experts in LIS held views that were in almost direct opposition to those of their practitioner peers. The practitioner experts in IT and LIS areas ranked the three main criteria identically as follows: “general moral imperatives” was the most important, followed by “organizational/workplace responsibilities” and then “specific professional responsibilities.” The academic IT

TABLE 3 Background information of the respondents providing valid responses.

Field	Academic	Practitioner	Total
Information Technology (IT)	8	7	15
Library and Information Science (LIS)	5	15	20
Total	13	22	35

TABLE 4 Three main criteria weights and rankings by experts in the IT field.

Main criteria	C.I.		Weights		Priority	
	Academic	Practitioner	Academic	Practitioner	Academic	Practitioner
General Moral Imperatives	0.01	0.03	0.491	0.606	1	1
Specific Professional Responsibilities	0.02	0.02	0.192	0.166	3	3
Organizational/workplace Responsibilities	0.009	0.03	0.317	0.228	2	2

In the aspect, the academic experts' C.I. was 0.00057; the practitioners' C.I. was 0.00126.

TABLE 5 Three main criteria weights and ranking by experts in the LIS field.

Main criteria	C.I.		Weights		Priority	
	Academic	Practitioner	Academic	Practitioner	Academic	Practitioner
General Moral Imperatives	0.03	0.00593	0.349	0.556	2	1
Specific Professional Responsibilities	0.02	0.00614	0.245	0.193	3	3
Organizational/workplace Responsibilities	0.01	0.00581	0.406	0.251	1	2

In this aspect, the academic experts' C.I. was 0.0011; the practitioners' C.I. was 0.00008.

experts possessed the same emphasis as the practitioner, but the academic LIS experts regarded organizational/workplace responsibilities as the most important. The results implied the different impacts of the technical and contextual features of information service on decision-making. While technical characteristics were applicable across organizations, contextual factors could mostly vary from organization to organization. Therefore, academic experts in the LIS area, who often needed to consider the overall library service across institutions, could place more emphasis on individual organizational/workplace factors. This is also in line with previous studies that have found that LIS professionals have more ethical issues and considerations due to the human-centered nature of their discipline and professional conduct (Fallis, 2007; ALA, 2021; Lin et al., 2022).

In addition, all respondents considered “specific professional responsibilities” to be the least important, indicating that both academic experts and practical experts mostly considered the critical component of ethics that affected information personnel to be general moral imperatives. It is possible that, when people face ethical dilemmas, the first consideration is the intrinsic value judgment accumulated from past experiences which shaped their moral cognition.

Analysis of the general moral imperatives criteria of two expert groups

For the analysis of the second level of the hierarchy, six evaluation sub-criteria were included in the general moral imperatives criteria, and the results of the evaluation are shown in Tables 6, 7. IT academics, practitioners, and LIS academic experts all ranked “avoid harm to others” as the most important

criterion for assessing general moral imperatives. This criterion was also ranked second only to “respect the privacy of others and honor confidentiality” by the LIS practitioner experts. The least important criterion was “Contribute to society and human well-being,” which was ranked sixth by all the experts except the practitioners in the IT field, who ranked it fourth.

In addition, the results showed that the practitioners in the field of LIS gave more importance to the financial value and added value of the products or intellectual property protected by property rights, copyrights, patents, and intellectual property rights, and further rated the criteria as a higher priority, while the criterion of “Honor property rights including copyrights and patents” was ranked 5th by the other three groups of experts and not given much importance.

Analysis of the specific professional responsibilities criteria of two expert groups

The specific professional responsibilities criteria contained six evaluation sub-criteria. As shown in Tables 8, 9, “achieve quality, effectiveness, and dignity of professional work,” “honor contracts, agreements, and assigned responsibilities,” and “access computing and communication resources with authorization” were all ranked in the top three in terms of importance by both academic and practitioner experts in the IT and LIS fields. The importance ranking of academic experts and practitioner experts in the LIS field was the same, from highest to lowest, for “honor contracts, agreements, and assigned responsibilities,” “access computing and communication resources with authorization,” and “achieve quality, effectiveness, and dignity of professional work,” respectively.

TABLE 6 Weights and ranking of general moral imperatives criteria by IT experts.

Criteria	Weights		Priority	
	Academic	Practitioner	Academic	Practitioner
A1 Contribute to society and human well-being	0.069	0.134	6	4
A2 Avoid harm to others	0.319	0.297	1	1
A3 Be honest and trustworthy	0.198	0.213	2	2
A4 Be fair and take action not to discriminate	0.160	0.098	3	6
A5 Honor property rights including copyrights and patents	0.116	0.117	5	5
A6 Respect the privacy of others and honor confidentiality	0.138	0.142	4	3

In this aspect, the academic experts' C.I. was 0.03; the practitioners' C.I. was 0.03.

TABLE 7 Weights and ranking of general moral imperatives criteria by LIS experts.

Criteria	Weights		Priority	
	Academic	Practitioner	Academic	Practitioner
A1 Contribute to society and human well-being	0.095	0.080	6	6
A2 Avoid harm to others	0.233	0.193	1	2
A3 Be honest and trustworthy	0.203	0.143	2	5
A4 Be fair and take action not to discriminate	0.159	0.166	4	4
A5 Honor property rights including copyrights and patents	0.133	0.192	5	3
A6 Respect the privacy of others and honor confidentiality	0.177	0.226	3	1

In this aspect, the academic's C.I. was 0.03; the practitioners' C.I. was 0.00593.

TABLE 8 Weights and ranking of specific professional responsibilities criteria by IT experts.

Criteria	Weights		Priority	
	Academic	Practitioner	Academic	Practitioner
B1 Achieve the quality, effectiveness and dignity of professional work	0.179	0.218	3	1
B2 Acquire and maintain professional competence	0.124	0.140	5	4
B3 Know and respect existing laws pertaining to professional work	0.173	0.138	4	5
B4 Accept and provide appropriate professional review	0.084	0.114	6	6
B5 Honor contracts, agreements, and assigned responsibilities	0.203	0.214	2	2
B6 Access computing and communication resources with authorization	0.237	0.176	1	3

In this aspect, the academic experts' C.I. was 0.02; the practitioners' C.I. was 0.02.

In addition, academic experts in both domains ranked the importance of the “specific professional responsibilities” criteria in roughly the same order. Only the academic experts in the IT field considered “access computing and communication resources with authorization” to be more important than “honor contracts, agreements, and assigned responsibilities,” while the academic experts in the LIS field thought the opposite. It was inferred that the work nature and environment of the academic experts were similar, so the priority of the evaluation criteria for “specific professional responsibilities” related to the workplace would be almost the same.

Analysis of the organizational/workplace responsibilities criteria of two expert groups

The organizational/workplace responsibilities criteria contained six evaluation sub-criteria. As shown in [Tables 10, 11](#) below, “protect the privacy of organizational data” was the most important criterion for experts in both fields. The next most important criteria were more divergent and mainly included “protect confidentiality of products and services,”

TABLE 9 Weights and ranking of specific professional responsibilities criteria by LIS experts.

Criteria	Weights		Priority	
	Academic	Practitioner	Academic	Practitioner
B1 Achieve the quality, effectiveness and dignity of professional work	0.142	0.172	3	3
B2 Acquire and maintain professional competence	0.102	0.105	5	6
B3 Know and respect existing laws pertaining to professional work	0.131	0.155	4	4
B4 Accept and provide appropriate professional review	0.099	0.134	6	5
B5 Honor contracts, agreements, and assigned responsibilities	0.266	0.230	1	1
B6 Access computing and communication resources with authorization	0.259	0.205	2	2

In this aspect, the academic experts' C.I. was 0.02; the practitioners' C.I. was 0.00614.

TABLE 10 Weights and ranking of organizational/workplace responsibilities criteria by IT experts.

Criteria	Weight		Priority	
	Academic	Practitioner	Academic	Practitioner
C1 Protect the privacy of organizational data	0.250	0.344	1	1
C2 Manage the ownership of software, hardware and information	0.148	0.048	3	6
C3 Manage the accessibility of the organizational system	0.122	0.077	6	5
C4 Protect property rights of the organization	0.140	0.182	5	2
C5 Protect confidentiality of products and services	0.147	0.182	4	2
C6 Maintain and manage the accuracy of organizational data	0.193	0.166	2	4

In this aspect, the academic experts' C.I. was 0.02; the practitioners' C.I. was 0.03.

TABLE 11 Weights and ranking of organizational/workplace responsibilities criteria by LIS experts.

Criteria	Weight		Priority	
	Academic	Practitioner	Academic	Practitioner
C1 Protect the privacy of organizational data	0.275	0.212	1	1
C2 Manage the ownership of software, hardware and information	0.122	0.096	6	6
C3 Manage the accessibility of the organizational system	0.148	0.133	3	5
C4 Protect property rights of the organization	0.136	0.178	4	4
C5 Protect confidentiality of products and services	0.191	0.190	2	2
C6 Maintain and manage the accuracy of organizational data	0.128	0.190	5	2

In this aspect, the academic experts' C.I. was 0.01; the practitioners' C.I. was 0.00581.

“protect property rights of the organization,” and “maintain and manage the accuracy of organizational data.” The criteria of “manage the ownership of software, hardware and information” and “manage the accessibility of the organizational system” were ranked lowest.

Possible explanations for the aforementioned evaluation results could be that IT academic experts were more management-oriented (García-Holgado et al., 2021), so they focused on the protection of the privacy of the organization's products and services and emphasized the need to maintain the accuracy of management information. As for the IT practitioners, they adopted the perspective of employees in the organization, so the privacy of employee data, customer data, and vendor data that they dealt with every day was

most important to them (D'Arcy and Devaraj, 2012; Lu et al., 2018), followed by the intellectual property rights of organizational resources, such as internal training materials and online databases (Yueh et al., 2016).

The criteria analysis of information professionals in information technology and library and information science fields

This study further integrated the results of the AHP analysis between academic experts and practitioner experts in the two professional fields. As shown in Table 12,

TABLE 12 Weights and ranking of three main criteria by information professionals in IT and LIS fields.

Aspects	C.I.		Weight		Priority	
	IT	LIS	IT	LIS	IT	LIS
General Moral Imperatives	0.01	0.00874	0.545	0.503	1	1
Specific Professional Responsibilities	0.00967	0.00499	0.181	0.208	3	3
Organizational/workplace Responsibilities	0.00918	0.00401	0.274	0.288	2	2

In this aspect, the IT expert's C.I. was 0.000086; the LIS expert's C.I. was 0.00001.

TABLE 13 Information technology experts' local and global weights of all criteria of the hierarchy.

Aspects		Criteria
Information Ethics	General Moral Imperatives (L:0.545)	A1 Contribute to society and human well-being (L:0.095; G:0.052)
		A2 Avoid harm to others (L:0.310; G:0.169)
		A3 Be honest and trustworthy (L:0.208; G:0.113)
		A4 Be fair and take action not to discriminate (L:0.128; G:0.070)
		A5 Honor property rights including copyrights and patents (L:0.118; G:0.064)
		A6 Respect the privacy of others and honor confidentiality (L:0.141; G:0.077)
	Specific Professional Responsibilities (L:0.181)	B1 Achieve the quality, effectiveness and dignity of professional work (L:0.196; G:0.036)
		B2 Acquire and maintain professional competence (L:0.133; G:0.024)
		B3 Know and respect existing laws pertaining to professional work (L:0.157; G:0.028)
		B4 Accept and provide appropriate professional review (L:0.098; G:0.018)
		B5 Honor contracts, agreements, and assigned responsibilities (L:0.210; G:0.038)
		B6 Access computing and communication resources with authorization (L:0.206; G:0.037)
	Organizational/workplace Responsibilities (L:0.274)	C1 Protect the privacy of organizational data (L:0.296; G:0.081)
		C2 Manage the ownership of software, hardware and information (L:0.090; G:0.025)
		C3 Manage the accessibility of the organizational system (L:0.100; G:0.027)
		C4 Protect property rights of the organization (L:0.162; G:0.044)
		C5 Protect confidentiality of products and services (L:0.167; G:0.046)
		C6 Maintain and manage the accuracy of organizational data (L:0.185; G:0.051)

the criterion of “general moral imperatives” was given the highest priority in both the IT and LIS fields, followed by “organizational/workplace responsibilities” and finally “specific professional responsibilities.”

In addition to the above three main criteria weights, the product of the main criteria weights and the sub-criteria weights was further calculated to represent the global priority of each evaluation criterion, and the relative weight rankings of all evaluation criteria were obtained as shown in [Tables 13, 14](#).

After further comparing the 18 sub-criteria, this study found that the most important criterion ranked by IT experts was “avoid harm to others” (GW = 0.113), while the criterion ranked most important by LIS experts was “respect the privacy of others and honor confidentiality” (GW = 0.108), which was slightly higher than “avoid harm to others” (GW = 0.102). The results echoed the findings of the previous studies ([Eskens, 2020](#); [Zimmer et al., 2020](#)), supporting that informaticists, like

the general public, placed the highest value on privacy in information ethics issues.

Many past studies have argued that Orientalism or Asian culture, which is strongly influenced by Confucianism, has shaped certain Asian values, one of which is “interpersonal reciprocity and accommodation” (avoiding conflict with others) ([Nathan, 2012](#)). In Confucian relationships, interpersonal relationships that give special consideration to the law of human relations may create a dilemma of human relations leading to subjective decisions due to the law of equity ([Huang, 2001](#)). Therefore, people in this cultural context tend to avoid conflict (including in verbal and non-verbal communication), prevent hurting others, and take care not to offend others. The results of the professionals' evaluations of these ethical criteria in this study may be related to the aforementioned characteristics of these cultural contextual influences.

TABLE 14 Library and information science experts' local and global weights of all criteria of the hierarchy.

Aspects		Criteria
Information Ethics	General Moral Imperatives (L:0.503)	A1 Contribute to society and human well-being (L:0.084; G:0.042)
		A2 Avoid harm to others (L:0.203; G:0.102)
		A3 Be honest and trustworthy (L:0.157; G:0.079)
		A4 Be fair and take action not to discriminate (L:0.166; G:0.083)
		A5 Honor property rights including copyrights and patents (L:0.176; G:0.089)
		A6 Respect the privacy of others and honor confidentiality (L:0.214; G:0.108)
	Specific Professional Responsibilities (L:0.208)	B1 Achieve the quality, effectiveness and dignity of professional work (L:0.164; G:0.034)
		B2 Acquire and maintain professional competence (L:0.104; G:0.022)
		B3 Know and respect existing laws pertaining to professional work (L:0.149; G:0.031)
		B4 Accept and provide appropriate professional review (L:0.124; G:0.026)
		B5 Honor contracts, agreements, and assigned responsibilities (L:0.240; G:0.050)
		B6 Access computing and communication resources with authorization (L:0.219; G:0.046)
	Organizational/workplace Responsibilities (L:0.288)	C1 Protect the privacy of organizational data (L:0.227; G:0.065)
		C2 Manage the ownership of software, hardware and information (L:0.102; G:0.029)
		C3 Manage the accessibility of the organizational system (L:0.138; G:0.040)
		C4 Protect property rights of the organization (L:0.168; G:0.048)
		C5 Protect confidentiality of products and services (L:0.191; G:0.055)
		C6 Maintain and manage the accuracy of organizational data (L:0.174; G:0.050)

Conclusion

This study used the analytic hierarchical process method to develop an evaluation framework and criteria for information professional ethics, and it also attempted to understand the differences in the ranking of the importance of the criteria for evaluating information professional ethics between practitioners in field of the information technology and those in the field of library and information science. From the results of the study, it was found that the basic attitudes or perceptions of the practitioners in the two professional fields toward the ethical issues of information profession were similar, and both of them focused the most on the general moral imperatives aspect and the need to avoid hurting others or violating others' privacy.

Furthermore, the analysis of the general moral imperatives suggested that expert respondents gave priorities to aspects that were directly and specifically related to personal interests. They were most concerned with causing harm to others and themselves, and more general aspects that caused no immediate harm were weighted less. The expert respondents' evaluations of the specific professional responsibilities suggested that factors that directly related to their work were regarded as the most important. They also gave high priority to legal issues, since the codification of norms suggested clearer guidelines of action to follow. In the aspect of organizational/workplace responsibilities, factors involved with interests and conflicts with external institutions were given the most weight over the

others, such as the protection of organizational information of clients, suppliers, and employees. On the other hand, internal exchange or circulation of information was regarded as less critical. Overall, the findings suggested that moral imperatives as personal traits were regarded as the most important, followed by personal competence in work and finally social interactions with others.

The results of this study can provide information professional organizations, academic researchers and practitioners with important references on information ethics issues. In particular, the results of this study will provide a deeper understanding of the decision-making behaviors of professionals in the face of ethical issues and serve as a reference for planning a comprehensive framework for professional training programs in information ethics education and training. In addition, the cultural contexts and personal characteristics involved in these ethical decisions are worthy of further exploration in future research. Despite considerations of disciplinary diversity, a few limitations of this study should be noticed. Since this study analyzed and compared only the results of subjective questionnaires completed by experts and scholars, their reasons and justifications for the rankings inferred by the researchers could still be insufficient to include all possibilities. Based on the findings of this study, future studies should be conducted to develop teaching materials and assessment tools for information ethics so as to help nurture talents in related information professional fields.

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. The patients/participants provided their written informed consent to participate in this study.

Author contributions

H-PY: conceptualization, methodology, funding acquisition, and writing – original draft, review and editing. C-YH: investigation, formal analysis, and visualization. WL: methodology, investigation, validation, and writing – review and editing. All authors contributed to the article and approved the submitted version.

References

- ACM (1992). *ACM Code of Ethics and Professional Conduct*. New York, NY: ACM.
- ACM (2018). *ACM Code of Ethics and Professional Conduct*. New York, NY: ACM.
- ALA (2021). *Code of Ethics, Adopted at the 1939 Midwinter Meeting by the ALA Council*. Chicago, IL: ALA.
- Anthes, G. H. (1996). Hack attack: cyberthieves siphon millions from US firms. *Computerworld* 30:81.
- Arnold, D. G., and Bowie, N. E. (2019). *Ethical Theory and Business*. Cambridge, MA: Cambridge University Press. doi: 10.1017/9781108386128
- ASIS&T (1992). *Professional Guidelines*. Silver Spring, MD: ASIS&T.
- Buchanan, E. A., Henderson, K. A., and Hauptman, R. (eds). (2009). *Case studies in library and information science ethics*. Jefferson, NC: McFarland & Co.
- Carbo, T. (2008). Ethics education for information professionals. *J. Lib. Admin.* 47, 5–25. doi: 10.1080/01930820802186324
- Chang, C. L.-H. (2011). The effect of an information ethics course on the information ethics values of students—A Chinese guanxi culture perspective. *Comp. Hum. Behav.* 27, 2028–2038. doi: 10.1016/j.chb.2011.05.010
- D'Arcy, J., and Devaraj, S. (2012). Employee misuse of information technology resources: Testing a contemporary deterrence model. *Decision Sci.* 43, 1091–1124. doi: 10.1111/j.1540-5915.2012.00383.x
- Eskens, S. (2020). The personal information sphere: An integral approach to privacy and related information and communication rights. *J. Assoc. Inform. Sci. Technol.* 71, 1116–1128. doi: 10.1002/asi.24354
- Fallis, D. (2007). Information ethics for twenty-first century library professionals. *Library Hi Tech* 25, 23–36. doi: 10.1108/07378830710735830
- Fiesler, C., Garrett, N., and Beard, N. (2020). “What do we teach when we teach tech ethics? A syllabi analysis,” in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, New York, NY. doi: 10.1145/3328778.3366825
- Fleischmann, K. R., and Wallace, W. A. (2009). Ensuring transparency in computational modeling. *Commun. ACM* 52, 131–134. doi: 10.1145/1467247.1467278
- Floridi, L., and Sanders, J. W. (2001). Artificial evil and the foundation of computer ethics. *Ethics Inform. Technol.* 3, 55–66. doi: 10.1023/A:1011440125207
- García-Holgado, A., García-Peñalvo, F. J., Therón, R., Vázquez-Ingelmo, A., Gamazo, A., González-González, C. S., et al. (2021). “Development of a SPOC of computer ethics for students of computer science degree,” in *Proceedings of the 2021 XI International Conference on Virtual Campus (IICV)*, Salamanca. doi: 10.1109/IICV53222.2021.9600272
- Gotterbarn, D., Brinkman, B., Flick, C., Kirkpatrick, M. S., Miller, K., Vazansky, K., et al. (2018). *ACM Code of Ethics and Professional Conduct*. New York, NY: ACM.
- Huang, K. K. (2001). The theoretical construction of Confucian Relationalism and its philosophical foundation. *Formosan Educ. Soc.* 2, 1–33.
- Hunt, S. D., and Vitell, S. (1986). A general theory of marketing ethics. *J. Macromarket.* 6, 5–16. doi: 10.1177/027614678600600103
- IEEE (2020). *Code of Ethics*. Piscataway, NJ: IEEE.
- IFLA (2012). *Code of Ethics for Librarians and Other Information Workers (Long Version)*. The Hague: IFLA.
- Kelton, K., Fleischmann, K. R., and Wallace, W. A. (2008). Trust in digital information. *J. Am. Soc. Inform. Sci. Technol.* 59, 363–374. doi: 10.1002/asi.20722
- LAROC (2002). *Code of Professional Ethics for Librarians in Taiwan*. Brazil: LAROC.
- Lin, C.-H., and Chou, C. (2014). Ethics curricula of the information science departments in Taiwanese universities and colleges. *J. Res. Educ. Sci.* 59, 197–228.
- Lin, W., and Huang, C.-Y. (2013). Understand public libraries' decisions on eBook extension service: An AHP Analysis. *J. Lib. Inform. Stud.* 11, 117–148.
- Lin, W., Wang, J.-Y., and Yueh, H.-P. (2022). Learning information ethical decision making with a simulation game. *Front. Psychol.* 13:933298. doi: 10.3389/fpsyg.2022.933298

Funding

This study was supported by the Taiwan Ministry of Science and Technology (NSC102-2511-S-002-007-MY2).

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.

- Liu, C.-J., and Yang, S. C. (2012). Applying the practical inquiry model to investigate the quality of students' online discourse in an information ethics course based on Bloom's teaching goal and Bird's 3C model. *Comp. Educ.* 59, 466–480. doi: 10.1016/j.compedu.2012.01.018
- Lu, M.-H., Lin, W., and Yueh, H.-P. (2018). How do employees in different job roles in the insurance industry use mobile technology differently at work? *IEEE Trans. Prof. Commun.* 61, 151–165. doi: 10.1109/TPC.2018.2796998
- MacLagan, P. (2003). Varieties of moral issue and dilemma: A framework for the analysis of case material in business ethics education. *J. Bus. Ethics* 48, 21–32. doi: 10.1023/B:BUSI.0000004364.63317.73
- Mason, R. O. (1986). Four ethical issues of the information age. *MIS Q.* 10, 5–12. doi: 10.2307/248873
- McMenemy, D., Poulter, A., and Burton, P. (2014). *A Handbook of Ethical Practice: A Practical Guide to Dealing with Ethical Issues in Information and Library Work*. Amsterdam: Elsevier.
- Nathan, A. (2012). Confucius and the ballot box: Why Asian values do not stymie democracy. *Foreign Affairs* 91, 134–139.
- Nissenbaum, H. (2019). Contextual integrity up and down the data food chain. *Theor. Inq. Law* 20, 221–256. doi: 10.1515/til-2019-0008
- Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Math. Model.* 9, 161–176. doi: 10.1016/0270-0255(87)90473-8
- Saaty, R. W. (1990). How to make a decision: The Analytic Hierarchy Process. *Eur. J. Operat. Res.* 48, 9–26. doi: 10.1016/0377-2217(90)90057-1
- Thong, J. Y., and Yap, C.-S. (1998). Testing an ethical decision-making theory: The case of softlifting. *J. Manag. Inform. Syst.* 15, 213–237. doi: 10.1080/07421222.1998.11518203
- Trepanier, C., Shiri, A., and Samek, T. (2019). An examination of IFLA and Data Science Association ethical codes. *IFLA J.* 45, 289–301. doi: 10.1177/0340035219849614
- Witt, S. (2017). The evolution of privacy within the American Library Association, 1906–2002. *Lib. Trends* 65, 639–657. doi: 10.1353/lib.2017.0022
- Xu, K., and Xu, J. (2020). A direct consistency test and improvement method for the analytic hierarchy process. *Fuzzy Opt. Decision Making* 19, 359–388. doi: 10.1007/s10700-020-09323-y
- Yueh, H.-P., Lu, M.-H., and Lin, W. (2016). Employees' acceptance of mobile technology in a workplace: An empirical study using SEM and fsQCA. *J. Bus. Res.* 69, 2318–2324. doi: 10.1016/j.jbusres.2015.12.048
- Zimmer, M., Vitak, J., and Wu, P. (2020). Editorial introduction: "Information privacy in the digital age". *J. Assoc. Inform. Sci. Technol.* 71, 997–1001. doi: 10.1002/asi.24394



OPEN ACCESS

EDITED BY
Claudio Longobardi,
University of Turin, Italy

REVIEWED BY
Issah Abdulai,
University of Cape Coast, Ghana
Michael Agyemang Adarkwah,
Beijing Normal University, China
Bernard Gumah,
University of Electronic Science and
Technology of China, China

*CORRESPONDENCE
John Elvis Hagan Jr.
elvis.hagan@ucc.edu.gh

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

RECEIVED 29 April 2022
ACCEPTED 15 August 2022
PUBLISHED 29 September 2022

CITATION
Agormedah EK, Ankomah F,
Frimpong JB, Quansah F, Srem-Sai M,
Hagan JE Jr and Schack T (2022)
Investigating teachers' experience and
self-efficacy beliefs across gender in
implementing the new
standards-based curriculum in Ghana.
Front. Educ. 7:932447.
doi: 10.3389/feduc.2022.932447

COPYRIGHT
© 2022 Agormedah, Ankomah,
Frimpong, Quansah, Srem-Sai, Hagan
and Schack. 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.

Investigating teachers' experience and self-efficacy beliefs across gender in implementing the new standards-based curriculum in Ghana

Edmond Kwesi Agormedah¹, Francis Ankomah^{2,3},
James Boadu Frimpong⁴, Frank Quansah⁵, Medina Srem-Sai⁶,
John Elvis Hagan Jr.^{4,7*} and Thomas Schack⁷

¹Department of Business and Social Sciences Education, University of Cape Coast, Cape Coast, Ghana, ²Department of Education and Psychology, University of Cape Coast, Cape Coast, Ghana, ³Department of Education, SDA College of Education, Asokore, Ghana, ⁴Department of Health, Physical Education and Recreation, University of Cape Coast, Cape Coast, Ghana, ⁵Department of Educational Foundations, University of Education, Winneba, Ghana, ⁶Department of Health, Physical Education, Recreation and Sports, University of Education, Winneba, Ghana, ⁷Neurocognition and Action-Biomechanics-Research Group, Faculty of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany

The introduction of the new standards-based curriculum in Ghana required teachers to make adjustments, including teaching routines. Despite the challenges encountered in the introduction of this new curriculum, its passage was successful. Previous studies have revealed mixed reactions from teachers and how their experiences affect their efficacy during the implementation stage. This study examined whether teachers' experiences and self-efficacy beliefs influenced the implementation of the new curriculum reforms across gender in Ghana. The cluster sampling technique was used to survey 693 basic school teachers who responded to a questionnaire. Descriptive and regression-based inferential statistics were used to analyze the data. A preliminary assessment showed that teachers exhibited a moderate-to-high level of teaching efficacy in student engagement, instructional strategy, and classroom management. Furthermore, the study also revealed that teaching experience was positively related to efficacy. However, gender significantly moderated the relationship between teaching experience and teachers' efficacy in student engagement as well as teaching experience and efficacy in student engagement. Findings imply that even though teachers may be more likely to plan and deliver pedagogical content and carry out instructional activities, their sense of efficacy in managing their classroom is questionable. Emphasizing teaching experience in improving their efficacy in curriculum implementation is key, especially among female teachers. Programs or capacity-building training workshops that seek to improve teachers' ability to manage the classroom environment should be organized regularly to promote effective curriculum implementation.

KEYWORDS

beliefs, gender, self-efficacy, standard-based curriculum, teaching experience

Introduction

Quality delivery of education depends on several factors including a quality curriculum, which almost every nation, at some point in time, has carried out some reforms. The Ghanaian educational system, which has undergone several educational reforms, is clustered into three major phases: 9 years of basic education [i.e., early childhood education (kindergarten), primary, and junior high school], 3 years of secondary education (i.e., senior high school/vocational and technical schools), and 3–4 years of tertiary education (i.e., academic university, technical university, colleges of education, and nursing training colleges) (Adu-Gyamfi et al., 2016). Historically, educational reform initiatives in Ghana are both curricular and structural (Quainoo et al., 2020). Recently, in September 2019, the Government of Ghana (GoG), through the Ministry of Education (MoE), Ghana Education Service (GES), and the National Council for Curriculum and Assessment (NaCCA) introduced a new curriculum reform, “the Standards-Based Curriculum” (SBC) at the basic schools (i.e., from kindergarten to primary six) based on the 2017 educational reform. The SBC was implemented to replace the objective-based curriculum (OBC) that the country had adopted since the introduction of formal education as far back as the eighteenth century (National Council for Curriculum and Assessment [NaCCA], 2019). In 1961, Ghana implemented an Education ACT to formalize and regularize the education system (Anum, 2007). Since then, several educational reforms (i.e., 1974, 1987, and 2007) have been introduced to ensure the efficiency and effectiveness of the education system and to provide quality education to citizens (Anum, 2007). These reforms emphasized the OBC until the recent educational reform in 2017, which advocated for SBC (Adu-Gyamfi et al., 2016). According to NaCCA (2019), the rationale for SBC was to respond to the national priority of shifting the structure and content of the education system from merely passing examinations to building character, nurturing values, literacy, confidence, and critical thinking among citizens. The key features of SBC include the acquisition of the 4Rs, namely reading, writing, arithmetic, and creativity; development of core competencies/twenty-first century skills [foundational and lifelong skills]; learner-centered pedagogies (e.g., differentiation, scaffolding, and use of ICT); and inclusion and diversity in education (NaCCA, 2019).

The new curriculum reform has caused a complete shift in teaching activities, which required a radical change in teachers’ roles from the agents of knowledge transmission to the facilitators of student knowledge acquisition (NaCCA, 2019). This approach has led to several concerns among teachers, which have affected the quality of implementation of the new curriculum (SBC) (Apau, 2021). Previous studies on SBC implementation in Ghana established that teachers have informational concerns (i.e., need more information on reforms), personal concerns (i.e., worried about how

the new reform will affect their teaching beliefs, confidence, competencies, and knowledge), and management concerns (i.e., worried about instructional resources, absence of textbooks, time, and logistics) about the new curriculum reforms (e.g., Aboagye and Yawson, 2020; Apau, 2021; Mpuangnan and Adusei, 2021). These concerns could affect teachers’ teaching sense of self-efficacy beliefs in implementing the SBC. Some scholars have found that effective curriculum implementation in schools is a function of teachers’ instructional competencies (knowledge and skills), perceptions, and confidence (self-efficacy beliefs) (Smith, 1996; Marsh, 1997; Roehrig et al., 2007).

Drawing from social cognitive theory (Bandura, 1977), self-efficacy is considered as “beliefs in one’s capabilities to organize and execute the course of action required to manage prospective situations” (Bandura, 1986, p. 3). In line with this definition, Tschannen-Moran and Woolfolk Hoy (1998) defined a teacher’s sense of self-efficacy belief as “the teacher’s belief in his or her capability to organize and execute the course of action required to successfully accomplish a specific teaching task in a particular context” (p. 22). Tschannen-Moran and Woolfolk Hoy (2001) also conceptualized teachers’ sense of self-efficacy beliefs as a function of efficacy in instructional strategy, classroom management, and student engagement. Efficacy in instructional strategy is about confidence in the teacher’s ability to deploy appropriate methods, procedures, and materials (i.e., provide appropriate learning experiences) to transmit the curriculum content to the learners (Tschannen-Moran and Woolfolk Hoy, 2001). Efficacy in classroom management is about the teacher’s belief in their ability to use appropriate skills and techniques to establish classroom rules and routines, direct and control learners’ misbehaviors, and arrange the physical and social conditions to create a quality learning environment for learners (Tschannen-Moran and Woolfolk Hoy, 2001). Efficacy in student engagement refers to the teacher’s confidence in their ability to organize and implement learning experiences or activities that would draw learners’ attention and develop their curiosity, interest, optimism, and passion for the learning tasks during the instructional process (Tschannen-Moran and Woolfolk Hoy, 2001). These beliefs result in an individual teacher’s decision to meet the demands of the current task.

Empirical studies have shown that teachers with a high level of teaching self-efficacy beliefs are more likely to adapt to educational innovations and, consequently, foster their implementation of curriculum reforms (e.g., Ghaith and Yaghi, 1997; Evers et al., 2002; Weisel and Dror, 2006); accept and use new innovative strategies of teaching (e.g., Haverback and Parault, 2008; Pfitzner-Eden, 2016; Snyder and Fisk, 2016); provide classroom planning and organization (e.g., Dibapile, 2012); and ensure quality instruction (e.g., Klassen et al., 2011; Holzberger et al., 2012; Klassen and Tze, 2014; Zee and Koomen, 2016). Teachers’ sense of self-efficacy leads to effective curriculum implementation (Snyder et al., 1992; Fullan, 1994).

Fullan (1994) argued in his research that when teachers have a greater sense of self-efficacy, it leads them to act and persist in the effort required to ensure successful curriculum implementation.

Extant researchers have reported different levels of teachers' self-efficacy beliefs in implementing the curriculum. For example, Hodges et al. (2016) reported that teachers exhibited high confidence (self-efficacy) levels when implementing the problem-based science curriculum in the USA. In Ghana, teachers exhibited a high level of self-efficacy in teaching beliefs (i.e., instructional strategies, student engagement, and classroom management) toward implementing the Ghanaian Language and Culture curriculum (Bassah, 2020) and the Social Studies curriculum (Arko, 2021). Similar findings are reported in studies conducted in Turkey (Kabaoglu, 2015) and Kenya (Wang'eri and Otanga, 2014). However, teachers with a moderate level of self-efficacy beliefs in implementing the curriculum have been reported in Tanzania (Jumane, 2012), the Caribbean (Jameson-Charles and Jaggernauth, 2015), Malaysia (Abdullah and Kong, 2016; Yusof and Nor, 2017), and Indonesia (Susilanas et al., 2018).

A plethora of investigations have also produced contradictory results about the influence of gender and teaching experiences on teachers' self-efficacy in implementing the curriculum. For example, a significant difference in teaching self-efficacy beliefs (classroom management, instructional strategies, and student engagement) of teachers with respect to gender was reported in India (Chandrika and Varma, 2022), Ethiopia (Butucha, 2014), and Kenya (Wang'eri and Otanga, 2014). Female teachers reported significantly higher teaching self-efficacy than their male counterparts in Iran (Karimvand, 2011) and Pakistan (Ahmad et al., 2015) while male teachers were more efficacious than females in Malaysia (Abdullah and Kong, 2016). Other studies have also discovered no gender influence on self-efficacy beliefs of teachers in curriculum implementation (e.g., Odanga et al., 2015; Yusof and Nor, 2017; Bassah, 2020; Rezaeian and Abdollahzadeh, 2020; Arko, 2021).

Previous studies have also established a significant relationship between teaching experience and teachers' self-efficacy beliefs during curriculum implementation. For example, George et al. (2018) in Australia found that teachers' self-efficacy increased across all dimensions of self-efficacy as they progressed from their first to the fifth year of teaching. Also, in Iran, Karimvand (2011) established that teachers with more teaching experience (e.g., 5 years and above) had significantly higher self-efficacy beliefs than teachers with less teaching experience (e.g., 1–3 years). Similar findings were reported in the Caribbean (Jameson-Charles and Jaggernauth, 2015), Malaysia (Yusof and Nor, 2017), and Kenya (Wang'eri and Otanga, 2014).

Ghana practices a centralized education system where the government makes curriculum planning, designing and

development, implementation and evaluation decisions through its educational policymakers like NaCCA, MoE, and GES (Anum, 2007; Adu-Gyamfi et al., 2016). The key curriculum decisions are vested in NaCCA (a centralized body) under the supervision of the GES and MoE. This body is mandated to develop a comprehensive curriculum document to guide teachers in teaching (NaCCA, 2019). Since Ghana operates within the centralized educational system, the fidelity approach to curriculum implementation is endorsed. Teachers are expected to faithfully comply, use, and implement the official curriculum as intended by the authorities without any deviation (Cobbold, 2017). Any minimum degree of deviation might distort the true meaning of what is intended to be implemented. Despite teachers being recognized as the pinnacle of effective curriculum implementation, their roles in all these curriculum decisions are infinitesimal in Ghana (Cobbold, 2017). Teachers in Ghana are not directly involved in the curriculum decision-making process. This situation has raised major concerns among teachers with respect to the implementation of any innovation like the SBC.

In Ghana, the only official introduction of SBC was through the government's national address. Since then, the GES and MoE, through the regional and district directors of education, were made to organize workshops and training sessions for the teachers to brief them on the new curriculum (i.e., SBC) and what is expected of them during the implementation in September 2019 (Kpedator, 2019). The workshops and training sessions were organized on a regional and district basis and lasted only 5 days (Kpedator, 2019). However, the teachers were highly dissatisfied with the organization of the workshops due to some challenges they encountered. As a result, some teachers quit the workshop sessions because their needs were not met and addressed (Arhinful and Tijani, 2020). Since the introduction of the SBC to date, it appears no major measures were put in place to inform its implementation process/rolling out. These issues have made teachers worry or express concerns about the SBC (Apau, 2021). The concerns of the teachers could negatively affect their beliefs and judgments or confidence in their capabilities to bring about the desired outcomes of SBC. Thus, their ability to handle the tasks, obligations, and professional requirements in implementing the SBC.

Taken together, teachers play a pivotal role in the successful implementation of educational reforms, especially with curriculum reforms. Hence, teachers' self-efficacy beliefs and experiences may be an essential source of motivation to implement curriculum reform initiatives. Since the introduction of SBC in 2019, empirical studies on teachers' experience and self-efficacy beliefs across gender in implementing the curriculum have been lacking. This study aimed to examine whether teachers' experience and self-efficacy beliefs will influence the implementation of the new curriculum reforms across gender in Ghana. The following specific objectives were outlined:

1. Assess the level of teachers' sense of teaching self-efficacy beliefs in curriculum implementation
2. Establish the influence of gender on teachers' sense of teaching self-efficacy beliefs in curriculum implementation
3. Determine the moderating effect of gender in the relationship between teaching experience and teachers' sense of teaching self-efficacy beliefs in curriculum implementation.

The outcome of this inquiry provides valuable information on the degree of teachers' experience and self-efficacy beliefs in implementing the SBC in Ghana and how teachers' profiles influence their teaching self-efficacy beliefs during the implementation. Teachers' experiences and beliefs regarding curriculum reforms are important elements of educational change, as there is a relationship between what teachers believe and what they do in the classroom. Identified findings may provide important insights into how teachers approach the implementation of new curricula and may help explain variations in curriculum implementation. Investigating teachers' self-efficacy beliefs not only has the potential to result in successful curriculum implementation in Ghana but can also provide input for interesting and effective ways of teacher development programs with an emphasis on educating competent teachers. Based on the foregoing literature, a conceptual model was developed for the investigation (see Figure 1).

Materials and methods

Participants' selection

The population comprised 2,145 basic school teachers from seven (7) study centers clustered within the three zones of Ghana, namely the Northern Zone, the Middle Zone, and the Southern Zone. This is composed of 325, 1,100, and 720 teachers for the Northern, Middle, and Southern zones, respectively. The participants had completed a Diploma in Basic Education from their initial training at the Colleges of Education (CoE) in Ghana. It should be noted that all the teachers in the population had already been employed in GES for teaching. They were officially enrolled to pursue sandwich degree programs (as a top-up degree program) within two and a half years in different academic-related courses at the University of Cape Coast in the 2019/2020 academic year. This sandwich program aims to help basic school teachers develop professionally while keeping the majority of them in the classroom. The program commences as soon as basic schools vacate to enable these teachers to attend lectures at different study centers across various CoEs in Ghana, clustered within the three zones of Ghana. Any in-service basic school teacher enrolled in the UCC sandwich degree program at the time of data collection was qualified to be involved in

the study. The descriptive cross-sectional survey design was employed, through cluster sampling to sample 693 basic school teachers (male $n = 374$, 54%, female $n = 319$, 46%), with ages categorized as, below 25 years ($n = 61$, 8.8%), 25–29 years ($n = 458$, 66.1%), and 30 years and above ($n = 174$, 25.1%). The teaching experiences of participants were categorized into 1–5 years ($n = 563$, 81.2%), 6–10 years ($n = 107$, 15.4%), and 10 years and above ($n = 23$, 3.3%). The sample distributions across the three zones were 105, 355, and 233 participants from the Northern, Middle, and Southern zones, respectively.

Instrumentation

A questionnaire was used to collect data to answer the research objectives raised in the study. The questionnaire was structured to contain an informed consent letter and the survey items. The instrument had a socio-demographic section (i.e., gender, age, and teaching experience), followed by the teaching self-efficacy beliefs section. The three major variables of interest in this research include gender, teaching experience, and teaching self-efficacy beliefs. The measurements of these key variables are explained in the subsequent section.

Gender

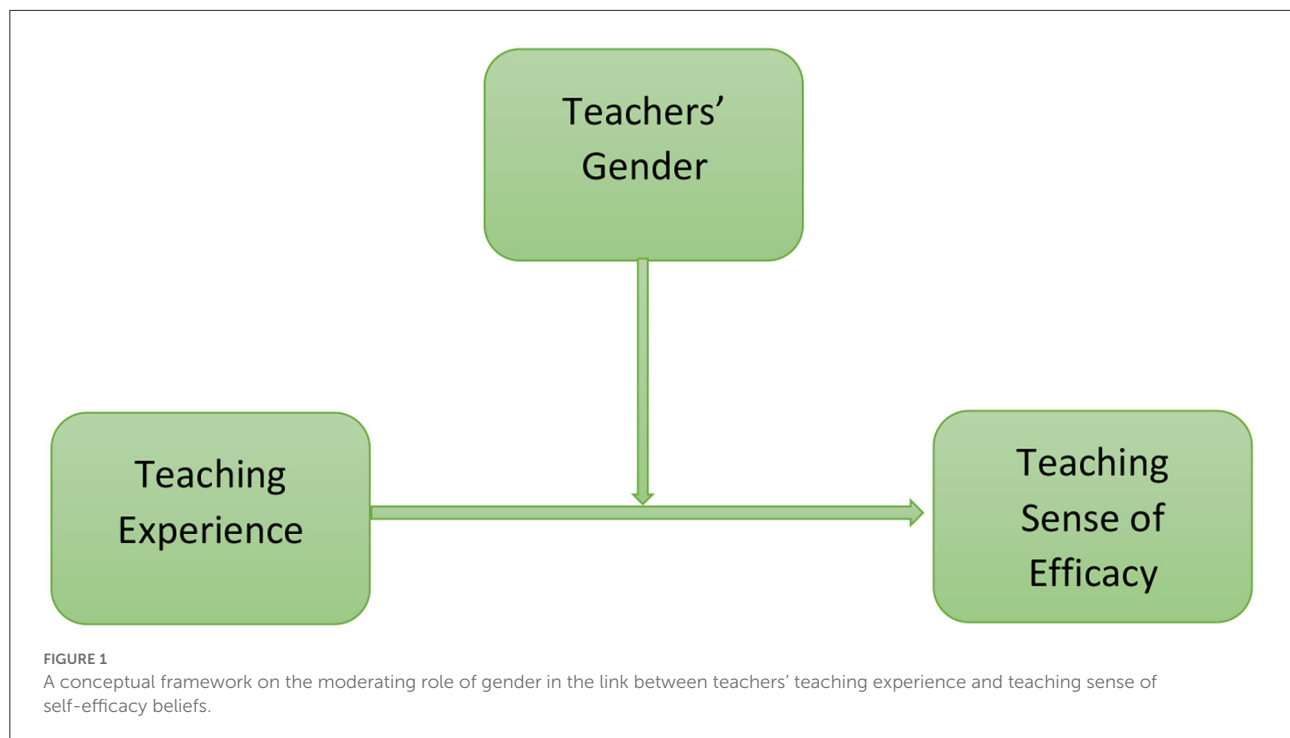
Gender was conceptualized as an individual's classification, generally as male or female or intersex, based on their reproductive organs and functions as informed by the chromosomal complement (Institute of Medicine, 2001; Short et al., 2013). As such, respondents were asked to indicate their gender by indicating either male, female, or intersex.

Teaching experience

The teaching experience was measured using the number of years that the respondent had taught as an in-service teacher, as endorsed by several previous studies (Adeyemi, 2008; Irvine, 2019; Ansah et al., 2020; Graham et al., 2020). Thus, respondents were required to indicate the number of years (approximated years) they had taught in the Ghanaian school system. Internships and off-campus teaching experiences were not considered.

Teachers' sense of teaching self-efficacy beliefs

The teachers' sense of efficacy scale (TSES) by Tschannen-Moran and Woolfolk Hoy (2001) was adapted to measure teachers' sense of teaching self-efficacy beliefs in curriculum



implementation. The scale had three dimensions, namely, student engagement, instructional strategy, and classroom management. Each dimension had eight items. Some of the questions on the scale include: “How much can you do to get through to the most difficult students” (efficacy in student engagement); “How much can you do to control disruptive behavior in the classroom” (efficacy in classroom management); and “How much can you gauge student comprehension of what you have taught” (efficacy in instructional strategies). The items were scored on a 6-point response scale, with ‘1’ nothing, ‘2’ very little, ‘3’ little, ‘4’ some influence, ‘5’ quite a bite, and ‘6’ a great deal. The scale had Cronbach’s alpha estimates of 0.87, 0.91, and 0.90 for the dimensions of student engagement, instructional strategy, and classroom management, respectively. The following values were obtained using the McDonald’s Omega (ω) coefficients: 0.568, 0.893, and 0.570 for student engagement, instructional strategy, and classroom management dimensions, respectively. Similarly, using Cronbach’s alpha, 0.618, 0.892, and 0.602 were obtained for student engagement, instructional strategy, and classroom management dimensions, respectively. The coefficients were deemed sufficient (Quansah, 2017; Pallant, 2020).

Data collection procedure

Following ethical approval from the University of Cape Coast’s Institutional Review Board (UCCIRB) with identification

number UCCIRB/CES/2020/81, further approval was sought from the Institute of Education at UCC which oversees the management of the sandwich program. Participants who volunteered to participate in the study signed written informed consent forms prior to data collection. Furthermore, the study participants were directly recruited from their various study centers with the help of their lecturers, who consented to assist in collecting the data during the 2019/2020 sandwich session. These lecturers were trained for 2 days to assist in this process because they had direct access to the in-service (sandwich) teachers and had already established rapport with them. Having direct contact with the in-service teachers, the lecturers were deemed helpful because they could facilitate the data collection process.

During the study, other ethical issues were also addressed, including confidentiality, willingness to be involved or otherwise in the study, and withdrawal from responding to the survey items anytime they deemed appropriate. Moreover, participants were assured that their information would be kept anonymous without any linkage with their personal identities. The research assistants thoroughly explained all items on the questionnaire to participants and asked them to seek further clarification if need be. The research assistants distributed the survey instruments to the participants to respond to before class each day to ensure that the participants’ lecture periods were not disrupted. Each questionnaire was answered within 10–15 min, while the entire data collection period took ~2 months (15 November 2020 to 8 January 2021). The

TABLE 1 Descriptive statistics on study variables.

No.	Variables	1	2	3	4	5
1	Student engagement	1				
2	Instructional strategy	0.346**	1			
3	Classroom management	0.671**	0.323**	1		
4	Teaching experience	0.745**	0.723**	0.712**	1	
5	Gender	0.117**	−0.015	0.063	0.080*	1
–	Mean	4.161	4.078	4.030	3.89	–
–	SD	0.858	1.278	0.846	0.674	–
–	Skewness	0.147	−0.337	0.359	0.237	–
–	Kurtosis	−0.693	−0.785	−0.323	−0.311	–

**Significant at $p < 0.001$, *Significant at $p < 0.005$.

research assistants collected all the answered questionnaires and gave them to the researchers in sealed brown envelopes for safekeeping.

Data analysis

Having screened and edited the data collected, outliers were checked, with none identified. Descriptive statistics such as mean, SD, skewness, and kurtosis were used to analyze the data. Pearson's correlation was used to assess the relationships among student engagement, instructional strategy, and classroom management dimensions. Frequency counts and percentages were used to analyze the levels of teaching self-efficacy beliefs, whereas multivariate simple linear regression analysis was used to examine the influence of gender on teachers' sense of self-efficacy beliefs. Considering the dependent variables, a stringent alpha of 0.017 was used. Using a simple moderation analysis, Hayes' PROCESS (model 1), with 5,000 bootstrap samples, was employed to test the moderation effect of gender in the relationship between teaching experience and teachers' sense of teaching self-efficacy beliefs in curriculum implementation. The gender of the teachers was dummy coded. In a dummy variable for males, all cases in which the respondent was male were coded as one, and all other cases in which the respondent was female were coded as 0 (i.e., the females were held as a reference group).

Results

Descriptive statistics on study variables

Descriptive information and correlations among the study variables are presented in Table 1.

TABLE 2 Level of teachers' sense of teaching self-efficacy beliefs.

Variable	Level	Frequency	Percent
Student engagement	Low	46	6.6
	Moderate	494	71.3
	High	153	22.1
Instructional strategy	Low	149	21.5
	Moderate	324	46.8
	High	220	31.7
Classroom management	Low	63	9.1
	Moderate	519	74.9
	High	111	16.0

From Table 1, the study variables were positively related. The relationship was highest between self-efficacy beliefs in student engagement and self-efficacy beliefs in classroom management ($r = 0.671$). A weak relationship was found between self-efficacy beliefs in student engagement and self-efficacy beliefs in instructional strategy ($r = 0.346$), self-efficacy beliefs in classroom management and self-efficacy beliefs in instructional strategy ($r = 0.323$). It was further shown that teachers demonstrated the highest sense of self-efficacy beliefs in student engagement ($M = 4.161$, $SD = 0.858$), followed by instructional strategy ($M = 4.078$, $SD = 1.278$). The teaching experience was positively associated with all the dimensions of teaching self-efficacy beliefs. The self-efficacy beliefs in classroom management ($M = 4.030$, $SD = 0.846$) were the least. Gender had a weak positive link with teachers' self-efficacy beliefs in student engagement, classroom management, and teaching experience but a weak negative association with self-efficacy beliefs in instructional strategy. The skewness and kurtosis for all the variables were within the recommended ranges of ± 2 and ± 7 , respectively (Hair et al., 2010).

Level of teachers' sense of teaching self-efficacy beliefs in SBC implementation

The study assessed the extent of teaching self-efficacy beliefs of teachers in the implementation of SBC. The efficacy scores ranged from 1 to 6. The scores were categorized as 1.0–2.99, 3.0–4.99, and 5.0–6.0 as low, moderate, and high, respectively. Table 2 presents the details.

As shown in Table 2, for all the dimensions of teaching self-efficacy beliefs, majority of the teachers exhibited moderate levels of teaching self-efficacy beliefs in the implementation of the SBC: student engagement ($n = 494$, 71.3%); instructional

strategy ($n = 324$, 46.8%); and classroom management ($n = 519$, 74.9%). However, quite a number of teachers possessed low self-efficacy beliefs in instructional strategy ($n = 149$, 21.5%).

Influence of gender on teachers' sense of teaching self-efficacy beliefs in curriculum implementation

Multivariate linear regression with a stringent alpha of 0.017 was used to examine the influence of gender on teachers' sense of teaching self-efficacy beliefs in SBC implementation. The results are summarized in Table 3.

Gender had no significant influence on teachers' sense of self-efficacy beliefs in student engagement, $B = 0.103$, $t = 1.607$, $p = 0.108$; instructional strategy, $B = 0.002$, $t = 0.024$, $p = 0.981$; and classroom management, $B = 0.106$, $t = 1.652$, $p = 0.099$ (Table 3). Practically, the magnitude of the educational importance of gender on teachers' sense of teaching self-efficacy beliefs in the implementation of the SBC was almost insignificant for all three dimensions of teaching self-efficacy beliefs (Cohen, 1988).

Moderating effect of teachers' gender in the relationship between teaching experience and self-efficacy beliefs in SBC implementation

The moderating role of gender in the relationship between teaching experience and each of the three dimensions of teachers' sense of teaching self-efficacy beliefs in SBC implementation is shown in Table 4.

Gender did not moderate the relationship between teaching experience and teachers' sense of self-efficacy beliefs in instructional strategies, $B = 0.092$, $t = 0.438$, Boot 95% CI $[-0.319, 0.503]$, $f^2 = 0.007$ (Table 4). However, the relationship between teaching experience and teachers' sense of self-efficacy beliefs in student engagement was significantly moderated by gender, $B = 0.323$, $t = 2.317$, boot 95% CI $[0.049, 0.596]$, $f^2 = 0.026$. Practically, the magnitude of the effect was very small. Similarly, gender significantly moderated the relationship between teaching experience and teachers' sense of self-efficacy beliefs in classroom management, with a very small effect size, $B = 0.312$, $t = 2.267$, boot 95% CI $[0.042, 0.583]$, $f^2 = 0.017$. Further probe into the moderations showed that, both in the case of student engagement and classroom management, the relationships were stronger for women than for men (see Figures 2A,B).

Discussion

The study's first objective assessed the level of teachers' sense of teaching self-efficacy beliefs in implementing the SBC. It was found that teachers possessed a moderate to high level of teaching self-efficacy beliefs in student engagement, instructional strategy, and classroom management. The trend of results was not so encouraging since a moderate level of teaching self-efficacy beliefs reflects a moderate degree of exactness in SBC implementation among teachers (Bassah, 2020). This pattern of results suggests that a high level of self-efficacy beliefs is required by teachers for successful implementation of the SBC. Teachers having a moderate self-efficacy belief implies that teachers do not have the full conviction of performing their roles regarding the implementation of the SBC (Kuyini et al., 2020). Undoubtedly, the implementation of the SBC in Ghana saw several challenges, which included non-availability and accessibility of teaching and learning materials, textbooks, syllabi, assessment literacy, and pedagogical logistics (Kpedator, 2019; Ansah et al., 2020; Quansah and Ankoma-Sey, 2020; Nugba et al., 2021; Quansah, 2021). Particularly, Kpedator (2019) showed that the 1-week workshop organized to train the teachers to implement the new curriculum was insufficient. Given these challenges, it is not surprising that teachers had a moderate level of teaching self-efficacy beliefs in the implementation of the SBC, with other previous studies showing similar findings (Jumanne, 2012; Jameson-Charles and Jaggernauth, 2015; Abdullah and Kong, 2016; Yusof and Nor, 2017; Susilanas et al., 2018). With the studies that showed contradictory findings (i.e., teachers exhibiting high levels of teaching self-efficacy beliefs), majority of the teachers sampled for such studies had between 6 and 10 years of teaching experience as compared to this study, where an overwhelmingly high number of the teachers had teaching experience ranging between 1 and 5 years. Arguably, the disparities in the years of teaching could explain the variations in the results.

Regarding the influence of gender on teachers' sense of teaching self-efficacy beliefs in SBC implementation, teachers' gender did not influence their sense of teaching self-efficacy beliefs in implementing the new curriculum. Practically, the magnitude of the educational importance of gender on teachers' sense of teaching self-efficacy beliefs in the implementation of the SBC was almost insignificant. This observation supports some studies that have been documented in literature (Odanga et al., 2015; Yusof and Nor, 2017; Bassah, 2020; Rezaeian and Abdollahzadeh, 2020; Arko, 2021). For instance, Arko (2021) and Bassah (2020) attested that gender did not have any substantial influence on the sense of teaching self-efficacy beliefs of teachers in the implementation of newly implemented curricula. Other studies in India

TABLE 3 Regression parameters of the influence of gender on teachers' sense of teaching self-efficacy beliefs in SBC implementation.

Dependent variable	Parameter	B	Std. error	t	p	f ²
Student engagement	Intercept	4.057	0.047	86.286	0.000	0.004
	Male	0.103	0.064	1.607	0.108	
Instructional strategy	Intercept	4.054	0.072	56.324	0.000	0
	Male	0.002	0.098	0.024	0.981	
Classroom management	Intercept	3.973	0.047	84.017	0.000	0.004
	Male	0.106	0.064	1.652	0.099	

Reference group = Female.

TABLE 4 Interaction effect of teachers' gender in the relationship between teaching experience and self-efficacy beliefs in SBC implementation.

Model	Parameter	B	SE	t	LLCI	ULCI	f ²
1	Constant	4.439	0.141	31.491	4.162	4.716	0.026
	Teaching experience	−0.328	0.113	−2.911	−0.549	−0.107	
	X1	−0.178	0.180	−0.989	−0.531	0.176	
	X1*Teaching experience	0.323	0.139	2.317	0.049	0.596	
2	Constant	4.418	0.212	20.845	4.002	4.834	0.007
	Teaching experience	−0.271	0.169	−1.599	−0.603	0.062	
	X1	−0.134	0.271	−0.493	−0.665	0.398	
	X1*Teaching experience	0.092	0.209	0.438	−0.319	0.503	
3	Constant	4.373	0.140	31.344	4.099	4.647	0.017
	Teaching experience	−0.340	0.111	−3.050	−0.559	−0.121	
	X1	−0.260	0.178	−1.459	−0.610	0.090	
	X1*Teaching experience	0.312	0.138	2.267	0.042	0.583	

Outcome variables: Model 1, student engagement; Model 2, instructional strategy; model 3, classroom management; X1: male; female, reference group; SE, standard error; LLCI, lower level of the confidence interval; ULCI, upper level of the confidence interval.

(Chandrika and Varma, 2022), Ethiopia (Butucha, 2014), and Kenya (Wang'eri and Otanga, 2014), however, noted the significant influence of gender on teachers' self-efficacy beliefs in the implementation of new curricula. For example, while female teachers in Iran (Karimvand, 2011) and Pakistan (Ahmad et al., 2015) were found to have higher teaching self-efficacy beliefs compared to their male peers, male teachers in Malaysia (Abdullah and Kong, 2016) were more efficacious than female teachers in the implementation of the new curriculum. These variations could be as a result of the variations in cultural codes. For example, in communities where males are seen as superior to their female counterparts, they are likely to exhibit higher levels of self-efficacy beliefs. Other factors that could explain the discrepancy in the findings include determinants such as teaching experience and pedagogical skills.

This study has revealed that teaching experience was positively related to teachers' sense of teaching self-efficacy beliefs. Although the SBC is new to teachers, the pedagogy,

subject matter knowledge, and classroom practices required for successful implementation may not be too different from the old curriculum. Given this premise, experienced teachers have been found to exhibit higher levels of teaching self-efficacy beliefs in implementing the SBC in Ghana (Apau, 2021). Several other studies have also supported the claim that teachers with more than 4 years of teaching experience demonstrate a higher level of teaching self-efficacy beliefs (Karimvand, 2011; Wang'eri and Otanga, 2014; Jameson-Charles and Jaggernauth, 2015; Yusof and Nor, 2017). Additionally, gender significantly moderated the relationship between teaching experience and teachers' sense of self-efficacy beliefs in student engagement and teaching experience and self-efficacy beliefs in student engagement. For the same number of years, male teachers are more likely to exhibit higher levels of self-efficacy beliefs in SBC implementation (i.e., classroom management and student engagement) than their female counterparts. This finding reflects the observations

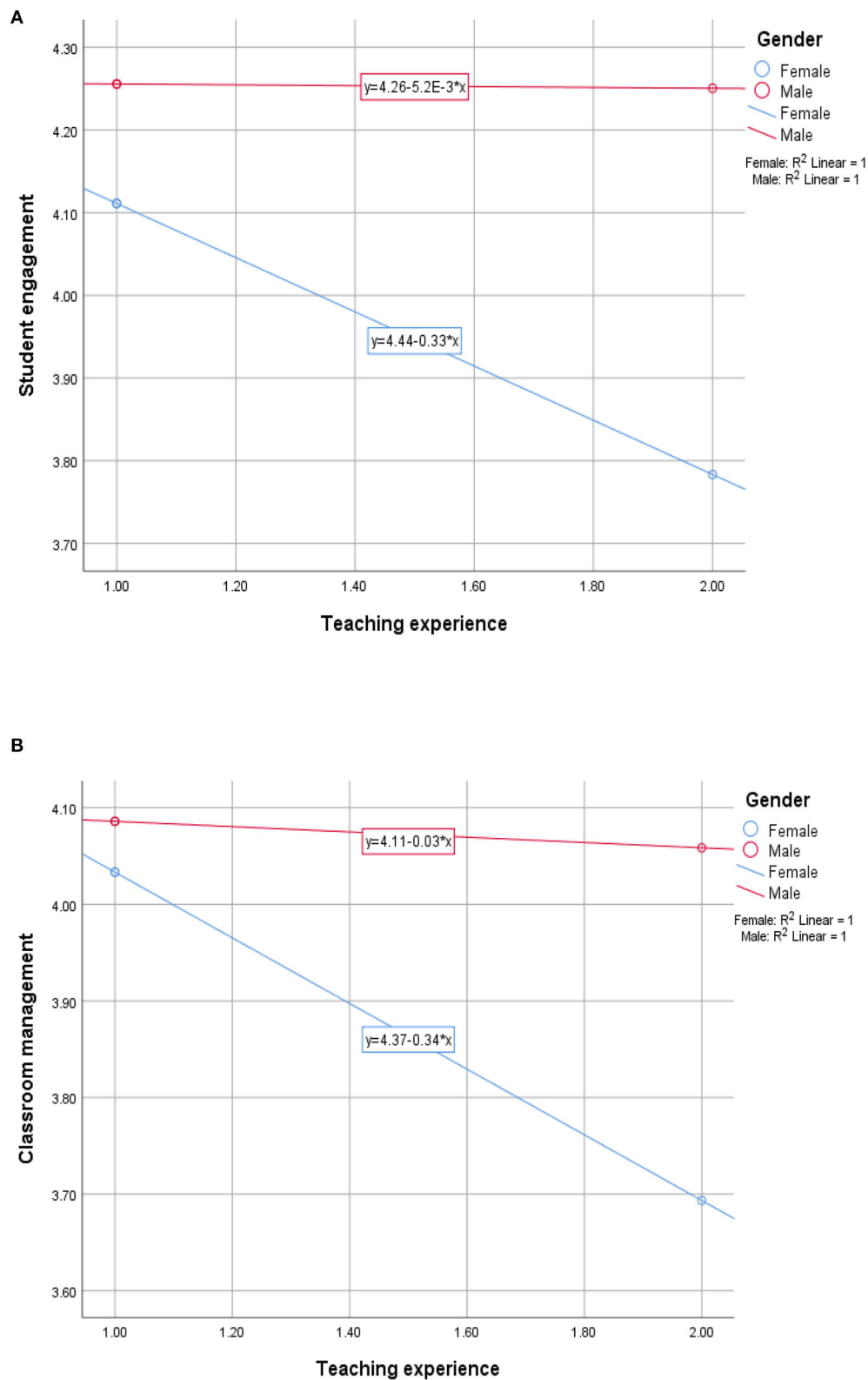


FIGURE 2

(A) Moderation effect of gender in the relationship between teaching experience and self-efficacy beliefs in student engagement. (B) Moderation effect of gender in the relationship between teaching experience and self-efficacy beliefs in classroom management.

of studies that compared male and female teachers with the same years of teaching, which indicated that male teachers often maintain classroom discipline, control disruptive actions, and can engage students in a class (Tabak et al., 2003; Gurbuzturk and Sad, 2009; Shaukat and Iqbal, 2012). However, other studies have also found contradictory results due to methodological limitations, samples used, and context-specific variations (Karimvand, 2011; Odanga et al., 2015; Bassah, 2020; Apau, 2021).

Strengths and limitations

The use of statistical analyses, a larger sample size ($n = 693$), and standardized instruments make it possible to verify the results through further statistical computations, generalize the findings to the population, and replicate the study in a different setting using different samples are the strengths of this study. The study acknowledges certain limitations. First, the descriptive nature of the study does not call for a cause-and-effect relationship among the variables. Again, the use of questionnaires might have introduced respondents' recall biases and other social desirability concerns (e.g., under or over-reporting). Moreover, as self-reported questionnaires were used to collect data from the participants simultaneously, common method variance or bias may be a concern. This approach might affect the reliability and validity of the findings. Furthermore, studying teachers in a sandwich degree program might lead to biased findings because such a group of individuals may not represent all the teachers who are teaching in the basic schools. Also, using lecturers as research assistants during the data collection may influence the participants to respond to the items in a biased manner. However, efforts were made to avoid these biases. The study could not establish and account for differences in teachers' self-efficacy beliefs across different subjects because the teachers who participated in the study were classroom teachers (i.e., they teach all the subjects in their respective classes).

Practical implications

The study stresses the role of teaching experience and its interaction with gender on the levels of teacher self-efficacy beliefs in the SBC implementation. The findings of the study call on the NaCCA and GES to organize more workshops to fully build teachers' capacity and/or sense of teaching self-efficacy beliefs to enhance the successful implementation of the SBC. Notably, the training programs and follow-up seminars need to focus on female and less experienced teachers. Taking a clue from Bandura's social cognitive theory in relation to the findings, teachers could develop their teaching self-efficacy beliefs through interactions

with relevant peers, imitation, observation, and modeling. Based on this tenet and the research findings, less experienced teachers should be paired with more experienced teachers as mentors in implementing the SBC. More mentor-mentee programs could be encouraged to build or promote teachers' self-efficacy beliefs, especially females, in implementing SBC. This approach seeks to build the teaching self-efficacy beliefs of less experienced teachers. Since teaching self-efficacy beliefs affect teachers' behavioral outcomes and students' academic achievement, much attention should be given to teachers with low self-efficacy beliefs. This suggestion is because teachers with common teaching self-efficacy beliefs are more likely to be dissatisfied and demotivated with their teaching job. This assumption could negatively affect their teaching behaviors, teaching quality, and students' academic performances.

Conclusion

The study found moderate levels of teaching self-efficacy beliefs in student engagement, instructional strategy, and classroom management. These findings imply that even though teachers may be more likely to plan and deliver content and carry out pedagogical and instructional activities, their teaching self-efficacy beliefs in managing their classrooms are questionable. Teachers' gender did not influence their sense of teaching self-efficacy beliefs in implementing the new curriculum. However, gender moderated the relationship between teaching experience and teachers' self-efficacy beliefs (i.e., student engagement and classroom management), with males showing considerable levels of teaching self-efficacy beliefs compared to their female counterparts. The findings draw on teaching experience in improving the teaching self-efficacy beliefs of teachers in the SBC implementation across gender. Programs or capacity-building training and workshops that seek to improve teachers' ability to manage their classroom affairs in connection with effective SBC implementation should be organized. Such programs should pay particular attention to female teachers and teachers with few years of teaching experience to help in the SBC implementation in the various schools.

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 Institutional Review Board of

University of Cape Coast, Ghana, with a reference number: UCCIRB/CES/2020/81. The participants provided their written informed consent to participate in this study.

Author contributions

EA and FA conceived the idea. FA performed the analysis. EA, FA, JF, FQ, MS-S, JH, and TS prepared the initial draft of the manuscript. All authors thoroughly revised and approved the final version of the manuscript.

Acknowledgments

The authors are sincerely thankful to Bielefeld University, Germany, for providing financial support for the article processing charge for the manuscript through the Open Access Publication Fund.

References

- Abdullah, M. S. B., and Kong, K. (2016). "Teachers' sense of efficacy with regards to gender, teaching options and experience," in *Proceedings of ISER 17th International Conference*. Auckland (2016).
- Aboagye, E., and Yawson, J. A. (2020). Teachers' perception of the new educational curriculum in Ghana. *Afr. Educ. Res. J.* 8, 6–12. Available online at: <https://eric.ed.gov/?id=EJ1242678> (accessed May 23, 2022).
- Adeyemi, T. O. (2008). Teachers teaching experience and students learning outcomes in secondary schools in Ondo state, Nigeria. *Educ. Res. Rev.* 3, 204–212. doi: 10.4314/ajesms.v3i1.38607
- Adu-Gyamfi, S., Donkoh, W. J., and Addo, A. A. (2016). Educational reforms in Ghana: past and present. *J. Educ. Hum. Dev.* 5, 158–172. doi: 10.15640/jehd.v5n3a17
- Ahmad, R. N., Khan, S. A., and Rehman, S. (2015). Comparative study to investigate the sense of teacher efficacy between male and female teachers. *Asian J. Manag. Sci. Educ.* 4, 29–35. Available online at: [http://www.ajmse.leena-luna.co.jp/AJMSEPDFs/Vol.4\(2\)/AJMSE2015\(4.2-04\).pdf](http://www.ajmse.leena-luna.co.jp/AJMSEPDFs/Vol.4(2)/AJMSE2015(4.2-04).pdf) (accessed May 23, 2022).
- Ansah, J. K., Quansah, F., and Nugba, R. M. (2020). 'Mathematics achievement in crisis': Modelling the influence of teacher knowledge and experience in senior high schools in Ghana. *Open Educ. Stud.* 2, 265–276. doi: 10.1515/edu-2020-0129
- Anum, A. K. (2007). *Educational reforms in Ghana, 1974-2007*. Available online at: <http://www.GhanaWebBlogEducationalReformsinGhana,1974-2007.htm> (accessed November 15, 2021).
- Apau, S. K. (2021). Teachers concerns about the implementation of the standards-based curriculum in Ghana: a case study of Effutu Municipality. *Educ. Res. Rev.* 16, 202–211. doi: 10.5897/ERR2020.4051
- Arhinful, E., and Tijani, H. (2020). *Teachers boycott GES' Curriculum Training Workshop Over Allowances*. Available online at: [https://citinewsroom.com/2020/12/a-r-teachers-boycott-ges-curriculum-training-workshop-over-allowances/#\\$sim\\$:text=A%2FR%3A%20Teachers%20boycott%20GES%27%20curriculum%20training%20workshop%20over%20allowances,-byErnest%20Arhinfulandtext\\$= \\$About%20400%20participants%20in%20the,is%20in%20its%20fourth%20day](https://citinewsroom.com/2020/12/a-r-teachers-boycott-ges-curriculum-training-workshop-over-allowances/#sim:text=A%2FR%3A%20Teachers%20boycott%20GES%27%20curriculum%20training%20workshop%20over%20allowances,-byErnest%20Arhinfulandtext$= $About%20400%20participants%20in%20the,is%20in%20its%20fourth%20day) (accessed July 12, 2021).
- Arko, D. A. (2021). How confident are social studies teachers in curriculum implementation? Understanding teachers' self-efficacy beliefs. *Am. J. Humanit. Soc. Sci. Res.* 5, 186–198. Available online at: <https://www.ajhssr.com/wp-content/uploads/2021/11/X21511186198.pdf> (accessed May 21, 2022).
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bassah, M. G. (2020). *Teachers' self-Efficacy Beliefs and Ghanaian Language Curriculum Implementation in Senior High Schools in North and South Danyi Districts, Volta Region*. Unpublished master's thesis, University of Cape Coast.
- Butucha, K. (2014). Relationships between secondary school beginning teachers' perceptions of self-efficacy and professional commitment in Ethiopia. *Int. J. Acad. Res. Prog. Educ. Dev.* 3, 79–104. doi: 10.6007/IJARPED/v3-i3/955
- Chandrika, S. D., and Varma, K. (2022). Study of self-efficacy amongst higher secondary school teachers with respect to teaching experience and gender. *J. Posit. Sch. Psychol.* 6, 195–199. Available online at: <https://journalppw.com/index.php/jpsp/article/view/1162> (accessed May 23, 2022).
- Cobbold, C. (2017). Moving from page to playground: the challenges and constraints of implementing curriculum in Ghana. *Res. Humanit. Soc. Sci.* 7, 1–11. Available online at: <http://hdl.handle.net/123456789/5425> (accessed May 4, 2022).
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Lawrence Erlbaum Associates, Publishers.
- Dibapile, W. T. S. (2012). *Teacher Efficacy and Classroom Management Among Botswana Junior Secondary School Teachers*. Unpublished doctoral dissertation, University of Tennessee, Knoxville, USA.
- Evers, W. J. G., Brouwers, A., and Tomic, W. (2002). Burnout and self-efficacy: a study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *Br. J. Educ. Psychol.* 72, 227–243. doi: 10.1348/000709902158865
- Fullan, M. (1994). "Implementation of innovations," in *The International Encyclopedia of Education*, 2nd Edn., eds T. Husen and T. N. Postlethwaite (Oxford: Pergamon Press), 2839–2847.
- George, S. V., Richardson, P. W., and Watt, H. M. (2018). Early career teachers' self-efficacy: a longitudinal study from Australia. *Aust. J. Educ.* 62, 217–233. doi: 10.1177/0004944118779601
- Ghaith, G., and Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes towards the implementation of instructional innovation. *Teach. Teach. Educ.* 13, 451–458. doi: 10.1016/S0742-051X(96)00045-5
- Graham, L. J., White, S. L., Cologon, K., and Pianta, R. C. (2020). Do teachers' years of experience make a difference in the quality of teaching? *Teach. Teach. Educ.* 96, 103190. doi: 10.1016/j.tate.2020.103190
- Gurbuzturk, O., and Sad, S. N. (2009). Student teachers' beliefs about teaching and their sense of self-efficacy: a descriptive and comparative analysis. *Inonu*

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.

The reviewer IA declared a shared affiliation with the authors EA, FA, JF, and JH to the handling editor at the time of review.

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.

- University J Faculty Educ. 10, 201–226. Available online at: <https://dergipark.org.tr/en/pub/inuefd/issue/8704/108685> (accessed March 23, 2022).
- Hair, J., Black, W., Babin, B., and Anderson, R. (2010). *Multivariate Data Analysis, 7th Edn.* Pearson Upper Saddle River.
- Haverback, H. R., and Parault, S. J. (2008). Pre-service reading teacher efficacy and tutoring: a review. *Educ. Psychol. Rev.* 20, 237–255. doi: 10.1007/s10648-008-9077-4
- Hodges, C. B., Gale, J., and Meng, A. (2016). Teacher self-efficacy during the implementation of a problem-based science curriculum. *Contemp. Issues Technol. Teach. Educ.* 16, 434–451. Available online at: <https://www.learnedlib.org/primary/p/151934/> (accessed March 29, 2022).
- Holzberger, D., Philip, A., and Kunter, M. (2012). How teachers' self-efficacy is related to instructional quality: a longitudinal analysis. *J. Educ. Psychol.* 105, 774–786. doi: 10.1037/a0032198
- Institute of Medicine (2001). *Exploring the Biological Contributions to Human Health: Does Sex Matter?* Washington, DC: National Academies Press.
- Irvine, J. (2019). Relationship between teaching experience and teacher effectiveness: implications for policy decisions. *J. Instr. Pedagogies*, 22, 1–19. Available online at: <https://eric.ed.gov/?id=EJ1216895> (accessed April 13, 2022).
- Jameson-Charles, M., and Jaggernauth, S. (2015). An investigation of the influence of teacher variables on pre-training efficacy beliefs. *Caribb. Teach. Scholar* 5, 5–24. Available online at: <https://journals.sta.uwi.edu/ojs/index.php/cts/article/view/523> (accessed April 15, 2022).
- Jumanne, J. (2012). *Teacher Self-Efficacy in Teaching Science Process Skills: An Experience of Biology Teachers in Morogoro, Tanzania.* Tanzania: LAP LAMBERT Academic Publishing.
- Kabaoglu, K. (2015). *Predictors of Curriculum Implementation Level in Elementary Mathematics Education: Mathematics-Related Beliefs and Teacher Self-Efficacy Beliefs.* Unpublished master's thesis, Graduate School of Social Sciences, Middle East Technical University.
- Karimvand, P. N. (2011). The nexus between Iranian EFL teachers' self-efficacy, teaching experience and gender. *Engl. Lang. Teach.* 4, 171–183. doi: 10.5539/elt.v4n3p171
- Klassen, R., Tze, V., Betts, S., and Gordon, K. (2011). Teacher efficacy research 1998–2009: signs of progress or unfulfilled promise? *Educ. Psychol. Rev.* 23, 21–43. doi: 10.1007/s10648-010-9141-8
- Klassen, R. M., and Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: a meta-analysis. *Educ. Res. Rev.* 12, 59–76. doi: 10.1016/j.edurev.2014.06.001
- Kpedator, E. Y. (2019). *Introduction of a New Standards-Based Curriculum: Are We Ready?* Available online at: <https://www.modernghana.com/news/955689/introduction-of-a-new-standards-based-curriculum.html> (accessed April 27, 2022).
- Kuyini, A. B., Desai, I., and Sharma, U. (2020). Teachers' self-efficacy beliefs, attitudes and concerns about implementing inclusive education in Ghana. *Int. J. Incl. Educ.* 24, 1509–1526. doi: 10.1080/13603116.2018.1544298
- Marsh, C. J. (1997). *Key Concepts for Understanding Curriculum 2: Planning, Management and Ideology.* London: Falmer Press.
- Mpuangnan, K. N., and Adusei, O. (2021). Implementation of standards-based curriculum in Ghana: the concerns of basic school teachers. *Int. J. Educ. Res.* 9, 53–66. Available online at: <https://ijern.com/journal/2021/March-2021/04.pdf> (accessed April 19, 2022).
- National Council for Curriculum and Assessment [NaCCA]. (2019). *Teacher Resource Pack: Primary 1–6.* Accra: MoE, NaCCA.
- Nugba, M. R., Quansah, F., Ankamah, F., Tsey, E. E., and Ankoma-Sey, V. R. (2021). Prevailing status of educational process factors in junior high schools in the Central Region of Ghana. *Asian J. Educ. Soc. Stud.* 24, 42–53. doi: 10.9734/ajess/2021/v24i230575
- Odanga, S. J. O., Raburu, P. A., and Aloka, P. J. O. (2015). Influence of gender on teachers' self-efficacy in secondary schools of Kisumu County, Kenya. *Acad. J. Interdiscipl. Stud.* 4, 189–198. doi: 10.5901/ajis.2015.v4n3p189
- Pallant, J. (2020). *SPSS Survival Manual. A Step-by-Step Guide to Data Analysis Using IBM SPSS, 7th Edn.* London: Routledge.
- Pfytzner-Eden, F. (2016). Why do I feel more confident? Bandura's sources predict preservice teachers' latent changes in teacher self-efficacy. *Front. Psychol.* 7, 1486. doi: 10.3389/fpsyg.2016.01486
- Quainoo, E. A., Quansah, F., Adams, F., and Opoku, E. (2020). Comparative review of educational reforms, policies and systems: a case of China and Ghana. *J. Comp. Stud. Int. Educ.* 2, 5–21. Available online at: <https://www.jcsie.com/ojs/dir/index.php/JCSIE/article/view/20>
- Quansah, F. (2017). The use of Cronbach alpha reliability estimate in research among students in public universities in Ghana. *Afr. J. Teach. Educ.* 6, 56–64. doi: 10.21083/ajote.v6i1.3970
- Quansah, F. (2021). "Teaching and learning of mathematics in lower primary in basic schools in Ghana," in *Developing Mathematical Literacy in the Context of the Fourth Industrial Revolution*, eds E. O. C. Espinosa (Hershey: IGI Global Publishers), 136–154.
- Quansah, F., and Ankoma-Sey, V. R. (2020). Evaluation of pre-service education programme in terms of educational assessment. *Int. J. Res. Teach. Educ.* 11, 56–69. Available online at: <https://ijrte.penpublishing.net/makale/1385> (accessed November 15, 2021).
- Rezaeian, S., and Abdollahzadeh, E. (2020). Teacher efficacy and its correlates in the EFL context of Iran: the role of age, experience, and gender. *Int. Online J. Educ. Teach.* 7, 1533–1548. Available online at: <https://eric.ed.gov/?id=EJ1271075> (accessed March 18, 2022).
- Roehrig, G. H., Kruse, A. K., and Kern, A. (2007). Teacher and school characteristics and their influence on curriculum implementation. *J. Res. Sci. Teach.* 44, 883–907. doi: 10.1002/tea.20180
- Shaukat, S., and Iqbal, H. M. (2012). Teacher self-efficacy as a function of student engagement, instructional strategies and classroom management. *Pakistan J. Soc. Clin. Psychol.* 10, 82–85. Available online at: <https://gcu.edu.pk/pages/gcupress/pjsoc/volumes/pjsoc2012july-13.pdf> (accessed March 27, 2022).
- Short, S. E., Yang, Y. C., and Jenkins, T. M. (2013). Sex, gender, genetics, and health. *Am J Public Health.* 103, S93–S101.
- Smith, A. B. (1996). The early childhood curriculum from a sociocultural perspective. *Early Child Dev. Care* 115, 51–64. doi: 10.1080/030044396150105
- Snyder, J., Bolin, F., and Zumwalt, K. (1992). "Curriculum implementation," in *Handbook of Research on Curriculum*, ed P. W. Jackson (New York, NY: Macmillan), 402–433.
- Snyder, S., and Fisk, T. (2016). Applying Bandura's model to identifying sources of self-efficacy of teaching artists. *Res. Schools* 23, 38–50. Available online at: <https://eric.ed.gov/?id=EJ1134838> (accessed March 23, 2022).
- Susilanas, R., Asra, A., and Herlina, H. (2018). The contribution of the self-efficacy of curriculum development team and curriculum document quality to the implementation of diversified curriculum in Indonesia. *MOJES* 2, 31–40. Available online at: <https://mojes.um.edu.my/index.php/MOJES/article/view/12838> (accessed March 21, 2022).
- Tabak, R. S., Akyildiz, N., and Yildiz, S. (2003). Teachers' self-efficacy perception levels and environment awareness. *Egitim Arastirmalari* 10, 134–145.
- Tschannen-Moran, M., and Woolfolk Hoy, A. W. (2001). Teacher efficacy: Capturing and elusive construct. *Teach. Teacher Educ.* 17, 783–805. doi: 10.1016/S0742-051X(01)00036-1
- Tschannen-Moran, M., and Woolfolk Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Rev. Educ. Res.* 6, 202–248. doi: 10.3102/00346543068002202
- Wang'eri, T., and Otanga, H. (2014). Sources of personal teachers' efficacy and influence on teaching methods among teachers in primary schools in Coast Province, Kenya. *Global J. Interdiscipl. Soc. Sci.* 3, 190–195. Available online at: <https://www.walshmedicalmedia.com/open-access/sources-of-personal-teacher-efficacy-and-influence-on-teaching-methods-among-teachers-in-primary-schools-in-coast-provin.pdf> (accessed March 21, 2022).
- Weisel, A., and Dror, O. (2006). School climate, sense of efficacy and Israeli teachers' attitudes toward inclusion of students with special needs. *Educ. Citizenship Soc. Justice* 1, 157–174. doi: 10.1177/1746197906064677
- Yusuf, C. M., and Nor, M. M. (2017). Level of teacher's self-efficacy based on gender, teaching experience and teacher training. *Adv. Sci. Lett.* 23, 2119–2122. doi: 10.1166/asl.2017.8573
- Zee, M., and Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: a synthesis of 40 years of research. *Rev. Educ. Res.* 86, 981–1015. doi: 10.3102/0034654315626801



OPEN ACCESS

EDITED BY

Claudio Longobardi,
University of Turin,
Italy

REVIEWED BY

Efthymios Lampridis,
Democritus University of Thrace, Greece
Quan-Hoang Vuong,
Phenikaa University, Vietnam

*CORRESPONDENCE

Eunice Pui-Yu Yim
euniceyim@gmail.com

SPECIALTY SECTION

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

RECEIVED 26 March 2022

ACCEPTED 27 September 2022

PUBLISHED 17 October 2022

CITATION

Yim EP-Y (2022) Effects of Asian cultural
values on parenting style and young
children's perceived competence: A cross-
sectional study.
Front. Psychol. 13:905093.
doi: 10.3389/fpsyg.2022.905093

COPYRIGHT

© 2022 Yim. 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.

Effects of Asian cultural values on parenting style and young children's perceived competence: A cross-sectional study

Eunice Pui-Yu Yim*

School of Education and Languages, Hong Kong Metropolitan University, Kowloon, Hong Kong SAR, China

Authoritarian parenting has long been associated with Western individualism and improved child development. This study examined the relationship between cultural values, parenting styles, and children's perceived competence in Hong Kong. A total of 48 parents from local Chinese families, 49 parents from South Asian families, and 105 children (24 local Chinese and 81 South Asian) aged 5–6 years participated in the study. Self-report questionnaires on adherence to Asian cultural values and parenting style were administered to parents. The Pictorial Survey on Children's Perceived Competence was administered to children by trained research assistants. The results contradicted two long-standing assumptions on Asian cultural values and parenting styles. First, higher adherence to Asian cultural values increased the likelihood of having an authoritarian parenting style. Second, authoritative parenting practices were more likely to be associated with improved social-emotional competence in children. Multiple regression analyses revealed a strong positive correlation between Asian cultural values and authoritative parenting style ($R^2=0.597$). There was no association between parenting style and the development of competence in young children. However, a positive correlation was found between Asian cultural values and young children's perceived competence. This study showed that components of collectivism and humility in Asian cultural values could have functional values that are essential for developing competencies in South Asian young children but not in local Chinese young children. This study discussed the implications of cultural values in the terms of contextualization, functional relevance of cultural values for ethnic minorities, and ideal parenting practices.

KEYWORDS

parenting style, cultural values, perceived competence, child development, immigrant

Introduction

In recent decades, the ethnic diversity in Hong Kong has expanded (Fleming, 2019). In 2016, Hong Kong had a population of approximately 7.34 million, including 8% of the population that is non-Chinese compared to the 6% of non-Chinese in 2011 (Census and Statistics Department, 2011, 2016). South Asian ethnic groups, such as Indonesians, Nepalese, Pakistanis, and Filipinos, comprised 71% of the non-ethnic Chinese population in Hong Kong (Census and Statistics Department, 2016). South Asians and local Hong Kong citizens are categorized as Asians and assumed to have a collectivist cultural orientation (Fleming, 2019). However, there are variations in the different cultural sub-groups across the broader Asian cultures, which are reflected in every aspect of life, including parenting practices.

Bronfenbrenner's (1979) ecological systems model of development proposes that an individual is nested within interrelated social systems at different levels, ranging from the micro level, (referring to the cultural context in the home) to the macro level (the broader socio-cultural context). Childrearing is inevitably affected by the occurrences within these social levels and vice versa (Bronfenbrenner, 1986). Parenting comprises various practices that reflect the cultural values within a child's home that could differ from the broader local culture in which they live (Bornstein, 2012; Hosken et al., 2019).

Acculturation theory contends that the smooth transition across contexts and positive acculturation with the host culture is required among immigrants (Berry, 1980; Barker and Cornwell, 2019). The smooth transition across contexts is indicative of having knowledge of the social expectations of the involved contexts and behaving in accordance to these expectations (Schwartz et al., 2010; Li, 2013). Given this, an individual may engage himself or herself in forced adaptation, not yet positive acculturation, in order to obtain smooth transitions. Acculturation theory suggests that forced adaptation and positive acculturation represent two ends of a continuum (Berry, 1980; Barker, 2015). New immigrants frequently engage in "forced" adaptation wherein they conform to local cultural values at the early stage of the adaptation process where as positive acculturation refers when individuals can achieve a balance between host (macro-level) and home (micro-level) cultural values in which they exercise these values appropriately and adaptively without devaluing either of the cultural values (Schwartz et al., 2010). In other words, positive acculturated individuals can exercise more than one set of cultural values over time and within or across contexts in daily living smoothly and competently. Considering this, exploring the dynamics between Asian cultural values on parenting behaviors would provide insights into the heterogeneity of Asian cultural values in Asian sub-cultural groups and their relationship to parental behaviors, which in turn influence children's perceived competence.

Cultural studies have identified six key elements typically associated with an authoritarian parenting style that most Asian ethnic groups having a collectivist cultural orientation share

(Kim et al., 1999; Ng and Wang., 2019). These elements are: (1) conformity to social norms (i.e., recognition and adherence to social expectations, norms, and practices); (2) family recognition through achievements (i.e., a reflection of an individual's achievements and family status); (3) emotional self-control (i.e., behaving appropriately rather than acting on how one is feeling); (4) collectivism (i.e., placing group welfare before individual welfare); (5) humility (i.e., being humble without boasting about one's achievement); and (6) filial piety (i.e., being respectful toward one's living parents, ascertaining the rules of culturally defined intergenerational relationships, and placing familial needs before individual interests). Among the four parenting styles proposed by Diana Baumrind (1978), the authoritarian parenting style is characterized by high control and low warmth in which parents employ strict discipline, are insensitive to the child's emotional needs (Ang and Goh, 2006). Parenting behaviors of this style include expectations of obedience and compliance to parental demands with less responsiveness toward their children's needs (Rudy and Grusec, 2006). Existing research suggests that parents in Asian societies generally demand high obedience levels from their children and interact with family and community members through rules and orders based on social hierarchy (Rudy and Grusec, 2006; Dwairy and Achoui, 2010). On the other hand, while permissive parenting style cares about their children's needs but have difficulty setting limits in children's disciplines, the authoritative parenting style is characterized by high control and high warmth with high parental affection and responsiveness to children's needs as well as high expectations with increased independence and self-control (Rudy and Grusec, 2006). These characteristics are generally reflected individualism-oriented culture which is characterized by open discussions and interpersonal equality dominating the Western culture (Rudy and Grusec, 2006). It is considered the ideal parenting style to raise competent children, whereas the authoritarian style is more likely to raise less competent children (Sorkhabi, 2005; Xu et al., 2005; Chen-Bouck et al., 2019).

Chao (1994) explained the discrepancies between Western and Chinese cultures in their understanding of "control" and "love." Particularly, the meaning of "guan" (to control) in the Chinese culture includes the meaning of responsiveness to children's needs has been misinterpreted by non-ethnic Chinese cultural groups. "Guan" means "to govern" but also means "to care for and to love." Components of love and responsiveness are embedded in the notion of "control" in Chinese parenting. The underlying meaning of "control" for them is to protect children from harm and help them succeed (Ang and Goh, 2006; Mousavi and Juhari, 2019). Therefore, Chinese parents, view "control" as an act of parental love, care, involvement, concern, and support for children's needs rather than the negative connotations suggested in existing family studies (Chao, 1994; Sorkhabi, 2005; Xu et al., 2005).

Other studies on the association between parenting style and children's positive outcomes highlighted the need to contextualize cultural values with consideration of children's developmental

stage, engaged cultural contexts and meanings to the families and their children (Alegre, 2011; Huang and Gove, 2015; Hung, 2018; Olla et al., 2018). Studies demonstrated that Chinese youth living in authoritarian families performed better in school than those in authoritative families (Alegre, 2011; Hung, 2018). Other studies have investigated the association between authoritarian parenting in Asian families and their children's development (Huang and Gove, 2015; Hung, 2018; Olla et al., 2018). These studies suggested that parents had their own values and expectations about how they, as parents, should behave to raise competent children. Many of these parents described competence in terms of academic achievement (Huang and Gove, 2015; Hung, 2018; Olla et al., 2018). On the other hand, studies on authoritative parenting behaviors found that these behaviors significantly contributed to improved school grades for non-Chinese, European-Americans youths, and school-aged children more than the first-generation Chinese youth and school-aged children (Pong et al., 2010; Antony-Newman, 2019). These findings suggested that the characteristics of authoritative parenting may not exert the same effects on families from different cultural backgrounds. Adolescents and young children are at different developmental stages. Adolescents experience more academic pressure for school achievement than young children, while young children experience challenges related to socio-emotional development such as acceptance from others, including parents, caregivers, and peers. Besides focusing on academic achievements and decontextualizing the interpretations of Asian cultural values, the existing studies focus on the influence of Asian cultural values on parenting while targeting Asian families with adolescents in different countries instead of young children from different sub-cultural group in the same geographic region (Chao, 2001; Pong et al., 2005; Shoshani and Steinmetz, 2014) and measuring developmental outcomes in adolescents in terms of academic achievement (Chao and Kaochinda, 2010; Lin et al., 2021). Thus, assessments of young children's development should focus on social-emotional development such as perceived self-competence in daily life which is affected by the children's involved social systems (Brock et al., 1998; Huber et al., 2019).

Bronfenbrenner's ecological and bioecological model of development provides a theoretical framework that describes the interconnectedness across five major social systems and their influence on personal development from an individual, societal, and organizational level (Bronfenbrenner, 1986; Bronfenbrenner and Morris, 1998). These five systems include: (1) the *microsystem*, which is the context that is closest to an individual's life with the most direct interactions (e.g., family, school, religious institutions, neighborhood, and peers); (2) the *mesosystem*, referring to the interconnections between the different elements in an individual's microsystems, (e.g., interactions between the family and school or children's peers and the family); (3) the *exosystem*, representing the links between social settings in which an individual is not actively involved (e.g., a child's experience at home may be influenced by parental experiences at work, subsequently affecting parent-child interactions); (4) the *macrosystem*,

comprising the overarching societal culture (e.g., socioeconomic status, ethnicity, and laws and rules binding individuals); and (5) the *chronosystem*, which refers to how the person and environments change over time (Christensen, 2016; Xia et al., 2020). According to this model, instead of living within five separate social systems, individuals are nested within the five social systems that are embedded within one another. The five social systems are embedded such that individuals are affected by the progressively complex reciprocal interactions between an individual and his or her environment regularly over extended periods (Xia et al., 2020). The model was later enhanced by introducing Process–Person–Context–Time Model (PPCT) to highlight on the important role of time and developmental needs across the life span on an individual development (Bronfenbrenner, 1992; Xia et al., 2020).

Proximal processes (*Process*) are the reciprocal interactions between a developing human being and one or more of the persons, objects, and symbols in his or her immediate environment (Bronfenbrenner, 1992). These reciprocal interactions are expected to become “progressively more complex” with the expanded social network and more complicated life events (Bronfenbrenner and Morris, 1998). *Person* characteristics include personality traits that can foster or sustain proximal processes, or interfere with or prevent their occurrence. The sustainability of the proximal process can be further supported by those biological, mental, or experiential resources, such as physical attributes, and reading abilities, that individuals bring to the development process. Contextual demands that may pose barriers to the development and proximal processes in the individuals vary from time to time, such as the difficulties of tasks and whether the demands posed on the individual are beyond the individual's capabilities and resources available at the time. *Context* refers to the socio-cultural environment, including digital world where the individual engages. Johnson and Pupilampu (2008) proposed an ecological techno-subsystem highlighting the significant impacts of the availability of technologies in an individual's microsystem that could bring socio-cultural influences from virtual contexts in which the individual is not currently physically engaged. In other words, individuals are exposed to multiple cultures and norms rather than being confined to a single culture or norm (Johnson and Pupilampu, 2008; Schwartz et al., 2010). Regarding the last component of *Time*, the model suggested that development and reciprocal interaction can occur within a shorter period during a proximal process (microtime) or span over a longer period of days to months (mesotime) or decades within and across generations (macrotime; Bronfenbrenner, 1986; Bronfenbrenner and Morris, 1998).

While the refined bioecological model with the PCCT model and ecological techno-subsystem provided a general overview of how different factors could synergistically impact human development over extended periods, it fails to explain how an individual is affected by the socio-cultural context it is engaged in at a specific developmental stage. In some stages of development, some factors may exert more influence on the individual's

development than the others. Different competence weighs differently in the different developmental stages. In childhood, the socio-emotional competence is considered more important than academic related competences, which are weighted heavier in school-aged and later stages of development. In this study, young children of 5 years old who are in the early stage of development (*Time*) and they primarily engage themselves in home and school (*Context*). Young children generally spend considerable less amount of time on internet than older children (Hinkley et al., 2012; Fakhouri et al., 2013). The reciprocal interactions between the children (*Person*) and their major caregiver (s) are the major proximal processes (*Process*) to be measured.

Review of the existing research investigating associations between parenting styles and cultural values have the following limitations. First, the existing literature investigating the parenting styles of Asian sub-cultural groups mainly compared families of inter-Asian countries (Ang and Goh, 2006; Chuang et al., 2018). Second, studies investigating the effects of parenting styles on children's development mainly focus on adolescents or school-aged children than preschoolers as study samples with research emphasis on academic-related skills for school readiness rather than their socio-emotional development (Wong et al., 2018; Xia, 2020). Third, although there are studies on Hong Kong-Chinese and sub-Asian children in Hong Kong, these studies focus on parental expectations and satisfaction instead of the effects of their parenting on the socio-emotional development of young children (Chan et al., 2009; Chan and Li, 2020; Lau and Power, 2020). The limitations in the existing literature suggest a need to explore the differences in parenting practices among Asian sub-cultural groups in within the same geographical region like Hong Kong (Chuang et al., 2018). Hong Kong, previously a British colony, has its Chinese cultural origins that can be discerned along with influence of the Western cultures. Interpretations of Asian cultural values among local Hong Kong residents are not equivalent to those who immigrated to Hong Kong from neighboring Asian countries (Hue and Kennedy, 2014). With an increasing number of South Asian families are migrating to Hong Kong with young children., the intergroup differences between local Chinese and South Asian cultural groups must be investigated to understand how adherence to Asian cultural values contributes toward parenting styles and the influence of parenting style on children's development within these groups (Yuen, 2016). The study investigates the extent to which these sub-cultural groups adhere to Asian cultural values and how these values were related to their choice of parenting practices which in turn contribute to children's well-being. This study aimed to investigate (1) the intergroup differences between local Chinese parents and South Asian parents in their adherence to Asian cultural values and their parenting style; (2) the relationship between Asian cultural values and parenting styles; (3) the dynamics between the specific components of Asian cultural values and young children's perceived competence; and (4) relationship between parenting style and children's perceived competence.

Materials and methods

Participants

Local kindergartens with South Asian families in districts with households of lower socio-economic status were invited to participate in the study through open recruitment. Three kindergartens agreed to participate. Children aged between 5 and 6 years and their parents were recruited from the participating kindergartens. The average age of parents was 33.2 years old, ranging from 21 to 33 years old. One kindergarten school had over 90% South Asian children and 10% local Chinese children, while the second one had 50% South Asians and 50% local Chinese children, and the third had 90% local Chinese children with 10% South Asian children. The South Asian ethnic groups included Indonesians (33%), Nepalese (11%), Pakistanis (34%), and Filipinos (22%). Due to the limited sample size of South Asian samples, this study used South Asian as the collective term for the South Asian ethnic groups and local Chinese for Hong Kong local ethnic group.

According to the school principals and teachers of participating schools, the South Asian parents had settled in Hong Kong for over 5 years and could read and speak English, with their children being born and raised in Hong Kong. No formal assessment was administered to determine children's language proficiency in Cantonese. These South Asian children started their full-day schooling in the participating schools when they were 3 years old. After 2 years of schooling in local kindergartens, these children were proficient in Cantonese for daily communication. Major caregivers for the children were invited to participate in the study. All individuals participating voluntarily in the study comprised 97 parents (49 mothers from South Asian group; 48 mothers from local Chinese groups), and 105 children (24 local Chinese, 81 South Asian).

Materials

This study employed three scales to assess an individual's adherence to Asian values, parenting style, and children's perceived competence. These scales were: 1) the Asian Values Scale, 2) the Parenting Practice Questionnaire, and 3) the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (PSPCSA) Preschool-kindergarten (both the boys and girls versions).

Asian values scale

The Asian Values Scale was a self-administered questionnaire developed by Kim et al. (1999) to assess parents' adherence to Asian cultural values. The scale's 24 items were rated and scored on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*) and measured an individual's adherence to Asian cultural values. The six domains included: *conformity to norms* (eight items), *family recognition through achievement* (three items),

emotional self-control (three items), *collectivism* (three items), *humility* (three items), and *filial piety* (four items). The scale has been used in several countries (including America, Australia, Philippines and South Korea) and has demonstrated good reliability and validity, with Cronbach's alphas ranging from 0.72 to 0.80 (Kim et al., 2005; Shim and Schwartz, 2008; Oei and Raylu, 2009; Magno, 2010; Oh and Lee, 2014). Examples of items in the questionnaire include: one should consider the needs of others before considering one's own needs; one should not deviate from familial and social norms (Kim et al., 1999).

The parenting practices questionnaire

The Parenting Practices Questionnaire was a self-administered survey developed by Robinson et al. (1995) to assess parents' parenting practices. The scale was with 62-items scored on a five-point rating scale (1 = never; 5 = always) to assess parents' adherence to three major parenting styles: *authoritative* (27 items), *authoritarian* (20 items), and *permissive* (15 items). This questionnaire excluded items related to the rejecting/neglecting parenting style based on the assumption that participating parents consented to participate in the study willingly and were actively involved in their children's activities (Robinson et al., 1995). The scale has been tested in various countries (Utah, Canada, Iran and Italy) with good reliability and validity (Robinson et al., 1996; Maria et al., 2013; Delvecchio et al., 2020; Mojdehi et al., 2020). Example items in the questionnaire include: "I praise my child when he/she is good"; "I argue with my child," "I use threats such as punishment with little or no justification" (Robinson et al., 1995).

Pictorial scale of perceived competence and acceptance for young children preschool-kindergarten (male and female versions)

The PSPCSA was used to assess children's perceived competence. Harter and Pike (1984) developed the 24-item scale with pictures covering the domains of cognitive competence (six items), physical competence (six items), peer acceptance (six items), and maternal acceptance (6 items). The scale has been tested in various countries showing good reliability and validity (Fantuzzo et al., 1996; Venetsano et al., 2018; Heritage et al., 2020). Each item has two pictures: one image represents a child who can perform the task, and the second image denotes a child who cannot perform the same task. Children are asked to indicate the picture that was the closest or most similar to them when performing the illustrated task. If the picture with a capable child is selected, the children are asked to indicate the extent of similarity between the two choices. Examples of items in the survey include: "This girl is not very good at counting. Are you?"; "The boy is not very good at climbing. Are you?" (Harter and Pike, 1984). Children are asked to make a judgment about the level of similarity between the illustrated picture and their self-perception on the same task.

Procedures

The major caregiver of each child was invited to complete two sets of self-administered questionnaires. The children were interviewed in their classrooms during their free time by trained research assistants using the PSPCSA. The pictorial questionnaire for children took about 15 min to complete.

Before administering the questionnaires, participating parents and their children provided informed consent. The participants were informed about the study's purpose, anonymity, and their right to withdraw from the study at any time. The completed questionnaire responses were entered into SPSS for analysis (Field, 2013).

Data analysis

Responses to the survey items were coded such that higher scores reflected more positive responses to examine the relationships between Asian cultural values and parenting styles, relationships between Asian cultural values and children's perceived competence, and the relationship between parenting style and children's perceived competence. Multiple linear regression was found appropriate for this study as it allows relationships between variables to be constructed and explored (Draper and Smith, 1998).

Results

Reliabilities of scales

The reliabilities of the scales were good. For the Parenting Practices Questionnaire, the Cronbach's alphas of authoritative, authoritarian, and permissive subscales in the parenting scale are 0.902, 0.743, and 0.616, respectively. For the Asian Values Scale, the Cronbach's alphas for subscale family recognition through achievement, collectivism, and humility were 0.772, 0.688, and 0.649, respectively. The internal consistencies of subscale conformity to norms, emotional self-control and filial piety were low, with 0.454, 0.401, and 0.307, respectively. Therefore, these three subscales were removed in subsequent data analyses. For PSPCSA, the Cronbach's alphas of cognitive competence, physical competence, peer acceptance, maternal acceptance are 0.615, 0.448, 0.526, and 0.561, respectively.

Intergroup differences in Asian cultural values and parenting style adoption

The study shows that South Asian parents maintained higher adherence to Asian cultural values ($R^2 = 0.597$) than local Chinese, indicating a strong positive correlation with authoritative parenting style. The regression models of South Asian parents and the other two parenting styles, namely authoritarian ($R^2 = 0.145$)

TABLE 1 Multiple linear regression of Asian cultural values adherence on parenting style for south Asians parents (N=49).

Model	B	SE	β	t	R ²	F	df	p
Dependent variable: Parent – Authoritative					0.597	10.134	6	0.000
AVS–Conformity	0.179	0.121	0.168	1.481				0.146
AVS–Achievement	−0.010	0.055	−0.021	−0.179				0.859
AVS–Emotion	0.151	0.066	0.281	2.292				0.027
AVS–Collectivism	0.064	0.069	0.121	0.938				0.354
AVS–Humility	0.310	0.069	0.490	4.483				0.000
AVS–Filial Piety	−0.023	0.067	−0.038	−0.344				0.733
Dependent variable: Parent – Authoritarian					0.145	1.156	6	0.348
AVS–Conformity	0.037	0.173	0.035	0.212				0.833
AVS–Achievement	0.012	0.078	0.025	0.148				0.883
AVS–Emotion	−0.025	0.094	−0.047	−0.262				0.794
AVS – Collectivism	−0.107	0.098	−0.205	−1.088				0.283
AVS – Humility	0.097	0.099	0.157	0.982				0.332
AVS – Filial Piety	−0.226	0.095	−0.383	−2.380				0.002
Dependent variable: Parent – Permissive					0.268	2.498	6	0.037
AVS – Conformity	0.077	0.141	0.083	0.545				0.589
AVS – Achievement	−0.055	0.064	−0.134	−0.856				0.397
AVS – Emotion	−0.134	0.076	−0.291	−1.756				0.086
AVS – Collectivism	0.000	0.080	−0.001	−0.005				0.996
AVS – Humility	−0.015	0.080	−0.027	−0.184				0.855
AVS – Filial Piety	−0.235	0.077	−0.453	−3.038				0.004

AVS, Asian values scales.

and permissive ($R^2=0.268$) failed to obtain satisfactory R^2 scores, indicating the absence of a meaningful relationship between these two parenting styles and Asian cultural values. For local-Chinese parents, no regression models had higher R^2 scores, as there were no meaningful relationships between any form of parenting style and their Asian cultural values. Regarding the influences of Asian cultural values on parenting style adoption, in South Asian parent group, authoritative parenting style had stronger positive correlation with *humility* ($\beta=0.490$) than *collectivism* ($\beta=0.121$). Authoritarian parenting style had negative correlation with *collectivism* ($\beta=-0.205$). Permissive parenting style had very weak negative correlations with the measured values including *achievement* ($\beta=-0.134$), *humility* ($\beta=-0.027$) and *collectivism* ($\beta=-0.001$). For local Chinese parents, authoritative parenting style had negative correlation with *achievement* ($\beta=-0.316$) whereas had the weakest correlation with *humility* ($\beta=0.013$). Authoritarian parenting style had positive correlations with *achievement* ($\beta=0.315$) and *collectivism* ($\beta=0.162$). Permissive parenting style had weak correlations with measured values including *achievement* ($\beta=0.266$), *humility* ($\beta=-0.143$) and *collectivism* ($\beta=0.092$) (Tables 1–3).

Asian cultural values on children's perceived competence

The correlation between Asian cultural values and perceived competence among children from local Chinese

parents was significantly stronger than among South Asian parents. Specifically, the association between Asian cultural values and *peer* competence ($R^2=0.643$) was significantly stronger among local Chinese than their South Asian counterparts ($R^2=0.097$).

This study shows that all six domains of Asian cultural values were positively related to the four aspects (*cognitive*, *physical*, *peer*, and *maternal competence*) of children's competence in both groups. For the local Chinese group, *collectivism* ($\beta=0.848$) and *humility* ($\beta=-0.838$) have the strongest correlation with *peer* competence. A similar magnitude of these two values was also found in its strong correlation with *maternal* competence (*collectivism* $\beta=0.624$; *humility* $\beta=-0.758$). For the South Asian group, *cognitive competence* and *collectivism* ($\beta=-0.324$) had stronger correlations with the *physical* competence of children's perceived competence (Tables 4–6).

Relationship between parenting style and children's perceived competence

The three parenting styles were not associated with children's perceived competence in both groups. The regression models on parenting style and children's perceived self-competence were low ($R^2<0.5$; Table 7).

From the table above, all the parenting style models affecting children's perceived competence were lower than 0.5, indicating the lack of a relationship between them. The models also showed

TABLE 2 Multiple linear regression of Asian cultural values adherence on parenting style for local-Chinese parents ($N=48$).

Model	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>R</i> ²	<i>F</i>	<i>df</i>	<i>p</i>
Dependent variable: Parent					0.275	2.596	6	0.032
– Authoritative								
AVS – Conformity	0.054	0.106	0.082	0.506				0.616
AVS – Achievement	−0.099	0.047	−0.316	−2.090				0.043
AVS – Emotion	0.090	0.073	0.190	1.236				0.223
AVS – Collectivism	0.035	0.076	0.068	0.463				0.646
AVS – Humility	0.010	0.110	0.013	0.089				0.930
AVS – Filial piety	0.127	0.101	0.196	1.253				0.217
Dependent variable: Parent					0.17	1.398	6	0.239
– Authoritarian								
AVS – Conformity	−0.033	0.122	−0.048	−0.273				0.786
AVS – Achievement	0.106	0.054	0.315	1.948				0.058
AVS – Emotion	−0.036	0.083	−0.071	−0.434				0.667
AVS – Collectivism	0.090	0.087	0.162	1.036				0.306
AVS – Humility	0.095	0.127	0.122	0.751				0.457
AVS – Filial piety	−0.059	0.116	−0.085	−0.508				0.614
Dependent variable: Parent					0.125	0.973	6	0.456
– Permissive								
AVS – Conformity	−0.101	0.104	−0.174	−0.973				0.336
AVS – Achievement	0.074	0.046	0.266	1.602				0.117
AVS – Emotion	0.059	0.071	0.142	0.838				0.407
AVS – Collectivism	0.042	0.074	0.092	0.575				0.569
AVS – Humility	−0.092	0.108	−0.143	−0.857				0.397
AVS – Filial piety	0.019	0.099	0.034	0.197				0.844

AVS, Asian values scales.

that eliminating items with low correlation scores from the children's perceived competence factors would not improve the regression results. Notably, during the second run, local Chinese children had an even lower R^2 , even though all R^2 scores were less than 0.5.

Discussion

Intergroup differences in Asian cultural values and parenting style adoption

Authoritative parenting has long been associated with individualism dominating in the Western culture. It is considered the ideal parenting style with clear guidelines and a high level of responsiveness toward children's needs (Huang and Gove, 2015). However, our results demonstrated that South Asian families with higher adherence to Asian culture values had greater authoritative parenting behaviors than their local counterparts. This contradictory finding against the traditional belief of "Asian–authoritarian parenting style" suggesting a need to re-examine the long-standing assumptions. Particularly, parents' responsiveness to children's needs in the Asian cultural context is needed to explore how parents apply Asian cultural values to pragmatically and meaningfully satisfy their children's needs that determine

parenting style rather than attributing the authoritarian parenting style to Asian cultural values.

South Asian parents

The relationship between Asian cultural values and parenting styles was evident in South Asian parents. However, it could not be discerned among the local Chinese parents, which could be explained by the goals and functional values of Asian cultural values among South Asian parents from an ecological perspective. Smooth transitions are required between various living contexts to promote a child's well-being (Bronfenbrenner and Ceci, 1994). In this study, transition experiences between micro- and macro-context for ethnic minorities could be different from their local counterparts (Barker and Cornwell, 2019). Being ethnic minorities residing in Hong Kong, South Asian group would pay more emphasis on cultural accommodation than local Chinese group (Berry, 1980; Vu et al., 2019). With this, children being obedient and conforming to rules become particularly important to these families as these values could promote safety from a cultural clash with the mainstream and assume smoother adaptation to the host culture (Barker, 2015; Chan and Li, 2020). Hence, South Asian group would exercise the Asian cultural values of obedience and conformity as means to practice love and care to address children's needs during the initial process of cultural accommodation and later acculturation (Bornstein, 2012; Stamkou et al., 2019; Chan

TABLE 3 Correlations between Asian values and parenting style.

			Correlations					
Ethnic group			AVS - Family recognition through achievement	AVS - Collectivism	AVS - Humility	Parenting - Authoritative	Parenting - Authoritarian	Parenting - Permissive
South Asian	AVS - Family recognition through achievement	Pearson correlation	1	−0.278	−0.106	−0.135	−0.055	−0.204
		Sig. (2-tailed)		0.056	0.473	0.359	0.712	0.164
		N	48	48	48	48	48	48
	AVS - Collectivism	Pearson correlation	−0.278	1	0.342*	0.493**	−0.077	−0.003
		Sig. (2-tailed)	0.056		0.017	0.000	0.604	0.982
		N	48	48	48	48	48	48
	AVS - Humility	Pearson correlation	−0.106	0.342*	1	0.660**	0.078	−0.077
		Sig. (2-tailed)	0.473	0.017		0.000	0.594	0.599
		N	48	48	49	49	49	49
	Parenting - authoritative	Pearson correlation	−0.135	0.493**	0.660**	1	−0.005	−0.123
		Sig. (2-tailed)	0.359	0.000	0.000		0.973	0.400
		N	48	48	49	49	49	49
	Parenting - authoritarian	Pearson correlation	−0.055	−0.077	0.078	−0.005	1	0.588**
		Sig. (2-tailed)	0.712	0.604	0.594	0.973		0.000
		N	48	48	49	49	49	49
	Parenting - permissive	Pearson correlation	−0.204	−0.003	−0.077	−0.123	0.588**	1
		Sig. (2-tailed)	0.164	0.982	0.599	0.400	0.000	
		N	48	48	49	49	49	49
Local-Chinese	AVS - Family recognition through achievement	Pearson correlation	1	0.224	−0.260	−0.360*	0.340*	0.271
		Sig. (2-tailed)		0.125	0.074	0.012	0.018	0.062
		N	48	48	48	48	48	48
	AVS - Collectivism	Pearson correlation	0.224	1	0.208	0.039	0.240	0.102
		Sig. (2-tailed)	0.125		0.156	0.795	0.101	0.491
		N	48	48	48	48	48	48
	AVS - Humility	Pearson correlation	−0.260	0.208	1	0.237	0.018	−0.196
		Sig. (2-tailed)	0.074	0.156		0.104	0.903	0.182
		N	48	48	48	48	48	48
	Parenting - authoritative	Pearson correlation	−0.360*	0.039	0.237	1	−0.231	−0.213
		Sig. (2-tailed)	0.012	0.795	0.104		0.114	0.147
		N	48	48	48	48	48	48
	Parenting - authoritarian	Pearson correlation	0.340*	0.240	0.018	−0.231	1	0.287*
		Sig. (2-tailed)	0.018	0.101	0.903	0.114		0.048
		N	48	48	48	48	48	48
	Parenting - permissive	Pearson correlation	0.271	0.102	−0.196	−0.213	0.287*	1
		Sig. (2-tailed)	0.062	0.491	0.182	0.147	0.048	
		N	48	48	48	48	48	48

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

TABLE 4 Correlations between Asian values and children competence.

			Correlations									
Ethnic group			AVS - Conformity to norms	AVS - Family recognition through achievement	AVS - emotional self-control	AVS - Collectivism	AVS - Humility	AVS - Filial piety	C_Cognitive	C_Physical	C_Peer	C_Maternal
South Asian	AVS - Conformity to norms	Pearson correlation	1	0.311*	0.157	0.170	0.242	0.183	0.235	−0.008	0.122	0.018
		Sig. (2-tailed)		0.032	0.287	0.247	0.094	0.208	0.116	0.956	0.419	0.908
		N	49	48	48	48	49	49	46	46	46	46
	AVS - Family recognition through achievement	Pearson correlation	0.311*	1	−0.238	−0.278	−0.106	0.371**	0.250	0.359*	0.140	−0.052
		Sig. (2-tailed)	0.032		0.103	0.056	0.473	0.009	0.098	0.016	0.358	0.734
		N	48	48	48	48	48	48	45	45	45	45
	AVS - emotional self-control	Pearson correlation	0.157	−0.238	1	0.568**	0.297*	−0.174	−0.094	−0.046	−0.110	−0.036
		Sig. (2-tailed)	0.287	0.103		0.000	0.041	0.237	0.540	0.763	0.470	0.815
		N	48	48	48	48	48	48	45	45	45	45
	AVS - Collectivism	Pearson correlation	0.170	−0.278	0.568**	1	0.342*	−0.266	−0.302*	−0.198	−0.225	−0.093
		Sig. (2-tailed)	0.247	0.056	0.000		0.017	0.068	0.044	0.193	0.137	0.542
		N	48	48	48	48	48	48	45	45	45	45
	AVS - Humility	Pearson correlation	0.242	−0.106	0.297*	0.342*	1	−0.002	−0.106	0.007	−0.037	−0.301*
		Sig. (2-tailed)	0.094	0.473	0.041	0.017		0.988	0.485	0.964	0.809	0.042
		N	49	48	48	48	49	49	46	46	46	46
	AVS - Filial Piety	Pearson correlation	0.183	0.371**	−0.174	−0.266	−0.002	1	0.225	0.377**	0.220	0.052
		Sig. (2-tailed)	0.208	0.009	0.237	0.068	0.988		0.133	0.010	0.142	0.732
		N	49	48	48	48	49	49	46	46	46	46
	C_Cognitive	Pearson correlation	0.235	0.250	−0.094	−0.302*	−0.106	0.225	1	0.126	0.284	−0.007
		Sig. (2-tailed)	0.116	0.098	0.540	0.044	0.485	0.133		0.405	0.056	0.961
		N	46	45	45	45	46	46	46	46	46	46

(Continued)

TABLE 4 (Continued)

			Correlations									
Ethnic group			AVS - Conformity to norms	AVS - Family recognition through achievement	AVS - emotional self-control	AVS - Collectivism	AVS - Humility	AVS - Filial piety	C_Cognitive	C_Physical	C_Peer	C_Maternal
Local-Chinese	C_Physical	Pearson correlation	−0.008	0.359*	−0.046	−0.198	0.007	0.377**	0.126	1	−0.279	−0.042
		Sig. (2-tailed)	0.956	0.016	0.763	0.193	0.964	0.010	0.405		0.060	0.781
		N	46	45	45	45	46	46	46	46	46	46
	C_Peer	Pearson correlation	0.122	0.140	−0.110	−0.225	−0.037	0.220	0.284	−0.279	1	0.008
		Sig. (2-tailed)	0.419	0.358	0.470	0.137	0.809	0.142	0.056	0.060		0.955
		N	46	45	45	45	46	46	46	46	46	46
	C_Maternal	Pearson correlation	0.018	−0.052	−0.036	−0.093	−0.301*	0.052	−0.007	−0.042	0.008	1
		Sig. (2-tailed)	0.908	0.734	0.815	0.542	0.042	0.732	0.961	0.781	0.955	
		N	46	45	45	45	46	46	46	46	46	46
	AVS - Conformity to norms	Pearson correlation	1	0.116	0.396**	0.220	0.268	0.424**	−0.041	−0.264	0.231	−0.170
		Sig. (2-tailed)		0.431	0.005	0.133	0.066	0.003	0.860	0.248	0.314	0.462
		N	48	48	48	48	48	48	21	21	21	21
	AVS - Family recognition through achievement	Pearson correlation	0.116	1	−0.193	0.224	−0.260	−0.145	−0.197	−0.166	0.094	−0.165
		Sig. (2-tailed)	0.431		0.190	0.125	0.074	0.325	0.391	0.471	0.685	0.476
		N	48	48	48	48	48	48	21	21	21	21
	AVS - Emotional self-control	Pearson correlation	0.396**	−0.193	1	0.134	0.236	0.354*	−0.093	0.120	0.300	0.038
		Sig. (2-tailed)	0.005	0.190		0.365	0.107	0.014	0.690	0.604	0.187	0.872
		N	48	48	48	48	48	48	21	21	21	21
	AVS - Collectivism	Pearson correlation	0.220	0.224	0.134	1	0.208	−0.023	0.127	0.131	0.433*	0.062
		Sig. (2-tailed)	0.133	0.125	0.365		0.156	0.876	0.585	0.572	0.050	0.789
		N	48	48	48	48	48	48	21	21	21	21

(Continued)

TABLE 4 (Continued)

		Correlations										
Ethnic group			AVS - Conformity to norms	AVS - Family recognition through achievement	AVS - emotional self-control	AVS - Collectivism	AVS - Humility	AVS - Filial piety	C_Cognitive	C_Physical	C_Peer	C_Maternal
AVS - Humility	Pearson correlation		0.268	−0.260	0.236	0.208	1	0.309*	0.480*	0.343	−0.330	−0.326
	Sig. (2-tailed)		0.066	0.074	0.107	0.156		0.033	0.028	0.128	0.143	0.150
	N		48	48	48	48	48	48	21	21	21	21
AVS - Filial piety	Pearson correlation		0.424**	−0.145	0.354*	−0.023	0.309*	1	0.182	0.216	−0.120	−0.031
	Sig. (2-tailed)		0.003	0.325	0.014	0.876	0.033		0.429	0.348	0.604	0.894
	N		48	48	48	48	48	48	21	21	21	21
C_Cognitive	Pearson correlation		−0.041	−0.197	−0.093	0.127	0.480*	0.182	1	0.406	−0.349	0.108
	Sig. (2-tailed)		0.860	0.391	0.690	0.585	0.028	0.429		0.068	0.121	0.642
	N		21	21	21	21	21	21	21	21	21	21
C_Physical	Pearson correlation		−0.264	−0.166	0.120	0.131	0.343	0.216	0.406	1	−0.203	0.133
	Sig. (2-tailed)		0.248	0.471	0.604	0.572	0.128	0.348	0.068		0.377	0.565
	N		21	21	21	21	21	21	21	21	21	21
C_Peer	Pearson correlation		0.231	0.094	0.300	0.433*	−0.330	−0.120	−0.349	−0.203	1	0.297
	Sig. (2-tailed)		0.314	0.685	0.187	0.050	0.143	0.604	0.121	0.377		0.191
	N		21	21	21	21	21	21	21	21	21	21
C_Maternal	Pearson correlation		−0.170	−0.165	0.038	0.062	−0.326	−0.031	0.108	0.133	0.297	1
	Sig. (2-tailed)		0.462	0.476	0.872	0.789	0.150	0.894	0.642	0.565	0.191	
	N		21	21	21	21	21	21	21	21	21	21

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

TABLE 5 Multiple linear regression of Asian cultural values on children's perceived competence for local Chinese children ($N=24$).

Model	<i>B</i>	<i>SE</i>	β	<i>t</i>	R^2	<i>F</i>	<i>df</i>	<i>p</i>
Dependent variable: Children					0.285	0.822	6	0.571
– Cognitive competence								
AVS – Conformity	0.218	0.286	0.230	0.760				0.933
AVS – Achievement	−0.032	0.103	−0.103	−0.315				0.727
AVS – Emotion	−0.066	0.163	−0.103	−0.406				0.622
AVS – Collectivism	0.008	0.160	0.016	0.050				0.968
AVS – Humility	0.314	0.214	0.418	1.469				0.140
AVS – Filial Piety	0.077	0.189	0.108	0.407				0.833
Dependent variable: Children					0.229	1.014	6	0.455
– Physical competence								
AVS – Conformity	0.022	0.295	0.023	0.074				0.107
AVS – Achievement	−0.061	0.106	−0.195	−0.575				0.503
AVS – Emotion	−0.008	0.167	−0.013	−0.051				0.446
AVS – Collectivism	0.242	0.165	0.477	1.467				0.694
AVS – Humility	0.035	0.220	0.047	0.160				0.374
AVS – Filial piety	−0.026	0.195	−0.037	−0.136				0.202
Dependent variable: Children					0.643	3.865	6	0.017
– Peer competence								
AVS – Conformity	0.075	0.232	0.069	0.321				0.772
AVS – Achievement	−0.114	0.083	−0.315	−1.365				0.127
AVS – Emotion	−0.032	0.132	−0.043	−0.243				0.709
AVS – Collectivism	0.499	0.130	0.848	3.835				0.002
AVS – Humility	−0.720	0.173	−0.838	−4.162				0.003
AVS – Filial piety	0.110	0.154	0.134	0.714				0.663
Dependent variable: Children					0.424	1.351	6	0.299
– Maternal competence								
AVS – Conformity	−0.349	0.449	−0.211	−0.778				0.423
AVS – Achievement	−0.263	0.161	−0.476	−1.628				0.186
AVS – Emotion	−0.250	0.255	−0.223	−0.979				0.598
AVS – Collectivism	0.559	0.252	0.624	2.222				0.056
AVS – Humility	−0.994	0.335	−0.758	−2.968				0.022
AVS – Filial piety	0.128	0.297	0.103	0.430				0.851

AVS, Asian values scales.

TABLE 6 Multiple linear regression of Asian cultural values on children's perceived competence for south Asian children ($N=81$).

Model	<i>B</i>	<i>SE</i>	β	<i>t</i>	R^2	<i>F</i>	<i>df</i>	<i>p</i>
Dependent variable: Children – Cognitive competence					0.151	1.645	6	0.162
AVS – Conformity	0.218	0.194	0.191	1.123				0.085
AVS – Achievement	0.020	0.086	0.041	0.230				0.766
AVS – Emotion	−0.018	0.105	−0.031	−0.168				0.615
AVS – Collectivism	−0.138	0.108	−0.251	−1.277				0.78
AVS – Humility	−0.040	0.108	−0.062	−0.373				0.634
AVS – Filial piety	0.069	0.107	0.106	0.645				0.663
Dependent variable: Children – Physical competence					0.233	1.952	6	0.097
AVS – Conformity	−0.158	0.236	−0.108	−0.669				0.300
AVS – Achievement	0.139	0.104	0.222	1.328				0.061
AVS – Emotion	0.218	0.127	0.304	1.710				0.416
AVS – Collectivism	−0.229	0.132	−0.324	−1.735				0.581
AVS – Humility	−0.028	0.132	−0.034	−0.213				0.624

(Continued)

TABLE 6 (Continued)

Model	B	SE	β	t	R ²	F	df	p
AVS – Filial piety	0.172	0.130	0.208	1.325				0.079
Dependent variable: Children – Peer competence					0.097	0.667	6	0.677
AVS – Conformity	0.130	0.211	0.108	0.616				0.451
AVS – Achievement	0.009	0.093	0.017	0.095				0.941
AVS – Emotion	0.038	0.114	0.065	0.337				0.986
AVS – Collectivism	−0.185	0.118	−0.319	−1.575				0.294
AVS – Humility	0.021	0.118	0.031	0.182				0.956
AVS – Filial piety	0.031	0.116	0.046	0.271				0.388
Dependent variable: Children – Maternal competence					0.125	0.864	6	0.530
AVS – Conformity	0.066	0.169	0.068	0.394				0.488
AVS – Achievement	−0.019	0.075	−0.046	−0.256				0.389
AVS – Emotion	0.073	0.091	0.154	0.808				0.787
AVS – Collectivism	−0.024	0.094	−0.051	−0.257				0.833
AVS – Humility	−0.207	0.094	−0.370	−2.203				0.047
AVS – Filial piety	0.006	0.093	0.010	0.061				0.623

AVS, Asian values scales.

TABLE 7 Comparison of the first and second run of R² multiple linear regression of parenting style affecting children's perceived competence in local Chinese and south Asian groups.

	First run R ²	Second run R ²
Non-local Asian children		
Cognitive competence	0.018	0.028
Physical competence	0.06	0.054
Peer competence	0.058	0.059
Maternal competence	0.04	0.028
Local Chinese children		
Cognitive competence	0.202	0.069
Physical competence	0.222	0.132
Peer competence	0.118	0.062
Maternal competence	0.408	0.364

and Li, 2020). Conversely, local Chinese do not have the pressure to accommodate to the host culture, and the functional values of Asian cultural values in meeting their children's well-being are less incentivizing than their non-local counterparts. Finally, unlike the traditional beliefs of permissive indulgent parents who care about their children's needs but have difficulty setting limits in children's disciplines, the weak correlations Asian cultural values between and permissive parenting style in in both groups suggested that the characteristics of permissive parenting style had no alignment with Asian cultural values that are characterized with social order and hierarchy (Xu et al., 2005; Mojdehi et al., 2020).

Local Chinese parents

The relationship between authoritarian parenting style and the Asian cultural value of *achievement* (family recognition through achievement) was stronger than the other Asian cultural values

in local Chinese parents, despite considerably less adhering to Asian cultural values, in our study. The finding suggested that the local Chinese parents may take authoritarian parenting behaviors as functional means to achieve their expectations on children's achievement. The higher fidelity to the plans, the more control and monitoring would be needed, the more likely the expected outcomes can be achieved. If parents would like the child to obtain achievement to certain level, the closer the child follow their parents' plans, the more likely the child can achieve the expected level of achievement (Huang and Gove, 2015).

Similarly, local Chinese parents with permissive parenting style demonstrated a comparably higher positive correlation with *achievement* than the other cultural values. This result suggested that parents with indulgent parenting behaviors have expectation in children's achievement which implies these parents believe that indulgent parenting behaviors could be an optimal parenting style for their families and their children. A number of studies on indulgent parenting style suggested that its chances to bring positive developmental outcomes in children could be no less than authoritative (Chen et al., 2000; García and Gracia, 2014). Children raised with permissive parenting experienced less nervousness and could learn competent attitudes and behaviors from social behaviors of their parents (Garcia et al., 2021).

Regarding the role of *achievement* played in parents with authoritative parenting style, the moderate negative correlation between this style and *achievement* indicates that parents who expressed warmth and were sensitive to their children's needs were less likely to consider *achievement* more important than other Asian cultural values such as *collectivism* and *humility*. These parents are could more timely knowledge and understanding about the priority needs at early years development. Moral and character development could be considered as more important

values than achievement for parents with authoritative parenting styles (Huang and Gove, 2015; Hung, 2018).

Components of Asian cultural values and parenting styles

A closer examination a closer examination of the association of specific components of Asian cultural values on parenting style shows that *collectivism* and *humility* contributed to authoritative parenting behaviors in South Asian families. From a cultural and pragmatic perspective, *collectivism* and *humility* are not merely values cherished by Asian families but are also recognized as protective factors for the survival and well-being of ethnic minorities in the local cultural context (Ang and Goh, 2006). In a collectivist society, a clear social order, interdependence, and mutual good are emphasized and considered essential for maintaining social harmony (Rudy and Grusec, 2001). Collectivism provides spiritual, physical, social, and emotional support to community members amid adversities, which is critical for ethnic minorities (Sorkhabi, 2005). Collectivistic behavior provides children with strong spiritual resources, protects them from harm, and promotes their well-being in the mainstream culture (Rudy and Grusec, 2001; Lansford, 2021). Moreover, parents with more humility than others are likely to recognize their limitations, become aware of their strengths, encourage their children to acknowledge their needs, and identify themselves without devaluing their home values. According to cultural studies on adaptation and well-being, forced assimilation can be detrimental to individual well-being as people who experience enforced cultural assimilation tend to devalue their cultural background and family values to integrate into the dominant culture (Berry, 1980; Liu et al., 2009; Barker, 2015). Such devaluation can be predominantly seen among children who have a high need for acceptance and peer relationships during childhood because children's experiences and sense of achievement tend to be social, as opposed to adolescents who are academically oriented (Seligman, 2011; Barker, 2015). The perceived notion of the association between Asian cultural values and authoritarian parenting style or less desirable parenting practices must be re-visited and re-examined from the perspective of the significance of cultural values for the betterment of individuals' well-being (Xu et al., 2005).

Components of Asian cultural values and children's perceived competence

This study found that all the three measured domains of Asian cultural values were positively related to all four aspects of children's competence that were measured in this study. However, the specific Asian cultural values contributing to the specific aspects of competence varied between South Asian and local Chinese groups.

Collectivism

For the South Asian group, *collectivism* had negative correlations with *physical* competence, and *peer* competence, weak negative correlation with *cognitive* competence, and non-significant association with *maternal* competence. This suggests that the value of interdependency for *collectivism* could negatively affect children's school-related competence. In contrast, *collectivism* had significantly stronger positive correlations with *physical* competence, *peer* competence, and *maternal* competence, and non-significant association with *cognitive* competence in the local Chinese group than their South Asian counterparts. Such findings can be explained by the functional meaning of collectivism in practice in these two groups. South Asian children are an ethnic minority in the local population. Group-oriented practices among members of a minority group may highlight their "differences" and pose an obstacle to their accommodation and integration process. Practicing group cohesiveness and prioritization of the group over the individual goals or needs may not be an optimal integration strategy for a minority group to adapt and integrate into the mainstream (Barker, 2015). Hence, to facilitate the accommodation and integration process, children of the minority group may tend to their needs to adapt to the host culture. Furthermore, they may develop a feeling of belonging to the mainstream group by being behaviorally and cognitively more similar to their mainstream peers.

In contrast, collectivism in the local Chinese group would positively contribute to children's competence could be attributed to the identity of being a member of mainstream group (Rhee et al., 1995). Being mainstreamers would bring a strong sense of security to the communities. Mainstream values and practices are commonly considered the standard of societal expectations about how an individual should behave in society (Barker and Cornwell, 2019). Children with group-oriented identities and practices contribute to their competence of interacting with peers in everyday physical and cognitive interaction. The higher the group identity as mainstreamers, the more likely the individual would be accepted by the society, including his or her major caregivers (*maternal* competence in this study). When collectivism could lead to high social-related competence, including peer and maternal competence, group goal and achievement could be more important than individual achievement in the local Chinese group where mainstream values are exemplified as a group.

Achievement

South Asian and local Chinese groups could have different interpretations on *family recognition through achievement* (*achievement*). For South Asian parents, *achievement* was associated with children's *physical* competence but it had very weak associations with the rest of the children's competence. Physical development and capabilities are considered as more important than socio-emotional related development in South Asian group which can be attributed to their disadvantaged living environment (Huang et al., 2012). The environment commonly associated with poor hygiene and crimes (Nguyet and Leung,

2014). Therefore, it is reasonable to assume these families are more likely to take pride if their children are able to grow up with both physical and mentally resilient strengths. On the other hand, for local Chinese group, *achievement* had negative relationships with *maternal* acceptance and *peer* competence than other competences suggested that local Chinese parents valued socio-emotional related competence more than achievement which traditionally referred to academic achievement. In recent years, the importance of holistic development in young children has been widely promoted in public education delivered through parent education programs at schools, community health clinics, and media channels (Ng et al., 2020). With more voices shared in public media asking for non-competitive curriculum and learning environment, local Chinese parents have developed high awareness about the non-academic needs in nurturing competent children (Ng et al., 2020).

Humility

For the South Asian group, *humility* had a moderate negative association with *maternal* competence and very weak associations with the rest of the measured children's competence. On the contrary, *humility* was found to have very strong negative correlations with *peer* competence, *maternal* competence, and positive correlation with cognitive competence but had a weak correlation with *physical* competence in the local Chinese group. The phenomenon can be explained by the functional meaning of *humility* in these two groups. Humility in Confucian doctrine is interpreted as being humble, tolerant and unostentatious without boasting about one's achievement (Shek, 2006; Chuang et al., 2018). For South Asian group, parents may take pride of their children's achievement to promote the recognition by the mainstreamers and enhance the social status as member of ethnic minorities (Nguyet and Leung, 2014).

For both groups, the value of *humility* could impose different functions to the groups. For South Asian groups who are living in disadvantageous living environment, children being tolerate without showing off one's strengths, both intellectual and physical strengths, could put their children at the risk of bullying (Nguyet and Leung, 2014). For ethnic minorities residing in a local population while socializing in more than one cultures (home and host), survival and adaptation in these social contexts are needed for smoother transition across contexts (Bronfenbrenner and Morris, 1998). It does not necessarily mean South Asian group does not value humility in raising their children. Rather, humility could be applied to various extent depending on the needs of these families in contexts (Nguyet and Leung, 2014). Similarly, the strong negative association between *humility*, children's *peer* competence and *maternal* competence in local Chinese group does not necessarily mean local Chinese parents do not value humility in children nurturance. The traditional expectation for the humility in collectivist context may imply hindering one's strengths and even uniqueness (Chan et al., 2009). A more adaptable application humility that value both humility and

individual uniqueness would help strike the balance between the two emphasis. Such application would require the individuals to analyze and reflect (*cognitive* competence) on the needs and appropriateness in accordance to the social norm of their engaged context (Ang et al., 2007).

Relationship between parenting style and children's perceived competence

The null relationship between parenting style and children's perceived competence suggests that parenting style does not lead to more competent children. Ideal parenting practices should be interpreted from within meaningful contexts. With reference to PPCT model, South Asian parents' parenting practices that allow children to take pride in their homes' cultural values and uniqueness while being aware of adjusting and adapting to their host culture, are likely to promote easier transition between social contexts with multiple norms (Bronfenbrenner and Morris, 1998). No individual is exposed solely to a single culture or norm. Social norms are dynamic, varying across social contexts and between people (Johnson and Puplampu, 2008). Parents must nurture children with skills and competencies to respond to the ever-changing social norms (Schwartz et al., 2010). Therefore, ideal parenting practices are behaviors responsive to changes that emerge in various contexts and enable children to be open and flexible to embrace diversity at home, in school, and within their community (Xu et al., 2005).

Implications

The three important implications of this study include, contextualization of cultural values, the functional value of cultural norms in parenting practices, and reconceptualization of ideal parenting styles.

Contextualization of cultural values for Asian sub-cultural groups

Asian cultural value is a generic and inclusive term that homogenizes the complexity and intergroup diversity among various sub-cultural groups in any Asian city. The study's findings provide deeper insights into typical Asian cultural values that have long been literally interpreted and associated with negative connotations in the Western literature, specifically in the context of involved culture and familial needs (Chan et al., 2009; Wong et al., 2018). The motives underlying values such as obedience, collectivism, and social hierarchy, exercised by South Asian parents, aim at smoother cultural accommodation and meaning of love, which is responsive to children's needs rather than blindly exercising control over the dependents. Values in cultures have different meanings and perform varied functions for different goals to meet developmental needs in different immediate social contexts across one's lifespan. Further research is required to

examine the meaning, role, and functions of Asian cultural values at different stages of people's lives in Asian sub-cultural groups.

Pragmatism of Asian cultural values for ethnic minorities

Cultural studies claim that an ideal integration is acculturation (Barker, 2015). Acculturation attenuates individuals' pressure of adjusting to a new cultural environment and enables them to maintain and value their homes' cultural values while positively integrating into the new culture such that smooth transition between social contexts can be promoted (Barker, 2015). A positive cultural integration process allows for maintaining cultural integrity alongside participation in the host cultural context (Yuen, 2016). In this study, the pragmatic value of collectivism and humility have higher tempting values for adoption and adherence for vulnerable people, i.e., those who are subject to social exclusion and are under pressure for social adaptation or acceptance. The study's results can alert clinical researchers in family studies and family counseling practitioners to help ethnic minorities who blindly assimilate the host culture's values. They should be helped in a caring and responsive way to increase their awareness of their needs and uniqueness. Parents of ethnic minorities should also be helped to nurture young children with attitudes, skills and resilient strength that enable them to face cultural clashes, be flexible and open to multi-cultural or norm acculturation in the broader cultural context is essential for the well-being of all community members, both ethnic minorities and local population (Stamkou et al., 2019).

Rethinking ideal parenting

Ideal or appropriate parenting, regardless of the types of parenting styles, is determined by identifying if parenting practices can foster children's well-being, including addressing the needs of both family and children in meaningful contexts and promoting smooth transition across daily living contexts (Darling and Steinberg, 2017). The study reiterates to clinical researchers in family studies and family counseling practitioners that authoritarian parenting style, as well as other so called less ideal parenting styles, is not necessarily associated with negative outcomes in children. In addition, awareness and competence in applying culturally sensitive intervention strategies are critical to improving service quality and appropriateness for families from diverse backgrounds.

Limitations

This study has several limitations. First, the questionnaires for adherence to Asian cultural values and parenting practices were self-reported, and participating parents rated their parenting style based on the host cultural norms. Second, due to the caregivers in families with lower socioeconomic status are mothers, the gender of caregivers has not been investigated in this study. With the unitary gender type of the participating caretakers, the gender of the child was not included in this study as the exploration on dynamic between both genders of caretakers and children could be very

limited. Third, this study did not investigate the relationships between indulgent-neglectful parenting styles on child development as parents with this parenting style paid the least effort to understand children's needs and set limits in children's behaviors. Investigations on the influences of these parents on their children development could be very limited. Fourth, the difference in the number of children in South Asian and local Chinese group could post limitation on findings generalization. Finally, although existing cultural research suggested that Asian cultural values and practices have been highly influenced by Confucianism, Taoism and Buddhism (Vuong et al., 2018, 2020), this study did not explore the relationships between these doctrines and the values held by the members of each participating ethnic group for the following reason. Apart from local Chinese, since the South Asian participants in this study were collectively called "South Asian" without differentiating each ethnic group due to limited sample size with uneven ethnic group distribution and limited statistical power, the investigation on the influences of the three doctrines on participants' cultural values and parenting behaviors could be very limited.

Considering the above limitations, future research is suggested to explore the role of Asian cultural values adherence in fathers on parenting style and children's development. Also, in view of diverse ethnicities among Asian, particularly South Asian, population, future research could enhance the statistical power by expanding sample size with more balanced ethnic make up in the sample and investigating the relationship between each specific component of cultural values, including the three doctrines, parenting behaviors in each ethnic group with its unique socio-cultural history to explore the influences of cultural values in parenting practices of different Asian ethnic groups on children development. After all, although the present study with collected quantitative data can be improved and supplemented by interviewing parents, conducting on-site observations of children's social behaviors, expanding the sample size of both South Asian and local Chinese group, and enhancing the differentiation of different cultural groups among the participants, our results have invaluable reference values and provide pioneer insights in exploring contextualization of parenting styles for different ethnic groups within a geographical region rather than across countries and valuable reference for future research on parenting amid cultural diversities.

Conclusion

Globalization facilitates cultural mobility within and across regions and countries. Parenting in multicultural societies is challenging and must be responsive to the changing dynamics in children's socio-cultural contexts that enable them to live with competence and dignity. This study complements the literature on parenting young children from diverse Asian sub-cultural groups. Clinical researchers in family studies and family counseling practitioners could use these findings to gain insight into the existing influences of home and host culture on parenting practices. Further, they could use these findings to integrate

cultural sensitivity into family interventions by contextualizing cultural values and interpreting them as pragmatic parenting practices for families living in metropolitan cities, such as Hong Kong. Longitudinal research could explore the impact of specific parenting behaviors on local and immigrant children's development and well-being at different life stages.

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 Hong Kong Metropolitan University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

References

- Alegre, A. (2011). Parenting styles and children's emotional intelligence: what do we know. *Fam. J.* 19, 56–62. doi: 10.1177/1066480710387486
- Ang, R. P., and Goh, D. H. (2006). Authoritarian parenting style in Asian societies: a cluster analytic investigation. *Contemp. Fam. Ther.* 28, 131–151. doi: 10.1007/s10591-006-9699-y
- Ang, S., Van Dyne, L., Koh, C., Ng, K., Templer, K., Tay, C., et al. (2007). Cultural intelligence: its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance. *Manag. Organ. Rev.* 3, 335–371. doi: 10.1111/j.1740-8784.2007.00082.x
- Antony-Newman, M. (2019). Parental involvement of immigrant parents: a meta-synthesis. *Educ. Rev.* 71, 362–381. doi: 10.1080/00131911.2017.1423278
- Barker, G. G. (2015). Choosing the best of both worlds: the acculturation process revisited. *Int. J. Intercult. Relat.* 45, 56–69. doi: 10.1016/j.ijintrel.2015.01.001
- Barker, G. G., and Cornwell, T. L. (2019). Acculturation, communication, and family relationships: challenges and opportunities. *West. J. Commun.* 83, 624–646. doi: 10.1080/10570314.2019.1620961
- Baumrind, D. (1978). Parental disciplinary patterns and social competence in children. *Youth Soc.* 9, 239–267. doi: 10.1177/0044118x7800900302
- Berry, J. (1980). "Acculturation as varieties of adaptation" in *Acculturation: Theory, Models, and Some New Findings*. ed. A. Padilla (Boulder, CO: Westview), 9–25.
- Bornstein, M. H. (2012). Cultural approaches to parenting. *Parent. Sci. Pract.* 12, 212–221. doi: 10.1080/15295192.2012.683359
- Brock, D. M., Sarason, I. G., Sanghvi, H., and Gurung, R. A. R. (1998). The perceived acceptance scale: development and validation. *J. Soc. Pers. Relat.* 15, 5–21. doi: 10.1177/0265407598151001
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, Massachusetts and London: Harvard University Press.
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: research perspectives. *Dev. Psychol.* 22, 723–742. doi: 10.1037/0012-1649.22.6.723
- Bronfenbrenner, U. (1992). "Ecological systems theory," in *Six Theories of Child Development: Revised Formulations and Current Issues*. ed. R. Vasta (London, Bristol, PA: Jessica Kingsley Publishers), 187–249.
- Bronfenbrenner, U., and Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: a bioecological model. *Psychol. Rev.* 101, 568–586. doi: 10.1037/0033-295X.101.4.568
- Bronfenbrenner, U., and Morris, P. A. (1998). "The ecology of developmental processes," in *Handbook of Child Psychology: Theoretical Models of Human Development*. eds. W. Damon and R. M. Lerner (New York: John Wiley & Sons Inc.), 993–1028.
- Census and Statistics Department (2011). Data from: Hong Kong 2011 Population Census – Summary Results. Available at: <https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B1120055&scode=170> (Accessed February 1, 2022).
- Census and Statistics Department (2016). Data from: Hong Kong 2016 Population by-Census—Thematic Report: Ethnic Minorities. Available at: <https://www.censtatd.gov.hk/hkstat/sub/sp459.jsp?productCode=B1120100> (Accessed February 1, 2022).
- Chan, S. M., Bowes, J., and Wyver, S. (2009). Chinese parenting in Hong Kong: links among goals, beliefs and styles. *Early Child Dev. Care* 179, 849–862. doi: 10.1080/03004430701536525
- Chan, Y., and Li, J. (2020). An early look at parental expectation towards early childhood education among Pakistani parents in Hong Kong: the role of culture and acculturation. *Child Youth Serv. Rev.* 119:105652. doi: 10.1016/j.childyouth.2020.105652
- Chao, R. K. (1994). Beyond parental control and authoritarian parenting style: understanding Chinese parenting through the cultural notion of training. *Child Dev.* 65:1111. doi: 10.2307/1131308
- Chao, R. K. (2001). Extending research on the consequences of parenting style for Chinese Americans and European Americans. *Child Dev.* 72, 1832–1843. doi: 10.1111/1467-8624.00381
- Chao, R. K., and Kaeochinda, K. F. (2010). "Parental sacrifice and acceptance as distinct dimensions of parental support among Chinese and Filipino American adolescents," in *Asian American Parenting and Parent-Adolescent Relationships, Part of the Advancing Responsible Adolescent Development, Book Series (ARAD)*. eds. S. Russell, L. Crockett and R. Chao (New York, NY: Springer), 61–77.
- Chen, X., Liu, M., and Li, D. (2000). Parental warmth, control, and indulgence and their relations to adjustment in Chinese children: a longitudinal study. *J. Fam. Psychol.* 14, 401–419. doi: 10.1037/0893-3200.14.3.401
- Chen-Bouck, L., Patterson, M. M., and Chen, J. (2019). Relations of collectivism socialization goals and training beliefs to Chinese parenting. *J. Cross-Cult. Psychol.* 50, 396–418. doi: 10.1177/0022022118822046
- Christensen, J. (2016). A critical reflection of Bronfenbrenner's development ecology model. *Probl. Educ. 21st. Century.* 69, 22–28. doi: 10.33225/pec/16.69.22
- Chuang, S. S., Glazman, J., Green, D. S., and Rasmi, S. (2018). Parenting and family relationships in Chinese families: a critical ecological approach. *J. Fam. Theory Rev.* 10, 367–383. doi: 10.1111/jftr.12257
- Darling, N., and Steinberg, L. (2017). "Parenting style as context: an integrative model," in *Interpersonal Development Routledge*. eds. B. Laursen and R. Žukauskienė (Routledge), 161–170.

Author contribution

The author confirms being the sole contributor of this work and has approved it for publication.

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.

- Delvecchio, E., Germani, A., Raspa, V., Lis, A., and Mazzeschi, C. (2020). Parenting styles and child's well-being: the mediating role of the perceived parental stress. *Eur. J. Psychol.* 16, 514–531. doi: 10.5964/ejop.v16i3.2013
- Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis*. New York: John Wiley.
- Dwairy, M., and Achoui, M. (2010). Parental control: a second cross-cultural research on parenting and psychological adjustment of children. *J. Child Fam. Stud.* 19, 1–7. doi: 10.1007/s10826-009-9334-2
- Fakhouri, T. H. I., Hughes, J. P., Brody, D. J., Kit, B. K., and Ogden, C. L. (2013). Physical activity and screen-time viewing among elementary school-aged children in the United States from 2009 to 2010. *JAMA Pediatr.* 167, 223–229. doi: 10.1001/2013.jamapediatrics.122
- Fantuzzo, J. W., McDermott, P. A., Manz, P. H., Hampton, V. R., and Burdick, N. A. (1996). The pictorial scale of perceived competence and social acceptance: does it work with low-income urban children? *Child Dev.* 67:1071. doi: 10.1111/j.1467-8624.1996.tb01783.x
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. London: Sage Publications Ltd.
- Fleming, K. (2019). “Who is “diverse”?: (in)tolerance, education, and race in Hong Kong” in *Education, Ethnicity and Equity in the Multilingual Asian Context. Multilingual Education*. eds. J. Gube and F. Gao (Singapore: Springer), 89–103.
- García, F., and Gracia, E. (2014). “The indulgent parenting style and developmental outcomes in south European and Latin American countries,” in *Parenting across cultures. Science across cultures: The history of non-Western Science*. ed. H. Selin, vol. 7 (Dordrecht: Springer)
- García, O. F., Lopez-Fernandez, O., and Serra, E. (2021). Raising spanish children with an antisocial tendency: do we know what the optimal parenting style is? *J. Interpers. Violence* 36, 6117–6144. doi: 10.1177/0886260518818426
- Harter, S., and Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Dev.* 55:1969. doi: 10.2307/1129772
- Heritage, B., Manchini, V., Rigoli, D., and Piek, J. (2020). Measure reliability of the perceived competence and social acceptance for children scale via rasch analysis. *Br. J. Educ. Psychol.* 90, 130–151. doi: 10.1111/bjep.12268
- Hinkley, T., Salmon, J., Okely, A. D., Crawford, D., and Hesketh, K. (2012). Preschoolers' physical activity, screen time, and compliance with recommendations. *Med. Sci. Sports Exerc.* 44, 458–465. doi: 10.1249/mss.0b013e318233763b
- Hosken, D. J., Hunt, J., and Wedell, N. (2019). *Genes and Behavior: Beyond Nature-Nurture*. New York, NY: Wiley.
- Huang, K., Calzada, E., Cheng, S., and Brotman, L. M. (2012). Physical and mental health disparities among young children of Asian immigrants. *J. Pediatr.* 160, 331–336.e1. doi: 10.1016/j.jpeds.2011.08.005
- Huang, G., and Gove, M. (2015). Asian parenting styles and academic achievement: views from eastern and Western perspectives. *Education (Chula Vista)* 135, 389–397.
- Huber, L., Plötner, M., and Schmitz, J. (2019). Social competence and psychopathology in early childhood: a systematic review. *Eur. Child Adolesc. Psychiatry* 28, 443–459. doi: 10.1007/s00787-018-1152-x
- Hue, M., and Kennedy, K. J. (2014). Creating culturally responsive environments: ethnic minority teachers' constructs of cultural diversity in Hong Kong secondary schools. *Asia. Pac. J. Educ.* 34, 273–287. doi: 10.1080/02188791.2013.823379
- Hung, J. (2018). Parenting styles, academic demands and children's psychosocial well-being: Why today's Hong Kong Chinese students are so stressed. *J. Global Health* 8:2. doi: 10.7916/thejgh.v8i2.4854
- Johnson, G. M., and Puplampu, K. P. (2008). Internet use during childhood and the ecological techno-subsystem. *Canadian J. Learn. Technol.* 34, 19–28. doi: 10.21432/T2CP4T
- Kim, B. K., Atkinson, D. R., and Yang, P. H. (1999). The Asian values scale: development, factor analysis, validation, and reliability. *J. Couns. Psychol.* 46, 342–352. doi: 10.1037/0022-0167.46.3.342
- Kim, B. K., Li, L. C., and Ng, G. F. (2005). The Asian American values scale—multidimensional: development, reliability, and validity. *Cultur. Divers. Ethnic Minor. Psychol.* 11, 187–201. doi: 10.1037/1099-9809.11.3.187
- Lansford, J. E. (2021). Annual research review: cross-cultural similarities and differences in parenting. *J. Child Psychol. Psychiatry* 63, 466–479. doi: 10.1111/jcpp.13539
- Lau, E., and Power, T. (2020). Coparenting, parenting stress, and authoritative parenting among Hong Kong Chinese mothers and fathers. *Parent. Sci. Pract.* 20, 167–176. doi: 10.1080/15295192.2019.1694831
- Li, M. (2013). Acculturation, filial responsibilities and living arrangements: an empirical study of the acculturative experiences of elderly Chinese immigrants in New Zealand. *Intercult. Commun. Stud.* 22, 301–322.
- Lin, X., Liao, Y., and Li, H. (2021). Parenting styles and social competence in Chinese preschoolers: a moderated mediation model of singleton and self-regulation. *Early Educ. Dev.* 33, 437–451. doi: 10.1080/10409289.2021.1940643
- Liu, Y., Fang, X., Cai, R., Wu, Y., and Zhang, Y. (2009). The urban adaptation and adaptation process of urban migrant children: a qualitative study. *Frontiers. Education. China.* 4, 365–389. doi: 10.1007/s11516-009-0020-3
- Magno, C. (2010). Looking at Filipino pre-service teachers' value for education through epistemological beliefs about learning and Asian values. *Asia-Pac. Educ. Res.* 19, 61–78.
- Maria, G. O., Tagliabue, S., and Confalonieri, E. (2013). Parenting style and dimensions questionnaire: a review of reliability and validity. *Marriage Fam. Rev.* 49, 465–490. doi: 10.1080/01494929.2013.770812
- Mojdehi, A. S., Shohoudi, A., and Talwar, V. (2020). Children's moral evaluations of different types of lies and parenting practices and across cultural contexts. *Curr. Psychol.* 41, 5420–5433. doi: 10.1007/s12144-020-01059-7
- Mousavi, A., and Juhari, R. (2019). Systematic review of parenting style and children's emotional intelligence: recent updates on Western and non-western families. *Malays. J. Med. Health. Sci.* 15, 147–152.
- Ng, F. F., and Wang, Q. (2019). “Asian and Asian American parenting,” in *Handbook of Parenting*. ed. M. H. Bornstein. 3rd Edn. Vol. 4: Social conditions and applied parenting (Mahwah, NJ: Routledge), 108–169.
- Ng, D. T., Fisher, J. W., Au, M. L., and Lo, S. K. (2020). Parental perceptions of holistic early childhood education in Hong Kong. *Educ. Plan.* 27, 49–60.
- Nguyen, E. J., and Leung, L. Y. (2014). *Understanding South Asian Minorities in Hong Kong. 1st Edn.* Hong Kong: Hong Kong UP, HKU, 2014.
- Oei, T., and Raylu, N. (2009). The relationship between cultural variables and gambling behavior among Chinese residing in Australia. *J. Gambl. Stud.* 25, 433–445. doi: 10.1007/s10899-009-9148-6
- Oh, J. H., and Lee, J. H. (2014). Asian values, ethnic identity, and acculturation among ethnic Asian wives in South Korea. *J. Int. Migr. Integr.* 15, 73–91. doi: 10.1007/s12134-012-0269-x
- Olla, M. G., Daulima, N. H. C., and Putri, Y. S. P. (2018). The experience of parents implementing authoritarian parenting for their school-age children. *Enferm. Clin.* 28, 122–125. doi: 10.1016/S1130-8621(18)30050-0
- Pong, S., Hao, L., and Gardner, E. (2005). The roles of parenting styles and social capital in the school performance of immigrant Asian and Hispanic adolescents. *Soc. Sci. Q.* 86, 928–950. doi: 10.1111/j.0038-4941.2005.00364.x
- Pong, S., Johnston, J., and Chen, V. (2010). Authoritarian parenting and Asian adolescent school performance: insights from the US and Taiwan. *Int. J. Behav. Dev.* 34, 62–72. doi: 10.1177/0165025409345073
- Rhee, E., Uleman, J. S., Lee, H. K., and Roman, R. J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. *J. Pers. Soc. Psychol.* 69, 142–152. doi: 10.1037/0022-3514.69.1.142
- Robinson, C. C., Hart, C. H., Mandelco, B. L., Olsen, S. F., Russel, A., Aloa, V., et al. (1996). Psychometric Support for a New Measure of Authoritative, Authoritarian, and Permissive Parenting Practices: Cross-Cultural Connections. 23 p; Paper Presented at the Biennial Conference of the International Society for the Study of Behavioral Development (14th, Quebec City, Quebec, Canada, August 12–16, 1996). Speeches/Conference Papers (150) Reports Evaluative / Feasibility (142)
- Robinson, C. C., Mandelco, B., Olsen, S. F., and Hart, C. H. (1995). Authoritative, authoritarian, and permissive parenting practices: development of a new measure. *Psychol. Rep.* 77:3. doi: 10.2466/pr0.1995.77.3.81
- Rudy, D., and Grusec, J. E. (2001). Correlates of authoritarian parenting in individualist and collectivist cultures and implications for understanding the transmission of values. *J. Cross-Cult. Psychol.* 32, 202–212. doi: 10.1177/0022022101032002007
- Rudy, D., and Grusec, J. E. (2006). Authoritarian parenting in individualist and collectivist groups: associations with maternal emotion and cognition and children's self-esteem. *J. Fam. Psychol.* 20, 68–78. doi: 10.1037/0893-3200.20.1.68
- Schwartz, S. J., Unger, J. B., Zamboanga, B. L., and Szapocznik, J. (2010). Rethinking the concept of acculturation. Implications for theory and research. *Am. Psychol.* 65, 237–251. doi: 10.1037/a0019330
- Seligman, M. E. P. (2011). *Flourish: A Visionary New Understanding of Happiness and Well-Being* New York: Free Press.
- Shek, D. L. (2006). Perceived parental behavioral control, psychological control, and parent-child relational qualities in Chinese adolescents in Hong Kong. *Am. J. Fam. Ther.* 34, 163–176. doi: 10.1080/01926180500357891
- Shim, Y. R., and Schwartz, R. C. (2008). Degree of acculturation and adherence to Asian values as correlates of psychological distress among Korean immigrants. *J. Ment. Health* 17, 607–617. doi: 10.1080/09638230701506838

- Shoshani, A., and Steinmetz, S. (2014). Positive psychology at school: a school-based intervention to promote adolescents' mental health and well-being. *J. Happiness Stud.* 15, 1289–1311. doi: 10.1007/s10902-013-9476-1
- Sorkhabi, N. (2005). Applicability of Baumrind's parent typology to collective cultures: analysis of cultural explanations of parent socialization effects. *Int. J. Behav. Dev.* 29, 552–563. doi: 10.1080/01650250500172640
- Stamkou, E., Kleef, G. A., Homan, A. C., Gelfand, M. J., Van de Vijver, F. J. R., Van Egmond, M. C., et al. (2019). Cultural collectivism and tightness moderate responses to norm violators: effects on power perception, moral emotions, and leader support. *Personal. Soc. Psychol. Bull.* 45, 947–964. doi: 10.1177/0146167218802832
- Venetsano, F., Kossyva, I., Valentini, N., Afthentopoulou-Evangelia, A., and Barnett, L. M. (2018). Validity and reliability of the pictorial scale of perceived movement skill competence for young Greek children. *J. Mot. Learn. Dev.* 6, S239–S251. doi: 10.1123/jmld.2017-0028
- Vu, K. T. T., Castro, K. M., Cheah, C. S. L., and Yu, J. (2019). Mediating and moderating processes in the associations between Chinese immigrant mothers' acculturation and parenting styles in the United States. *Asian Am. J. Psychol.* 10, 307–315. doi: 10.1037/aap0000150
- Vuong, Q. H., Bui, Q. K., La, V. P., Vuong, T., Nguyen, V. T., Ho, M., et al. (2018). Cultural additivity: behavioural insights from the interaction of Confucianism, Buddhism and Taoism in folktales. *Palgrave Commun.* 4:143. doi: 10.1057/s41599-018-0189-2
- Vuong, Q. H., Ho, M. T., Bui, Q., La, V., Vuong, T., Nguyen, V. T., et al. (2020). On how religions could accidentally incite lies and violence: folktales as a cultural transmitter. *Palgrave Commun.* 6:82. doi: 10.1057/s41599-020-0442-3
- Wong, D., Zhuang, X., and Ng, T. (2018). Is parental control beneficial or harmful to the development of young children in Hong Kong? *J. Child Fam. Stud.* 28, 831–838. doi: 10.1007/s10826-018-1301-3
- Xia, X. (2020). Parenting style and Chinese children's school readiness outcomes: the moderating role of socioeconomic status. *Child Youth Serv. Rev.* 118:105381. doi: 10.1016/j.childyouth.2020.105381
- Xia, M., Li, X., and Tudge, J. R. H. (2020). Operationalizing Urie Bronfenbrenner's process-person-context-time model. *Hum. Dev.* 64, 10–20. doi: 10.1159/000507958
- Xu, Y., Farver, J. A., Zhang, Z., Zeng, Q., Yue, L., and Cai, E. (2005). Mainland Chinese parenting styles and parent-child interaction. *Int. J. Behav. Dev.* 29, 524–531. doi: 10.1080/01650250500147121
- Yuen, C. Y. M. (2016). Enhancing early childhood schooling of south Asian children in Hong Kong: beliefs and perceptions of kindergarten teachers and principals. *Early Child Dev. Care* 186, 403–418. doi: 10.1080/03004430.2015.1036420



OPEN ACCESS

EDITED BY

Douglas F. Kauffman,
Medical University of the Americas –
Nevis, United States

REVIEWED BY

Luxin Yang,
Beijing Foreign Studies
University, China
Ester Trigo-Ibáñez,
Universidad de Cádiz, Spain

*CORRESPONDENCE

Reza Kafipour
kafipour@sums.ac.ir

SPECIALTY SECTION

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

RECEIVED 03 May 2022

ACCEPTED 29 September 2022

PUBLISHED 20 October 2022

CITATION

Moslehi S and Kafipour R (2022)
Syntactic structure and rhetorical
combinations of Iranian English
research article titles in medicine and
applied linguistics: A cross-disciplinary
study. *Front. Educ.* 7:935274.
doi: 10.3389/feduc.2022.935274

COPYRIGHT

© 2022 Moslehi and Kafipour. 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.

Syntactic structure and rhetorical combinations of Iranian English research article titles in medicine and applied linguistics: A cross-disciplinary study

Shadab Moslehi and Reza Kafipour *

Department of English Language, School of Paramedical Sciences, Shiraz University of Medical Sciences, Shiraz, Iran

This corpus-based cross-disciplinary study investigated the syntactic structures and rhetorical combinations of 200 research article titles in two disciplines, namely, medicine (100) and applied linguistics (100). The RA titles were selected from four reputable Iranian English journals. The authors were all Iranian researchers in these two disciplines. The titles were analyzed in light of Dietz taxonomy. First, the frequencies and percentages of each syntactic and rhetorical construction occurrence have been calculated. Next, the authors performed a *t*-test regarding title length and the chi-square test to decide whether syntactic or rhetorical construction is a discipline-specific convention. The findings revealed that medical titles were longer than the linguistic ones. The frequency and percentage of both single- and multi-unit RA titles were essentially the same in both disciplines. Concerning the syntactic components of single-unit RA titles, the most frequently used structure was the nominal construction, followed by verbal and prepositional ones. The most recurrent syntactic components of the nominal structure in both disciplines were post- and pre-modified, with medical titles overtaking the linguistic titles in all nominal categories. In terms of verbal constructions, the dominant structure in medicine was the full sentence, and in applied linguistics, gerund phrases. Regarding the rhetorical components of multi-unit RA titles, medical titles took precedence over the linguistic titles in using the topic method. The topic scope and topic description organization are mostly reported in applied linguistics titles. In this study, however, two new rhetorical combinations were identified. The chi-square test results only confirmed the verbal structures of single-unit and the rhetorical combinations of multi-unit RA titles as distinctive features for each discipline.

KEYWORDS

applied linguistics, medicine, research article titles, rhetorical organization, syntactic structure

Introduction

Over the last three decades, titles and their features in different genres (e.g., dissertations, research articles, and review papers) have enticed most researchers and readers into researching them (Jalilifar et al., 2012). In fact, a title brings about an identity for any academic piece of work; that is why it is worth investigating. Furthermore, as Haggan (2004) mentioned, authors work on the title, and readers read it as the first part of an article. Therefore, the title has a critical role in readers' decision to read a paper or ignore it; in other words, it can motivate or demotivate the readers to read an article (Archibald, 2005; Hartley, 2005, 2007).

Since the 1990s, the genre has found a prominent role in different types of text analysis. Genre-based analysis originated from discourse analysis and has been widely used in English for a specific purpose (Swales, 1990; Dudley-Evans, 1994; Bhatia, 2008; Martín and León Pérez, 2014). Research articles (RAs) received considerable attention in genre-based analysis research because they are considered primary mediums for scientific communication and the worldwide distribution of academic knowledge (Peacock, 2002). Therefore, a growing appeal emerged to present patterns of scientific and literary texts for genre-based research in RAs (Martin, 2003). Swales (1990) book on genre analysis motivated researchers to study different aspects of RAs: micro-structures such as voice, tense, and pronouns, and macro-structures like introduction, method, result, and discussion. Researchers, in the last three decades, conducted numerous research on the titles in various academic genres, like review articles (Soler, 2007), dissertations (Dudley-Evans, 1984), and original research articles (Goodman, 2000; Haggan, 2004; Wang and Bai, 2007). However, the syntactic construction of titles has not been widely researched in comparative RA studies, especially in medicine and applied linguistics in an EFL context. To the best of the researcher's knowledge, little research has been conducted in this domain. Therefore, the purpose of the present study is to highlight the syntactic structures of RAs in the two fields, namely, medicine and applied linguistics, and shed light on the differences, if any, between these disciplines in terms of syntactic constructions of RA titles.

Literature review

As Cheng et al. (2012) stated, several studies have been conducted until now to examine research titles syntactically, highlighting the title length (e.g., Yitzhaki, 1994; Anthony, 2001; Haggan, 2004), structural organizations of titles (e.g., Haggan, 2004; Wang and Bai, 2007; Salager-Meyer and Alcaraz Ariza, 2013; Archibald, 2017), titles in different genres (e.g., Hamp-Lyons, 1987; Soler, 2007; Méndez and Alcaraz, 2017), titles in various fields (e.g., Buxton and Meadows, 1977; Moattarian

and Alibabae, 2015; Nagano, 2015; Shahidpour and Alibabae, 2017), and titles in different languages (e.g., Soler, 2007).

According to the existing literature, researchers used mono- and cross-disciplinary methods to study various aspects of titles in academic texts. However, cross-disciplinary studies do not have mono-disciplinary research restrictions; therefore, they have set the scene in interpreting extant phenomena by highlighting similarities and differences. Furthermore, cross-disciplinary research brings novelty to language studies and integrates new knowledge (Herieg, 2011). Buxton and Meadows (1977); Fortanet Gómez et al. (1998); Haggan (2004), Lewinson and Hartley (2005), Soler (2007), and Haggan (2004), Wang and Bai (2007), Moattarian and Alibabae (2015); Shahidpour and Alibabae (2017) are among those researchers who have conducted cross-disciplinary studies on RA titles in terms of syntactic structures and other aspects of RA titles. All of these studies have shed light on the correlations between different facets and patterns of titles across a range of disciplines as RA titles reveal the researchers' approach to their disciplinary practices and how well they can reach their potential readers.

In fact, the research titles mirror how the authors appeal to their readers. Researchers want their academic papers to be read and cited; however, with the tsunami of research papers in every discipline, this becomes a challenging job. Thus, to fulfill this objective, the authors have resorted to various means, including length, style, cultural allusions, compounds, and questions. In truth, some researchers have conducted various studies to tap into different methods of writing research titles in different fields.

Title formats

Among all the researchers, Buxton and Meadows (1977) pioneered research on RA titles. They investigated RA titles related to natural and social sciences in English, French, and German journals by observing content words. The researchers reported more informativity in natural sciences RA titles, especially chemistry and botany, than in social sciences. In addition, they reported frequent use of nouns and readers' inclination to retrieve information as the primary difference between RA titles in these fields.

Fortanet Gómez et al. (1997) investigated the structure and content of RA titles in different disciplines, including computer science, applied linguistics, business, economics, and chemistry. The study results indicated that chemistry RA titles had the highest number of words, and linguistics had the lowest. In addition, the study showed that the length of titles varied across the different disciplines. For example, most titles in chemistry and the minority in linguistics represented the general topic and specific focus of the research. By contrast, one-third of the titles indicated the nature of the study conducted. Since then, a trend has appeared toward investigations of syntactic

structures of RA titles (Moattarian and Alibabae, 2015). A year later, Fortanet Gómez et al. (1998) studied 800 titles in terms of their syntactic structures in computer science, applied linguistics, business and economics, and chemistry. The authors came up with noun phrases as the most frequently used syntactic structures, including a pre-modifier, head, and post-modifier, with the highest frequency in chemistry and computer science. Also, the scrutiny revealed that both applied linguistics and business and economics indicated a majority of -ing forms of the verbs (gerunds) in their titles.

Busch-Lauer (2000) also examined linguistic and medical titles of research articles and conference proceedings written in German and English. The results showed 8.4-word titles in linguistics and 9.9-word titles in medicine, which implies that titles in linguistics are shorter than those in medicine. Appiah et al. (2019) also worked on both natural and social sciences. After investigating title structures in a corpus of 574 titles in three disciplines, namely, gynecology/obstetrics, business, and law, Appiah et al. (2019) found that noun phrases were observed widely in the titles across the three fields; however, business titles were longer than those in gynecology and law and more probable to have compound units including a colon. Syntactically, nominal structures largely dominated single-unit titles in the three disciplines, and both pre- and post-modified titles were dominant in all the disciplines.

Nagano (2015) and Milojevic (2017) also argued that the length of research titles differs between disciplines. Nagano (2015) investigated the structures of a 3,200-title corpus of RAs published in authoritative journals in botany, fluid engineering, geology, and medicine as hard sciences, and economics, education, history, and sociology as soft sciences. The findings revealed the soft sciences had shorter titles, more multi-unit titles, fewer titles with noun phrases, a lower word rate, and higher application of the article “the” to initiate a title unit than the hard sciences. Along the same lines, Milojevic (2017) investigated the features of 500 research article titles in nursing and argued that the title length varies between disciplines. The results also indicated significant differences represented in article titles of internationally high-impact factor journals regarding the four features of style, length, structure, and content. Milojevic also found that the titles are longer in hard sciences such as medicine than in discursive sciences like sociology. Soler (2007) noted the same point after investigating syntactic title structures in biological and social sciences research and review papers. The findings indicated that soft sciences titles were shorter than hard sciences titles. The primary syntactic structures were nominal (the most frequently used), question, full-sentence, and compound constructions. The author claimed that the differences are not limited to disciplines, and they even exist between examined genres: research papers and review articles. For instance, the full-sentence structure was common in medicine, biology, and biochemistry RA titles, but there was no such construction in the review paper titles. Also, no incidence

of the full-sentence structure was observed in linguistics and psychology titles.

In their study, Hyland and Zou (2022) explored the key features of 5,070 titles in the highly ranked journals of six disciplines in the sciences, social sciences, and humanities to seek their typical structural patterns and characteristics of titles across a range of disciplines. A variety of disciplinary differences discovered were traced to distinct features of the fields. The findings are in contrast to the previous ones, showing that titles in the more discursive soft fields were lengthier, more interrogative, contained compound forms, and, except for history, were more likely to comprise the approach or findings. As it is evident, there is variation in interests and foci with titles regarding different topics, journals, and disciplines. One justification is that titles are intended to attract and inform the readers in various ways in different contexts (Hyland and Zou, 2022).

Title syntactic structures

Besides, some studies have taken a normative stance, encouraging the researchers the ways to entice their audience by employing title characteristics that are linked to high citation counts (e.g., Paiva et al., 2012; Thelwall, 2017). Examples of those structures are nominal, prepositional, and verbal, which are identified in different discipline-specific titles. For instance, Doykova (2016) examined the syntactic structures of 500 RA titles, written for reputable medicine and dental medicine journals (2010–2016) by non-native speakers of English. In addition to the title, word length and type were studied. This corpus-based study focused on the typical keywords and collocations employed as headings. It was revealed that the nominal title and its subtypes were the most stable structures.

Similarly, Haggan (2004) studied syntactic differences among research articles titles in linguistics, literature, and science. He classified the titles as full-sentence titles (especially in science), compound titles (especially in literature), and the remaining title structures. The last category included three subcategories: noun (especially in literature and linguistics), prepositional, and participial phrases. The author also found that science titles had 13.8 words, whereas literature and linguistics titles consisted of 8.8 words, implying science titles are longer than the other two fields. This finding is consistent with that claimed by Busch-Lauer (2000) and Fortanet Gómez et al. (1997). Furthermore, the results indicated that among phrasal structures, the prepositional phrase had the lowest frequency, with linguistics and literature reflecting a higher incidence of such structure than science. Finally, the author reported that compound or nominal structures occurred more frequently than complete sentences, prepositional structures, and V-ing phrases in research article titles.

Cheng et al. (2012) also examined the syntactic structures of 796 RA titles in applied linguistics journals indexed in the Social Sciences Citation Index (SSCI). Compound, nominal, full-sentence, V-ing, and prepositional phrases were the syntactic structures identified by the authors. Compound and nominal titles were the most frequently observed titles, respectively. However, other structures hardly occurred. For compound titles, a total of 11 categories were identified, including topic scope, topic method, topic description, topic source, metaphor topic, and topic question. Concerning nominal titles, the authors identified two types of discipline-specific and non-discipline-specific heads. Most of the discipline-specific heads characterize compound nouns. Post-modifiers are prepositional phrases.

In another study by Wang and Bai (2007), 417 titles were studied in medical research articles. They analyzed the structure of RA titles. The authors came up with an average length of 10.9 words. Most of the titles were nominal groups (99%), having single heads (75%), followed by prepositional post-modifiers (68%). In terms of rhetorical functions and syntactic structures, in their cross-disciplinary and cross-linguistic corpus study, Shahidpour and Alibabae (2017) studied 750 English RA titles and 750 Persian ones published in electrical engineering, psychiatry, and linguistics journals. They followed Dietz (1995) taxonomy and reported rhetorical functions and syntactic structures of RA articles. The findings indicated that the differences mainly occurred in title components, length, and style. Persian titles were longer than their English counterparts in length. Titles related to linguistics journals were the shortest, while those of psychiatry were the longest. Most titles were single-unit ones. The English researchers used multi-unit titles more than the Persian researchers. The incidence of multi-unit titles was most observed in linguistic RA titles. The authors found no significant difference concerning the combinations of multi-unit titles in the two languages. Yet, topic description structures were the most occurring combination in electrical engineering and linguistics titles, while in psychiatry titles, topic method construction was the most dominant combination. The most frequent syntactic structure among single-unit titles was the post-modified nominal group.

In another comprehensive cross-genre study, Gesuato (2008) analyzed the syntactic structure of 250 titles from four genres within applied linguistics, namely, books, dissertations, journal articles, and conference proceedings (CPs). She found that the noun phrase was the most frequently used structure in all examined genres, including books, dissertations, RAs, and CPs, with 87.5, 86.9, 83.5, and 82.0% frequency, respectively. Furthermore, dissertations had the longest titles, whereas books had the shortest titles. In another study, Moattarian and Alibabae (2015) investigated the syntactic structures of RA titles in applied linguistics, dentistry, and civil engineering. They analyzed 420 randomly selected RA titles using Dietz (1995) taxonomy for the syntactic structure of RA titles and Anthony's classifications for the compound-unit titles. They found some

discipline-specific differences in title components, mirroring academic conventions of title constructions of respective fields.

Title punctuation patterns

Punctuation has also attracted attention to see whether compound titles are correlated with higher citations. Jacques and Sebire (2010) and Buter and van Raan (2011), for example, advocated the use of compound structures including colons, while Hartley (2007) and Jamali and Nikzad (2011) noted that titles including a colon were followed by both fewer downloads and fewer citations. Regarding discipline-specific inclinations, Lewinson and Hartley (2005) argued that academicians are interested in using the colon as one type of punctuation, especially in the titles of single-authored papers in arts and social sciences. This finding aligns with Hartley (2007) findings that art and humanities research articles have longer titles compared with other disciplines. Lewinson and Hartley (2005) also reported discipline-specific differences in terms of structure (e.g., complete sentences preferred over sentences divided by colons), length (e.g., article titles are longer than book titles), and content (including a general subject, the methodology, a question, a precise theme, and the author's argument) of titles in different genres, such as books and RAs.

Dillon (1981) conducted a corpus study and investigated the use of the colon in journal articles titles. The study results showed that colon was more frequently observed in theoretical research journals, followed by empirical and pedagogical research journals, respectively. One year later, Dillon (1982) studied 1,150 journal article titles regarding the use of colons in three disciplines of education, psychology, and literary criticism over 100 years (1880–1980). The findings showed that colon use in articles developed gradually and steadily across the disciplines, initially in literacy criticism journals. Moreover, Kerans et al. (2020) worked on medical research titles, identifying the number of parts of the titles and their punctuation. They found 10 subtypes. Anthony (2001) found nine-word titles for research articles in computer science sub-disciplines. The findings revealed that title length noticeably varied within sub-disciplines. Moreover, it was disclosed that on average, 13% of the titles were two-unit titles, separated by a colon. The two most frequent rhetorical organizations observed in these compound titles were “name description” and “topic scope”. Cianflone (2010) analyzed a small corpus of 63 RA titles, including noun phrases, compounds, full-sentence declaratives, or questions from three reputable journals on veterinary medicine. These patterns were observed with the examples of compound sentences using punctuations like colon for indicating research methods in the second parts of the titles.

Although the literature review reveals a scant number of comparative empirical research articles on RA articles' syntactic structures and their rhetorical organizations written

by non-native authors in medicine and applied linguistics, especially in an EFL context, which is an often unnoticed aspect of academic discourse research, they play a key role in knowledge construction. In fact, they help readers notice and cite new research. Most large-scale analyses confirmed various syntactic structures of RA titles in different disciplines. Still, they did not investigate the differences, especially between these two fields of study in a cross-disciplinary course. Moreover, most studies suffice to report a descriptive analysis just reporting the frequencies and percentages of occurrences. However, inferential statistics is required to see if the cross-disciplinary differences in syntactic structures and rhetorical organizations are statistically significant and considered discipline-specific features.

The theoretical framework of the study

The comparison of the syntactic structures of titles in medicine and applied linguistics has been based on the continuum of hard (e.g., medicine) and soft (e.g., applied linguistics) disciplines, which was first presented by Hyland (2000). Their knowledge structures and intellectual inquiry manifested in the rhetorical conventions of disciplines differentiated them (Moattarian and Alibabae, 2015). The reason behind this recourse is that comparing the syntactic structures and rhetorical functions with these hard and soft science fields may help us better understand whether and how these structures are more geared toward hard or soft sciences. Moreover, this study analyzes the syntactic structures of RA titles in light of Dietz (1995), as cited in Bush-Lauer, (2000) taxonomy. This is attributed to the fact that this taxonomy captures different title lengths and focuses on different title styles and their classifications; consequently, it will give researchers a more comprehensive look at syntactic structures in titles. Figure 1 presents the framework.

As the repeated measuring of syntactic structures of multi-unit titles might make frequency counts and interpretation complicated and unclear, the researchers decided to analyze single-unit and multi-unit titles separately. This action aligns with Moattarian and Alibabae (2015) decision to avoid double-counting the same structures for different title styles. Thus, the statistical data analysis includes the analysis of the structures of single-unit and multi-unit RA titles. For the single-unit titles, Dietz (1995) taxonomy has been followed in which the syntactic features of length and style were identified. According to this taxonomy, the style of titles can be single or multiple. Analyzing the constituent components of the multi-unit RA titles, the researchers used Anthony (2001) taxonomy of compound constructions. The classification incorporates five categories, namely, name-description, description-name, topic-description, topic-scope, and topic-method combinations, closely examining the link

and association between title units. In terms of Dietz's taxonomy, single-unit title patterns were identified as nominal, verbal, prepositional, and adjectival/adverbial structures. The nominal titles are categorized as unmodified, pre-modified, post-modified, and pre- and post-modified grammatically. Verbal structures are indicated as either a full sentence or a gerund phrase, and there is no specific categorization specified for the prepositional and adjectival/adverbial structures in this taxonomy.

Significance of the study

The style manuals for writing scientific RA titles insufficiently specify what criteria should be followed for an appropriate title (Day, 1989; Ebel et al., 1993). There are some guidelines in the journals on how to write RA titles. Yet, considering original titles, we readily notice that they do not often follow the requirements for an appropriate title (Busch-Lauer, 2000). This issue makes classifying scientific literature in line with content areas relatively challenging by only relying on the titles. That is why the researchers study the titles and their constructions in scientific papers. As the literature review indicates, most corpus studies conducted tend to explore many article titles published in different journals with various disciplines. However, the studies are restricted to examining only a few features such as title length and punctuation use or, for example, colons like the studies conducted by Dillon (1981, 1982) and Lewinson and Hartley (2005). Or, if the study adopts a broader scope, the discipline investigated is limited, such as the research conducted by Gesuato (2008). Although the focus of this cross-disciplinary study revolves around a limited number of RA titles, the researchers pursued an in-depth analysis of the syntactic structures and rhetorical organization to have a more comprehensive, detailed, and in-depth look at the issue in an EFL context.

Besides, as the review of literature unveils, there has been a body of research denoting the differences between disciplines which bifurcate into natural and social sciences (Buxton and Meadows, 1977), biological and social sciences (Soler, 2007), hard and soft sciences (Nagano, 2015), etc. In this cross-disciplinary study, the logic behind the selection of titles from two distinct disciplines, namely, medicine and applied linguistics, was to divulge the discipline-specific differences in title components, length, and style, reflecting upon their academic conventions. The aim was to see how far the medical and non-medical titles might show the communicative and rhetoric differences and/or similarities characterized by the nature of their research. Such studies help one to cross over the boundaries of disciplinary research, integrate knowledge, and bring novelty to language research (Herieg, 2011). Furthermore, most studies in the literature reported only the counts and frequencies of the occurrences and did not examine the

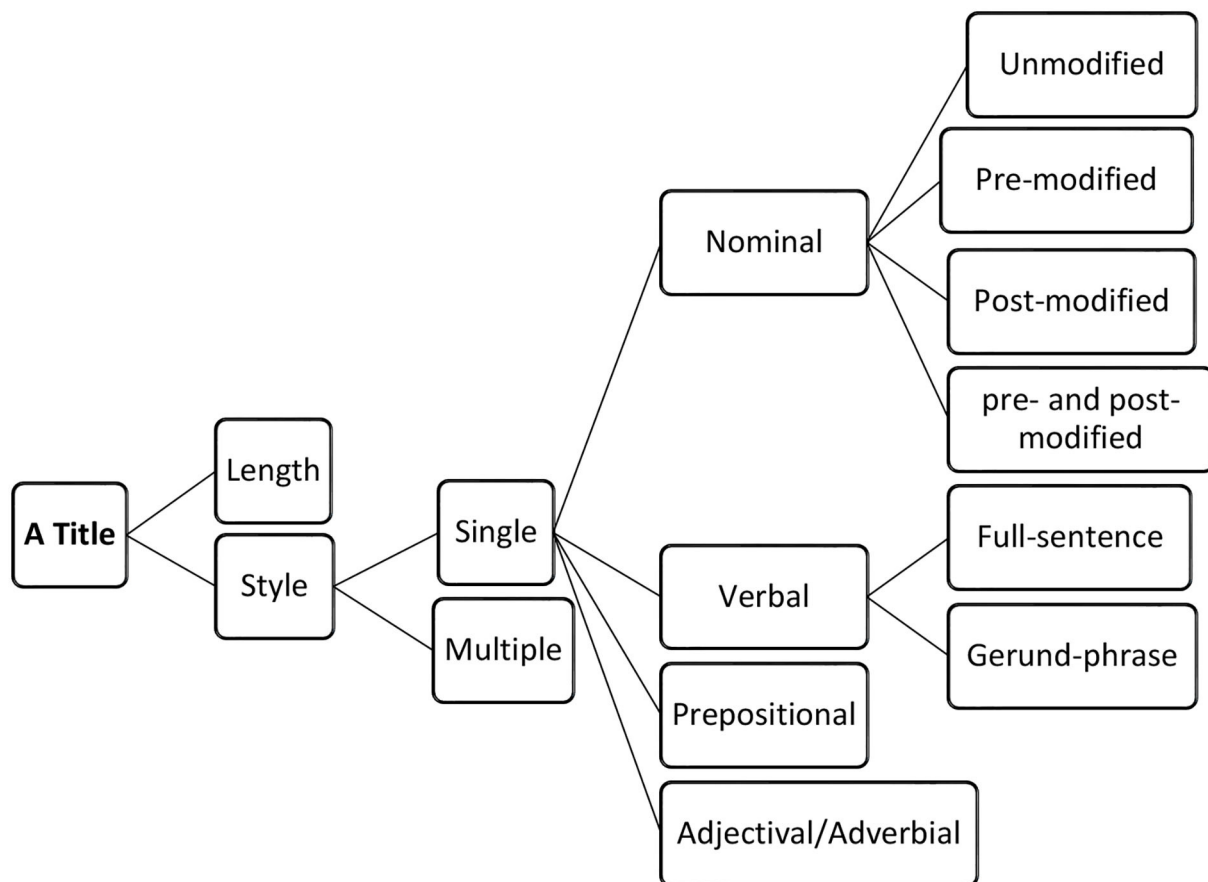


FIGURE 1

A framework adapted from Dietz (1995) taxonomy to analyze syntactic structures of RA titles.

significance of differences or the meaningful relationship between the intended structure and the discipline under investigation; therefore, it is difficult to see if cross-disciplinary studies, especially with small samples, have the statistical power for valid comparisons in any case (Kerans et al., 2016). As indicated earlier, many researchers studied different aspects of RAs. Although RA titles have a significant role in motivating or demotivating a reader to read a scientific article, there is a limited number of cross-disciplinary body of research articles on the syntactic structures of RA titles written in medicine and applied linguistics by non-native researchers in renowned English journals published in an EFL context, such as Iran.

Thus, the present study investigates the frequency and percentage of syntactic structures of RA titles written by Iranian researchers in two groups of disciplines, namely, medicine and applied linguistics, in high-quality and reputable Iranian research journals in English. The study explores any statistically significant difference between RA titles in these disciplines in terms of title length. Moreover, the study investigates whether there is a statistically meaningful relationship between various

syntactic structures and rhetorical combinations on the one hand and the two disciplines, medicine and applied linguistics, on the other hand. This study hopes to come up with conclusions to find syntactic structures and rhetorical organizations that scholars in those specialties frequently use to fulfill RA title requirements in reputable and renowned journals in each field, as mentioned earlier. Accordingly, the following research questions are raised:

1. Is there a statistically significant difference between medicine and applied linguistics RA titles in terms of title length?
2. Is there a statistically meaningful relationship between the two categories of RA title style (single and multiple units) and the two disciplines (medicine and applied linguistics)?
3. Is there a statistically meaningful relationship between syntactic components of single-unit RA titles (nominal, verbal, and prepositional) and the two disciplines (medicine and applied linguistics)?

4. Is there a statistically meaningful relationship between syntactic components of RA nominal structures (un-, pre-, post, and pre- and post-modifiers) and the two disciplines (medicine and applied linguistics)?
5. Is there a statistically meaningful relationship between syntactic components of verbal structures (full sentence and gerund phrase) and the two disciplines (medicine and applied linguistics)?
6. Is there a statistically meaningful relationship between the rhetorical organization of verbal structures (e.g., topic-scope, topic-method, and topic-description) and the two disciplines (medicine and applied linguistics)?

Methods

Research design

This corpus-based cross-disciplinary study investigates the syntactic structures and rhetorical organization of 200 RA titles in medicine (100 titles) and applied linguistics (100 titles). The authors selected reputable Iranian English journals indexed in Scopus. This study incorporated a framework developed by Dietz (1995) as cited in Busch-Lauer (2000) to analyze the title length, title styles, and title syntactic components of single-unit RA titles. Furthermore, the authors used a framework developed by Anthony (2001) to explore the rhetorical organization of multiple or compound RA titles. Although descriptive and inferential statistical analyses were carried out, the study is primarily quantitative. However, qualitative data analysis has contributed to more comprehensive views toward comparing the different syntactic structures and rhetorical functions of each structure in RA titles written by Persian authors of English scientific papers in these disciplines.

Materials

The materials in this study are research article titles in two significant fields of study, namely, medicine and applied linguistics, published in Iranian renowned English journals. First, the authors randomly selected two Iranian English journals in each discipline to collect data. Then, through random sampling, the authors selected 100 RA articles in each major: 50 RA titles from the *Journal of Research in Applied Linguistics*, 50 RA titles from *Language Related Research*, 50 RA titles from the *Iranian Journal of Medical Sciences*, and 50 RA titles from *Medical Journal of the Islamic Republic of Iran*. It is worth mentioning that these Iranian English journals were selected because they are authoritative and eminent journals indexed in Scopus. Furthermore, the authors selected the articles published in 2018–2021 to control and eliminate the effect of changes in language use during that time.

Data collection and analysis procedures

First, the authors applied the taxonomy adopted from Dietz (1995), as cited in Busch-Lauer (2000) to analyze the data. They incorporated this framework into the study to investigate the syntactic features of the RA titles as follows:

- 1) title length (number of words per title),
- 2) title style (single-unit or multi-unit structures), and
- 3) title unit components (single-unit syntactic structures).

The authors counted words as strings of letters preceded or followed by spaces or punctuation marks to know the title length. However, they considered capitalized abbreviations and hyphenated compounds a single word. For instance, they counted seven words in the title “Turn-Taking, Preference, and Face in Criticism Responses”. So, they considered the title constituent parts in analyzing the title style. Title styles according to the framework bifurcated into single-unit and multi-unit structures. Single-unit titles were made of a single phrase or sentence, while multi-unit structures were composed of two or more phrases or sentences separated by a colon. They were examined separately since the syntactic structures in these two styles differ. Dietz’s framework was employed to analyze the single-unit structures. This framework classifies single-unit titles into nominal, verbal, prepositional, and adjectival/adverbial syntactic constructions.

a) Nominal structures: A nominal title includes one or more nouns, often called head(s), with or without pre-modifiers and/or post-modifiers (Wang and Bai, 2007). The main section of the nominal phrase is the head. In Wang and Bai (2007) term, further specifications have grammatical or semantic relationships to the head (Richards et al., 1998). The four following titles can well reflect various nominal constructions of this study: Coercion and Construction Grammar (unmodified) Modified; Semiotic Analysis of Tejarat Bank Advertisements (pre-modified); Politeness and Impoliteness in Persian-speaking Youngsters’ Novels (post-modified); and Analysis of the Textual Cohesion of a Sonnet by Hafiz based on its Thematic Structure (pre- and post-modified).

b) Verbal structures: Gerund-phrase titles and full-sentence titles are two primary groups of verb phrase titles whose instances are indicated in the following titles: 9-cis-Retinoic Acid and 1,25-dihydroxy Vitamin D3 Improve the Differentiation of Neural Stem Cells into Oligodendrocytes through the Inhibition of the Notch and Wnt Signaling Pathways (full-sentence structure); and Estimating the Net Survival of Patients with Gastric Cancer in Iran in a Relative Survival Framework (gerund phrase structure).

c) Prepositional structures: When a title starts with a preposition followed by an object, it is called a prepositional title (e.g., “From Intertextual Relations to Intermediality Aspects”).

d) Adjectival or adverbial phrase titles: Adjectives and adverbs inform the readers of their general focus in the title. However, the authors did not observe any incidence of these types.

Anthony's classification of compound constructions was followed for the rhetorical analysis of constituent elements of multiple titles. According to this framework, multiple constructions are given as follows: a) Name–description titles bear a name or abbreviation in the first unit and describe it in the second part (e.g., “NIDDM: Noninsulin-Dependent Diabetes Mellitus”); b) description–name titles act on the reverse, that is, provide a description and then its name or abbreviation; c) topic–description titles introduce the main topic in the first unit and provide the specific description in the second part (e.g., “National Minimum Data Set for Antimicrobial Resistance Management: Toward Global Surveillance System”); d) topic–scope titles depict the primary topic and then the scope of the study (e.g., “Incorporating E-learning in teaching the English language to medical students: exploring its potential contributions” or “On the Validation of a Preliminary Model of Reading Strategy Using SEM: Evidence From Iranian ELT Postgraduate Students”); e) topic–method titles first bring the main topic into notice and then the study's research (e.g., “The Effect of Remote Ischemic Preconditioning on the Incidence of Acute Kidney Injury in Patients Undergoing Coronary Artery Bypass Graft Surgery: A Randomized Controlled Trial”).

The incidence of each structure was first identified, counted, and finally reported through frequencies and percentages. Subsequently, the comparison was made the syntactic and rhetorical structures of the RA titles in medicine and applied linguistics through a statistical procedure, namely, the chi-square test *via* SPSS, to see if a statistically significant relationship exists between the particular structure or rhetorical organization of each discipline.

Results and discussion

Title length

The first research question posed in this study revolves around the title length or the number of words per title. The question is whether there is a statistically significant difference between medicine and applied linguistics RA titles in terms of title length.

Table 1 shows that medical titles are longer than linguistic ones in terms of title words. There are 1,559 words in medicine titles and 1,291 words in linguistics ones. The average text length of RA titles in medicine is 15.59, whereas in applied linguistics the average is 12.91 words per title. However, linguistic titles capture minor variance or inconsistency.

Table 2 indicates the results of a *t*-test for equality of means in terms of the title length in medicine and applied linguistics. As the *p*-value is less than the significance level of 0.05 (i.e., $\text{sig} = 0.00 < 0.05$), there is a statistically significant difference between the two fields of medicine and applied linguistics. Therefore, the answer to the first research question is positive. As a result, the statistically meaningful difference between the two discipline title lengths should be considered. This finding corroborates other researchers' similar findings that title length could be captured as a cross-disciplinary distinct feature. The finding contradicts the results found by Hyland and Zou (2022) and aligns with the findings of Busch-Lauer (2000), Haggan (2004), Soler (2007), Nagano (2015), Moattarian and Alibabae (2015), Nagano (2015), Shahidpour and Alibabae (2017), Milojevic (2017), and Appiah et al. (2019), indicating that title length is not only contingent upon the disciplines but also becomes more pronounced in hard science disciplines such as medicine than in soft science fields like applied linguistics in this study. That is, hard disciplines have longer titles than soft ones. One reason might be that in hard science and, more

TABLE 1 Descriptive of title length across disciplines.

Disciplines	N	Minimum	Maximum	Sum	Mean		Std. Deviation	Variance
					Statistic	Std. Error		
Medicine	100	6	29	1,559	15.59	0.464	4.636	21.497
Applied linguistics	100	4	27	1,291	12.91	0.455	4.546	20.669
Total	200	4	29	2,850	14.25	0.338	4.773	22.781

TABLE 2 *T*-test for the equality of means.

t	df	Sig (2-tailed)	Mean difference	Std. error difference	95% Confidence interval of the difference	
					Lower	Upper
4.127	198	0.000	2.680	0.649	1.399	3.961
4.127	197.924	0.000	2.680	0.649	1.399	3.961

TABLE 3 Frequency and percentage of single and multi-unit titles in the two disciplines (chi-square test).

		Title style		Total	Sig
		Single unit	Multiple units		
Medicine	Count	58	42	100	0.557
	% within disciplines	58.0%	42.0%	100.0%	
	% within Title style	50.0%	50.0%	50.0%	
	% of Total	29.0%	21.0%	50.0%	
Applied linguistics	Count	58	42	100	
	% within disciplines	58.0%	42.0%	100.0%	
	% within Title style	50.0%	50.0%	50.0%	
	% of Total	29.0%	21.0%	50.0%	
Total	Count	116	84	200	
	% within disciplines	58.0%	42.0%	100.0%	
	% within Title style	100.0%	100.0%	100.0%	
	% of Total	58.0%	42.0%	100.0%	

specifically, in medical and clinical titles, there is a tendency toward elaborating on the opening statements or topics in terms of their scope, description, and methodology for the physicians. As a result, they sometimes draw conclusions based on merely looking at the RA titles (Goodman, 2000). More compound words have also lent themselves to longer titles in such disciplines. However, the point is that, unlike the previous research, this study once more ratifies the former findings in a new context, an Iranian EFL setting, where the authors are all non-native speakers.

Title style

The second feature captured in this study is title style, which is defined as the number of units in each title. According to the Dietz taxonomy, there are two style categories: single- and multi-unit RA titles. The single one is characterized by the occurrence of the title in one whole unit. The multi-unit or compound titles are those in which the title is manifested in more than one unit linked by an appropriate punctuation mark, commonly a colon (Nagano, 2015).

Table 3 displays the frequency and percentage of both single- and multi-unit titles that are unexpectedly similar in both disciplines. In both medicine and applied linguistics, single-unit title structures account for 58%, and multiple constructions constitute 42% of all. Therefore, in this corpus of 200 titles, no difference has been observed in terms of the title style. This finding is even evident within the title style in Table 3, where the percentage of occurrence (50%) is the same for each discipline. The chi-square test does not indicate any statistically meaningful relationship between the title style and the disciplines ($\text{sig} = 0.557 > 0/05$). Consequently, the answer to the second research question in this study is negative, that

is, there is no statistically meaningful relationship between the title style and the investigated disciplines. Unlike the previous studies, such as those conducted by Hartley (2007), Cheng et al. (2012), Moattarian and Alibabae (2015), Nagano (2015), Shahidpour and Alibabae (2017), and Hyland and Zou (2022), the findings of this study do not verify the nature of discipline-specific feature of title styles in medicine and applied linguistics. Most of the literature (e.g., the studies mentioned previously) lends itself to using multiple units more in the soft fields such as humanities (e.g., applied linguistics) than in sciences. However, the results in this study contradict those findings as no meaningful relationship was found. This striking similarity might be attributed to the fact that over time, the non-native Iranian physicians' and linguists' knowledge regarding the rhetorical organizations of the RA titles has changed toward a similar style. This finding may suggest that title style is no longer a discipline-specific feature in these two majors. However, more comprehensive corpus-based diachronic research should be conducted to investigate the development of patterns over time.

Syntactic components of single units

Dietz's taxonomy divides single-unit titles into nominal, verbal, prepositional, and adjective/adverb structures. As Table 4 indicates, the use of nominal, verbal, and prepositional structures in the field of medicine accounts for 89.7, 8.6, and 1.7%, respectively. In applied linguistics, the use of nominal constructions reduces to 87.9%, while verbal structures are raised to 12.1%. There is no incidence of prepositional structures in this discipline.

TABLE 4 Frequency and percentage of title components or syntactic structures in single-unit titles in the two disciplines (chi-square test).

		Nominal	Verbal	Prepositional	Total	Sig
Medicine	Count	52	5	1	58	0/511
	% within disciplines	89.7%	8.6%	1.7%	100.0%	
	% within syntactic structures	50.5%	41.7%	100.0%	50.0%	
	% of Total	44.8%	4.3%	0.9%	50.0%	
Applied linguistics	Count	51	7	0	58	
	% within disciplines	87.9%	12.1%	0.0%	100.0%	
	% within syntactic structures	49.5%	58.3%	0.0%	50.0%	
	% of Total	44.0%	6.0%	0.0%	50.0%	
Total	Count	103	12	1	116	
	% within disciplines	88.8%	10.3%	0.9%	100.0%	
	% within syntactic structures	100.0%	100.0%	100.0%	100.0%	
	% of Total	88.8%	10.3%	0.9%	100.0%	

As the findings represent, the most frequently used structure within single-unit RA titles in both disciplines is the nominal structure, followed by less observed constructions, that is, verbal and prepositional phrases. Moreover, there are no adjectival or adverbial structures in medicine and applied linguistics corpus. As the chi-square test in Table 4 reveals ($\text{sig} = 0.511 > 0.05$), despite the observed differences between medicine and applied linguistics in terms of nominal, verbal, and prepositional structures, there was no meaningful relationship between the disciplines and such structures. Therefore, the answer to the third research question is negative. For the Iranian authors in these majors, the structures mentioned here in RA titles have no meaningful relationship with their field of study. As most of the studies are cross-disciplinary, the authors have to see whether the differences across the majors regarding such syntactic constructions can be meaningfully attributed to the disciplines under investigation or not. As the literature review discloses, the most recurrent and prevalent syntactic construction corresponded to the nominal group structure, followed by verbal and prepositional structures (e.g., Haggan, 2004; Soler, 2007; Wang and Bai, 2007; Cheng et al., 2012; Moattarian and Alibabae, 2015; Nagano, 2015; Doykova, 2016; Shahidpour and Alibabae, 2017). The reason can be attributed to the fact that even the authors (here in this study, the non-native RA writers) are inclined to summarize the essence of their studies and pack more information through nominal construction. As Soler (2007) and Wang and Bai (2007) declared, nominal structures can bring about efficient information use through different modifiers pre- and/or post-ones. In nominal structures, heads, which are nouns, present the focus of the study, and modifiers provide more detailed specifications to better elaborate on the intended issue (Wang and Bai, 2007). That is why the next section is devoted to analyzing different subsections and specifications of the nominal structures in research articles.

Nominal structures

Table 5 illustrates the frequency and percentage of unmodified (UM), pre-modified (PRM), post-modified (POM), and pre- and post-modified (PPM) structures in RA nominal titles in medicine and applied linguistics.

The findings reveal that PPM accounts for most of the RA titles in medicine and applied linguistics (75.0 and 72.5%, respectively). The next dominant feature is the POM, which occurred in medicine and applied linguistics (19.2 and 17.6%, respectively). The other two nominal components, namely, PRM and UM, were the least occurred structures (medicine, 3.8% and applied linguistics, 7.8% for pre-modified construction) and (medicine, 1.9% and applied linguistics, 2.0% in terms of unmodified structures). The results reveal that medical nominal titles override linguistic nominal titles in terms of PPM and POM, although the difference is not considerable. Regarding PRM and UM, the linguistic nominal constructions take over the medical ones. However, as the chi-square test results uncover ($\text{sig} = 0.858 > 0/05$), there is no meaningful relationship between the use of the four categories of nominal construction by the non-native Iranian RA authors and the two disciplines. Therefore, the answer to the fourth research question is not positive. This study's findings on the frequency of the most dominant nominal PPMs, which accounts for 73.8% of all the cases, corroborate other studies (e.g., Fortanet Gómez et al., 1998) but are in contrast to those of Haggan (2004), Wang and Bai (2007), Cheng et al. (2012), Moattarian and Alibabae (2015); and Shahidpour and Alibabae (2017) who found post-modifiers as the most frequently used constructions. However, this research is in line with the mentioned studies in that the least frequent nominal title is an unmodified nominal construction that constitutes only 1.9% of all the nominal structures. The reason can be attributed to the tendency among the RA authors to provide more details and specifications, as Wang and Bai (2007) once stated, to better elaborate on the general focus of the

TABLE 5 Frequency and percentage of syntactic structures of nominal constructions in single-unit titles in the two disciplines (chi-square test).

		Syntactic structures				Total	Sig
		UM	PRM	POM	PPM		
Medicine	Count	1	2	10	39	52	0.858
	% within disciplines	1.9%	3.8%	19.2%	75.0%	100.0%	
	% within syntactic structures of nominal	50.0%	33.3%	52.6%	51.3%	50.5%	
	% of Total	1.0%	1.9%	9.7%	37.9%	50.5%	
Applied linguistics	Count	1	4	9	37	51	
	% within disciplines	2.0%	7.8%	17.6%	72.5%	100.0%	
	% within syntactic structures of nominal	50.0%	66.7%	47.4%	48.7%	49.5%	
	% of Total	1.0%	3.9%	8.7%	35.9%	49.5%	
Total	Count	2	6	19	76	103	
	% within disciplines	1.9%	5.8%	18.4%	73.8%	100.0%	
	% within syntactic structures of nominal	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	1.9%	5.8%	18.4%	73.8%	100.0%	

studies and have a more comprehensive view of the content of the papers. However, the current study shows that the frequency of these specifications is not necessarily linked to the disciplines under investigation. Therefore, these nominal components are not distinctive in medicine and applied linguistics.

Verbal structure

The second most recurrent syntactic structures employed by both disciplines are verbal structure. The full-sentence (FS) and gerund phrase (GPH) titles constitute RA verbal titles.

Table 6 shows that the incidence of FS in medicine is 60.0%, while this structure is absent in applied linguistic titles. The Iranian researchers in linguistics have not used full sentences. The results show that the use of GPHs accounts for all the structures used in applied linguistics. In other words, Iranian authors preferred to use gerundive phrases as the only verbal structure in writing their RA titles. The use of GPHs in medicine constitutes 40.0% of all the verbal titles in this discipline. The chi-square test results ($\text{sig} = 0.045 < 0.05$) reveal a statistically meaningful relationship between the use of FS and GPH verbal bifurcation, on the one hand, and the studied disciplines, on the other hand. It can, therefore, be concluded that these two verbal syntactic titles can be considered distinctive features in medical and linguistic RA titles. The findings are in accord with that reported by Soler (2007), who suggested that FS verbal titles are typical of medicine, biology, and biochemistry. Still, there was no incidence of full-sentence structures in linguistics and psychology titles.

Moreover, the incidence of full-sentence titles was reported by Cianflone (2010) in veterinary medicine and by Haggan (2004) concerning science titles. However, the findings of this research are in contrast to those of some studies. For example, Cheng et al. (2012) found full-sentence titles in

linguistics. Moattarian and Alibabae (2015) also came up with a disciplinary peculiarity relating to the incidence of full-sentence titles in dentistry RA titles, which suggested a marked contrast to the titles of applied linguistics and civil engineering. Hyland (2004) stated that it is to express claims through elaborate exposition and considerable tentativeness in social sciences. In biological sciences, introducing ideas through full-sentence constructions seems easier (cited in Soler, 2007). Moattarian and Alibabae (2015) mentioned that the authors could use the full-sentence construction to present their studies' results decisively and synthetically in one single structure. Yet, other studies such as Shahidpour and Alibabae (2017) stated that English linguistics authors used full-sentence titles more than V-ing phrases compared with researchers in hard sciences.

Concerning gerund or, in some studies, V-ing titles, the present study findings lend to those found by Moattarian and Alibabae (2015) in that such titles are more observed in applied linguistics than in medical sciences like dentistry. One reason can be attributed to briefly highlighting the research process or activity or presenting the primary topic and research process concurrently. According to Wang and Bai (2007), RA authors use such construction to enhance readers' conciseness and attractiveness.

Prepositional, adjective, and adverb structures

Prepositional structures are rarely used as the initiating constructions in the RA titles. This study did not show the incidence of such structures and adjectival or adverbial titles. This finding strongly corroborates other studies such as those conducted by Haggan (2004), Cheng et al. (2012), Moattarian and Alibabae (2015), and Shahidpour and Alibabae (2017). Accordingly, these structures are the least frequent ones in

TABLE 6 Frequency and percentage of syntactic structures of verbal constructions in single-unit titles in the two disciplines (chi-square test).

		q6		Total	Sig
		FS	GPH		
Medicine	Count	3	2	5	0.045
	% within disciplines	60.0%	40.0%	100.0%	
	% within syntactic structures of verbal	100.0%	22.2%	41.7%	
	% of Total	25.0%	16.7%	41.7%	
Applied linguistics	Count	0	7	7	
	% within disciplines	0.0%	100.0%	100.0%	
	% within syntactic structures of verbal	0.0%	77.8%	58.3%	
	% of Total	0.0%	58.3%	58.3%	
Total	Count	3	9	12	
	% within disciplines	25.0%	75.0%	100.0%	
	% within syntactic structures of verbal	100.0%	100.0%	100.0%	
	% of Total	25.0%	75.0%	100.0%	

TABLE 7 Rhetorical combinations in multi-unit titles in the two disciplines (chi-square test).

		Multiple units					Total	Sig
		TS	TM	TD	TM+TS	TM+TD		
Medicine	Count	5	23	6	7	1	42	0.001
	% within disciplines	11.9%	54.8%	14.3%	16.7%	2.4%	100.0%	
	% within in multiple unit	26.3%	79.3%	30.0%	46.7%	100.0%	50.0%	
	% of Total	6.0%	27.4%	7.1%	8.3%	1.2%	50.0%	
Applied linguistics	Count	14	6	14	8	0	42	
	% within disciplines	33.3%	14.3%	33.3%	19.0%	0.0%	100.0%	
	% within in multiple unit	73.7%	20.7%	70.0%	53.3%	0.0%	50.0%	
	% of Total	16.7%	7.1%	16.7%	9.5%	0.0%	50.0%	
Total	Count	19	29	20	15	1	84	
	% within disciplines	22.6%	34.5%	23.8%	17.9%	1.2%	100.0%	
	% within in multiple unit	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	22.6%	34.5%	23.8%	17.9%	1.2%	100.0%	

single-unit titles. The justification may be that they are not informative enough (Shahidpour and Alibabae, 2017).

Rhetorical organizations of multi-unit titles

Table 7 displays the counts and percentages of the occurrence of different rhetorical combinations, namely, topic-scope (TS), topic-method (TM), topic-description (TD), topic-method + topic-scope (TM+TS), and topic-method + topic-description (TM+TD), in multiple or compound RA titles in two majors of medicine and applied linguistics.

As the findings indicate, the most frequently employed combination in medicine RA titles is TM (54.8%). The lowest incidence is related to the TM+TD combination (2.4%). In applied linguistics, most of the rhetorical combinations are

TS and TD (both equally make up 33.3% of all cases), and the least rhetorical organization is the TM combination. There is no incidence of TM+TD combination in this discipline; however, there is only one occurrence of TM+TD. The most to the least dominant combinations used by the Iranian authors in both disciplines are TM (34.5%), TD (23.8%), TS (22.6%), TM+TS (17.9%), and TM+TD (1.2%). The chi-square test results indicate ($\text{sig} = 0.001 < 0.05$) that there is a meaningful relationship between different rhetorical categories of multi-unit titles and the disciplines under study.

Consequently, these combinations constitute distinctive disciplinary features in medicine and applied linguistics. Thus, the answer to the last research question is positive. Moattarian and Alibabae (2015) believed that topic-method combination is common and prevalent in medical titles, while

name–description and description–name combinations are not. This claim aligns with the current study findings. The often employed TM combination might be due to the physicians' willingness to introduce a general topic, including a medical problem, and then present how the study is conducted to better capture various aspects and methods through which that medical problem can be solved.

Moreover, this study confirms the findings revealed by Moattarian and Alibabae (2015). They claimed that applied linguistics RA titles mostly had a topic–scope organization. According to these authors, one reason for this is the linguists' tendency to use such comprehensive RA titles, giving the readers the required information regarding the study context, samples, and site. Also, the findings align with Anthony (2001) study, which indicates that hard sciences like computer engineering did not frequently use the topic–description combination. However, the current study findings conflict with Anthony's results that name–description and topic–scope are the frequently used combinations in such sciences. Shahidpour and Alibabae (2017) also found the topic–method combination as the most frequent rhetorical function in psychiatry, and topic–description as the dominant combination in linguistics, which can ratify the findings of this study.

Conclusion

This corpus-based cross-disciplinary study investigated the syntactic structures and rhetorical combinations of RA titles in terms of stylistic conventions dominant in two disciplines on the soft–hard continuum presented by Hyland (2000), namely, medicine as a hard science and applied linguistics as a soft science. The RA titles were randomly selected from four reputable Iranian English journals indexed in Scopus. The authors were all Iranian researchers in the target disciplines, and this has provided the study with a different fresh perspective to look at the same problem of title construction. This study presents the counts and percentages of the occurrences in each syntactic and rhetorical construction. The point is that, unlike the other similar studies, the authors used a chi-square test to see whether these frequencies led to a meaningful relationship between the intended title structure and the discipline to justify better if that syntactic or rhetorical construction can be defined as a discipline-specific feature. Findings through the application of the *t*-test revealed that there exists a statistically significant difference between the two fields of medicine and applied linguistics in terms of title length, with medical titles longer than the linguistic ones. However, the frequency and percentage of both single- and multi-unit RA titles were exceptionally similar in both disciplines.

Consequently, the chi-square test did not indicate any statistically meaningful relationship between the title style

and the investigated disciplines. Concerning the syntactic components of single-unit RA titles, the most frequently used structure was the nominal construction, followed by verbal titles. The prepositional ones are the last with the least frequency. However, the chi-square test results indicate these features are not discipline-specific since no meaningful relationship has been reported. The most to the least frequently employed syntactic components of a nominal structure in both disciplines were PPM, POM, PRM, and UM, with medical titles taking precedence over linguistic titles in all these aspects. However, the differences were not statistically meaningful. Medical titles included two syntactic structures of verbal constructions, with full sentences outweighing gerund phrases. Yet, in linguistic RA titles, the dominant structure was gerund phrases (100%), without any incidence of full-sentence titles.

Regarding the rhetorical components of multi-unit RA titles, medicine prioritized linguistic titles using TM, and applied linguistics titles included TS and TD rhetorical organizations. However, two new rhetorical combinations were recognized in this study, namely, TM+TS and TM+TD, while the former outnumbered the latter in both disciplines. Furthermore, the chi-square test results defined each discipline's distinctive rhetorical categories.

Consequently, the scrutiny of these findings indicates the inclination of the authors in each field of study to use some particular RA title structures as their academic conventions. Some but not all structures may be more geared to a specific discipline. An RA title's communicative and rhetorical effectiveness can mirror the essence of research in that discipline. Such discipline-induced conventions can develop English for specific purposes, especially in academic writing courses in both fields, so novice writers, especially the non-native ones in EFL contexts like Iran where exposure to English is limited, will become more conscious of syntactic structures and rhetorics for title design. Cheng et al. (2012) claimed that the title has a critical role in showing the priority and distinction of the article over diverse research done. For instance, Goodman (2000) believed that RA titles sometimes have a remarkable role in physicians' clinical decisions.

This study can bear some pedagogical implications for ESP reading and writing teachers to teach various title structures so that non-native medical and language students can better satisfy the discipline's rhetorical needs and title requirements. The cross-disciplinary studies give the researchers and teachers better insights into the potential differences across various disciplines' conventions, strengths, and weaknesses. Consequently, the ESP syllabi can be reformulated to encapsulate valuable instructions on how to write the titles in each field of study more efficiently. Along with its merits, this study had some specific limitations. The corpus was restricted in terms of the disciplines investigated. More disciplines with larger corpora taken from more

authoritative Iranian journals in English can enrich our studies, bringing more confidence in the generalizability of results. The reason behind selecting these four reputable Iranian English journals was that all of them have been indexed in Scopus. However, it is worth noting that greater diversity should not become an obstacle in performing in-depth analysis. Besides, variety in corpus development, methodology, or problem statement should not lead to difficulty in interpreting the results (Kerans et al., 2016). More diachronic studies are also suggested to better reflect the changes over time to be familiar with the recent trends in title construction. Furthermore, the study of titles in different genres such as conference paper titles, case report titles, book titles, and dissertation titles can enhance the quality of such studies. Hence, further research is recommended to be conducted at a greater breadth while not falling short in terms of depth and richness.

Data availability statement

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

References

- Anthony, L. (2001). Characteristic features of research article titles in computer science. *IEEE Trans. Prof. Commun.* 44, 187–194. doi: 10.1109/47.946464
- Appiah, K. R., Ankomah, C., Yaw Osei, H., and Hattoh-Ahiaduvor, T. (2019). Structural Organisation of research article titles: A comparative study of titles of business, gynaecology and law. *Adv. Lang. Liter. Stud.* 10, 145–54. doi: 10.7575/aiaa.v.10n.3p.145
- Archibald, A. J. B. (2005). *Dissertation Title: A Cross-Disciplinary Study? Unpublished Manuscript*. Singapore: National University of Singapore.
- Archibald, A. J. B. (2017). A linguistic analysis of conference titles in applied linguistics. *Int. J. For. Lang. Teach. Res.* 5, 11–25. doi: 10.4236/jamp.2017.58127
- Bhatia, V. J. (2008). Genre analysis, ESP and professional practice. *English Specific Purp.* 27, 161–174. doi: 10.1016/j.esp.2007.07.005
- Busch-Lauer, I. (2000). “Titles in English and German research papers in medicine and linguistics,” in A. Trosborg (Ed.), *Analyzing Professional Genres*, (Amsterdam: John Benjamins). p.77–97. doi: 10.1075/pbns.74.08bus
- Buter, R. K., and van Raan, A. F. J. (2011). Non-alphanumeric characters in titles of scientific publications: An analysis of their occurrence and correlation with citation impact. *J. Inform.* 5, 608–617. doi: 10.1016/j.joi.2011.05.008
- Buxton, A. B., and Meadows, A. J. (1977). The variation in the information content of titles of research papers with time and discipline. *J. Document.* 33, 46–52. doi: 10.1108/eb026633
- Cheng, S. W., Kuo, C., and Kuo, C. (2012). Research article titles in applied linguistics. *J. Acad. Lang. Learn.* 6, A1–A14. doi: 10.4236/jsea.2009.25050
- Cianflone, E. (2010). Scientific titles in Veterinary Medicine research papers. *English Specific Purp. World* 9, 1–8.
- Day, R. A. (1989). *How to Write and Publish a Scientific Paper*. 3rd edition. Cambridge: Cambridge University Press.
- Dietz, G. (1995). *Titel wissenschaftlicher Texte*. (Forum für Fachsprachen-Forschung 26). Tübingen: Narr
- Dillon, J. T. (1981). Functions of the colon: an empirical test of scholarly character. *Educ. Res. Quart.* 5, 71–75.
- Dillon, J. T. (1982). In pursuit of the colon: a century of scholarly progress: 1880–1980. *J. Higher Educ.* 53, 93–99. doi: 10.1080/00221546.1982.11780427
- Doykova, I. (2016). Title structures in research articles (EMP). *Int. J. Knowl.* 14, 770–774. Available online at: <https://www.researchgate.net/publication/315760130>
- Dudley-Evans, T. (1984). A Preliminary Investigation of The Writing of Dissertation Titles in The ESP Classroom. Methodology, Materials, Expectations. *Exeter Linguist. Stud Exeter.* 7, 40–6.
- Dudley-Evans, T. (1994). “Genre analysis: An approach to text analysis for ESP” in M. Coulthard (Ed.), *Advances in written text analysis*, (London: Routledge). p. 219–228.
- Ebel, H. F., Bliefert, C., and Avenarius, H. J. (1993). *Schreiben und Publizieren in der Medizin*. Weinheim, New York, Basel, Cambridge, Tokyo: VCH.
- Fortanet Gómez, I., Coll García, J. F., and Palmer Silveira, J. C., and Posteguillo Gómez, S. (1997). “The writing of titles in academic research articles,” in Chamorro R. M., Navarrete A. R. (eds.) *Lenguas aplicadas a las ciencias y la tecnología. Aproximaciones*, Cáceres, (Spain: Universidad de Extremadura, Servicio de Publicaciones) p. 155–158.
- Fortanet Gómez, I., Posteguillo Gómez, S., Coll García, J. F., and Palmer Silveira, J. C. (1998). “Linguistic analysis of research article titles: Disciplinary variations,” in I. Vázquez and I. Guillén (Eds.), *Perspectivas pragmáticas en lingüística aplicada*, (Zaragoza, Spain: Ambar Editorial). p. 443–447.
- Gesuat, S. (2008). “Encoding of information in titles: Practices across four genres in Linguistics,” In C. Taylor (Ed.). *Ecolingua: The role E-Copora in Translation and Language Learning*, (Trieste: EUT). p. 127–157.
- Goodman, N. W. (2000). Survey of active verbs in the titles of clinical trials reports. *Br. Med. J.* 320, 914–915. doi: 10.1136/bmj.320.7239.914
- Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attraction. *J. Pragm.* 36, 293–317. doi: 10.1016/S0378-2166(03)00090-0
- Hamp-Lyons, L. (1987). *Study Writing: A Course in Written English for Academic and Professional Purposes*. Cambridge, U.K.: Cambridge University Press.

Author contributions

SM conducted data collection, data analysis, and prepared the first draft of the manuscript. RK revised the first draft semantically and syntactically and prepared the manuscript for submission. 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.

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.

- Hartley, J. (2005). To attract or to inform: What are titles for? *J. Tech. Writ. Commun.* 35, 203–213. doi: 10.2190/NV6E-FN3N-7NGN-TWQT
- Hartley, J. (2007). Planning that title: Practices and preferences for titles with colons in academic articles. *Libr. Inf. Sci. Res.* 29, 553–568. doi: 10.1016/j.lisr.2007.05.002
- Herieg, C. (2011). “Cross-disciplinary approaches to the English language: Introduction,” in T. Popescu, R. Pioariu, and C. Herteg, (Eds.), *Cross-Disciplinary Approaches to the English Language: Theory and Practice*, (Newcastle: Cambridge Scholars Publishing). p. 2–5.
- Hyland, K. (2000). *Disciplinary Discourses: Social Interactions in Academic Writing*. London: Longman.
- Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *J. Second Lang. Writ.* 13, 133–151. doi: 10.1016/j.jslw.2004.02.001
- Hyland, K., and Zou, H. (2022). Titles in research articles. *J. English Acad. Purp.* 56, 1–13. doi: 10.1016/j.jeap.2022.101094
- Jacques, T. S., and Sebire, N. J. (2010). The impact of article titles on citation hits: An analysis of general and specialist medical journals. *JRSM Short Rep.* 1, 1–5. doi: 10.1258/shorts.2009.100020
- Jalilifar, A. R., Hayati, A. M., and Namdari, N. (2012). A comparative study of research article discussion sections of local and international applied linguistic journals. *J. Asia TEFL* 9, 1–29.
- Jamali, H. R., and Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics* 88, 653–661. doi: 10.1007/s11192-011-0412-z
- Kerans, M. E., Marshall, J., Murray, A., and Sabate, S. (2020). Research article title content and form in high-ranked international clinical medicine journals. *English Specific Purp.* 60, 127–139. doi: 10.1016/j.esp.2020.06.001
- Kerans, M. E., Murray, A., and Sabatè, S. (2016). Content and phrasing in titles of original research and review articles in 2015: range of practice in four clinical journals. *Publications* 4, 1–22. doi: 10.3390/publications4020011
- Lewinson, G., and Hartley, J. (2005). What’s in a title? Number of words and the presence of colons. *Scientometrics* 63, 341–356. doi: 10.1007/s11192-005-0216-0
- Martin, P., and León Pérez, I. K. (2014). Convincing peers of the value of one’s research: A genre analysis of rhetorical promotion in academic texts. *English Specific Purp.* 34, 1–13. doi: 10.1016/j.esp.2013.09.002
- Martin, P. M. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English Specific Purp.* 22, 25–43. doi: 10.1016/S0889-4906(01)00033-3
- Méndez, D. I., and Alcaraz, M. Á. (2017). Titles of scientific letters and research papers in astrophysics: a comparative study of some linguistic aspects and their relationship with collaboration issues. *Adv. Lang. Liter. Stud.* 8, 128–139. doi: 10.7575/aiaac.all.v.8n.5p.128
- Milojevic, S. (2017). The length and semantic structure of article titles- evolving disciplinary practices and correlations with impact. *Front. Res. Metr. Anal.* 2, 1–10. doi: 10.3389/frma.2017.00002
- Moattarian, A., and Alibabae, A. (2015). Syntactic structures in research article titles from three different disciplines: Applied Linguistics, Civil Engineering, and Dentistry. *J. Teach. Lang. Skills* 7, 27–50. doi: 10.22099/jtls.2015.3530
- Nagano, L. R. (2015). Research article titles and disciplinary conventions: A corpus study of eight disciplines. *J. Acad. Writ.* 5, 133–144. doi: 10.18552/joaw.v5i1.168
- Paiva, C. E., Lima, J. P., and Paiva, B. S. (2012). Articles with short titles describing the results are cited more often. *Clinics* 67, 509–513. doi: 10.6061/clinics/2012(05)17
- Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System* 30, 479–497. doi: 10.1016/S0346-251X(02)00050-7
- Richards, J. C., Platt, J., and Platt, H. (1998). *Longman Dictionary of Language Teaching and Applied Linguistics*. Hong Kong: Pearson.
- Salager-Meyer, F., and Alcaraz Ariza, M. A. (2013). Titles are “serious stuff”: a historical study of academic titles. *JAH* 4, 257–271. Available online at: <http://hdl.handle.net/10045/35747>
- Shahidpour, V., and Alibabae, A. (2017). Syntactic structures and rhetorical functions of electrical engineering, psychiatry, and linguistics research article titles in English and Persian: A cross-linguistic and cross-disciplinary study. *J. Teach. Lang. Skills (JTLS)* 36, 145–175. doi: 10.22099/jtls.2017.4066
- Soler, V. (2007). Writing titles in science: An exploratory study. *English Specific Purp.* 26, 90–102. doi: 10.1016/j.esp.2006.08.001
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Thelwall, M. (2017). Avoiding obscure topics and generalising findings produces higher impact research. *Scientometrics* 110, 307–320. doi: 10.1007/s11192-016-2159-z
- Wang, Y., and Bai, Y. (2007). A corpus-based syntactic study of medical research article titles. *System* 35, 388–399. doi: 10.1016/j.system.2007.01.005
- Yitzhaki, M. (1994). Relation of title length of journal articles to number of authors. *Scientometrics* 30, 321–332. doi: 10.1007/BF02017231



OPEN ACCESS

EDITED BY

Douglas F. Kauffman,
Medical University of the Americas – Nevis,
United States

REVIEWED BY

Shirli Werner,
Hebrew University of Jerusalem, Israel
Thomas P. Gumpel,
Hebrew University of Jerusalem, Israel

*CORRESPONDENCE

Alexander Röhm
alexander.roehm@udo.edu

SPECIALTY SECTION

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

RECEIVED 01 April 2022

ACCEPTED 05 October 2022

PUBLISHED 26 October 2022

CITATION

Röhm A, Grengel M, Möhring M,
Zensen-Möhring J, Nellen C and
Hastall MR (2022) Applying the mixed-
blessings model and labeling theory to
stigma in inclusive education: An
experimental study of student and trainee
teachers' perceptions of pupils with ADHD,
DLD, and intellectual disability.
Front. Psychol. 13:910702.
doi: 10.3389/fpsyg.2022.910702

COPYRIGHT

© 2022 Röhm, Grengel, Möhring, Zensen-
Möhring, Nellen and Hastall. 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.

Applying the mixed-blessings model and labeling theory to stigma in inclusive education: An experimental study of student and trainee teachers' perceptions of pupils with ADHD, DLD, and intellectual disability

Alexander Röhm^{1*}, Michelle Grengel¹, Michéle Möhring¹,
Johannes Zensen-Möhring², Cosima Nellen¹ and Matthias R.
Hastall¹

¹Qualitative Research Methods and Strategic Communication for Health, Inclusion, and Participation, Department of Rehabilitation Sciences, TU Dortmund University, Dortmund, Germany, ²Intellectual Disabilities in Rehabilitation and Education, Department of Rehabilitation Sciences, TU Dortmund University, Dortmund, Germany

Institutional and individual stigmatization represent major barriers that prevent children with disabilities from accessing education. It can be presumed that children with disabilities are labeled as such even in inclusive educational settings and that teachers' attitudes toward inclusive education and children with disabilities play a crucial role in this context. Against this background, the present study aims to (a) apply and conceptualize the mixed-blessings model in the context of stigma-related reactions to children's disability labels in inclusive education and (b) shed light on the causal attributions of teachers that underlie stigma-related attitudes toward children with various disabilities. A 3x2x2x2 online experiment examined the ways in which disability-specific causes and symptoms, the type of disability in question, the children's sex, and efficacy cues regarding educational efforts affect future teachers' attitudes toward and expectations of inclusive education as well as their social distance toward children with disabilities. The participants in this experiment were $N=605$ German student and trainee teachers representing different types of teaching professions. A multivariate analysis of variance (MANOVA) revealed that, in particular, the cause attributed to the disability, the depicted type of disability and the probability of learning success led to changes in attitudes. Respondents' teaching self-efficacy and their status as students or trainees emerged as moderators of the effect of pupils' type of disability. As a result, teacher education and training as well as communication regarding pupils with disabilities require a high degree of sensitivity to disability-specific and efficacy-related cues to prevent (accidental) professional or institutional stigmatization.

KEYWORDS

stigma, inclusive education, ADHD, DLD, intellectual disability, self-efficacy, mixed-blessings model, labeling theory

Introduction

Although inclusive education “constitutes an international policy imperative that promotes the rights of disabled children to be educated alongside their peers in mainstream classrooms” (Liasidou, 2012, p. 168), institutional and individual stigmatization remain major barriers that prevent children with disabilities from accessing education (Cooney et al., 2006; Scior et al., 2012). In the German school system, for instance, pupils with disabilities must be diagnosed and assigned special educational needs in order to receive support in accordance with their individual conditions. More than half of pupils with identified educational needs attend special education schools, which are often separated from ordinary schools in terms of both space and content, featuring different didactic concepts and curricula. Nonetheless, according to a recent report by the German *Conference of Ministers of Education* (Kultusministerkonferenz; KMK, 2022), in 2020, approximately 44.5% of such pupils attended mainstream schools with the aim to be taught alongside pupils without disabilities in an inclusive educational setting.

Accordingly, it can be presumed that children with disabilities continue to be labeled in terms of their disability, even in inclusive educational settings. Such labelling is likely to affect teachers' reactions to them as well as their interactions with them, which can result in stigmatization (Caslin, 2021). According to Link and Phelan's (2001) *labeling theory*, stigmatization emerges via a social process in which “elements of labeling, stereotyping, separation, status loss and discrimination co-occur in a power situation that allows them to unfold” (p. 367). In this regard, teachers' attitudes toward inclusive education and children with disabilities play a crucial role (Avramidis and Norwich, 2002; de Boer et al., 2011; Röhm et al., 2018) and are often shaped by information regarding and attributions of causal explanations to a disability (e.g., Lebowitz et al., 2016; Zensen and Röhm, 2021). Such information can be the result of personal experiences but can also be drawn from mass media and social media sources (Röhm et al., 2018). Building on media effects research and Zillmann and Brosius' (2000) *exemplification theory*, single-case descriptions (i.e., exemplars) such as case vignettes are known to influence recipients' attitudes toward certain issues (e.g., inclusive education and children with disabilities; Röhm et al., 2018). Such exemplars are perceived as a typical representative of the whole group (e.g., children with disabilities in general), and attitudes toward them (e.g., social distance) are thus generalized to the whole group (Zillmann, 2006).

Following the *mixed-blessings model* (Haslam and Kvaale, 2015), which combines assumptions drawn from Weiner's (1986) attribution theory with Gelman's (2009) essentialism framework, biogenetic explanations, in comparison to psychosocial explanations, are believed to reduce social distance toward affected individuals, but are also assumed to increase pessimism concerning the treatability and changeability of their condition (Kvaale et al., 2013; Lebowitz et al., 2016). To date, the mixed-blessings model has been widely used to improve our

understanding of mental illness stigma in the case of adults (e.g., Dittrich et al., 2021). Although the model represents a promising framework for research on stigmatization mechanisms, the reduction of stigmatization and the promotion of positive attitudes in the context of inclusive education, it has been adopted to investigate the stigmatization of children with disabilities only rarely (e.g., Zensen and Röhm, 2021). However, the question of which disability-specific causal information should be emphasized in the context of teacher education and communication regarding children with disabilities to reduce the likelihood of accidental stigmatization remains largely unanswered.

The present study aims to (a) apply and conceptualize the mixed-blessings model to stigma-related reactions to children's disability labels in the context of inclusive education and (b) shed light on the causal attributions by teachers that underlie stigma-related attitudes toward children with various disabilities. The model's applicability is tested by an experiment that employs single case descriptions (case vignettes) of children with different types of disabilities in the context of inclusive education. More precisely, the study examines the ways in which disability-specific causes and symptoms, the type of disability in question, the children's sex, and efficacy cues regarding educational efforts affect student and trainee teachers' overall attitudes toward and expectations of inclusive education and social distance toward children with disabilities.

Originally, the mixed-blessings model developed by Haslam and Kvaale (2015) postulated that information concerning the biogenetic causes of an illness contributes to either (1) an attribution of uncontrollability (e.g., disability as a consequence of fate; cf. Weiner, 1986; Weiner et al., 1988) or (2) perceptions of psychological essentialism, which ascribe an illness or disability to a person's personality (cf. Gelman, 2009). While the former attribution is known to reduce blame and social distance toward an affected person (Weiner et al., 1988; Dijkster and Koomen, 2003), the latter supposedly increases social distance as well as prognostic pessimism regarding the changeability and treatability of the disability as well as its perceived dangerousness. However, these patterns have been confirmed only partially by various studies concerning the stigma associated with mental illness, thereby highlighting the linkage between attributed uncontrollability and decreased social distance as well as between essentialist beliefs and increased prognostic pessimism (Kaushik et al., 2016; Lebowitz et al., 2016; Lebowitz and Appelbaum, 2019). For instance, Dittrich et al. (2021) did not observe significant associations among biogenetic causes (vs. psychosocial causes), essentialist beliefs, and an increase in social distance toward persons with schizophrenia. In contrast, participants who were presented with a psychosocial causal explanation for schizophrenia indicated a relation between their biogenetic causal beliefs and increased social distance, which was mediated by their essentialist beliefs. According to those authors, their “differential findings can be accounted for by the subjects' different readiness to subscribe to biogenetic and psychosocial causal beliefs” (Dittrich et al., 2021,

p. 8), thus highlighting the importance of examining the model's implications for anti-stigma interventions in further detail.

To apply the mixed-blessings model to stigma in the context of inclusive education (Figure 1), the present study focuses on the empirically confirmed relations between biogenetic causes and, on the one hand, attributed uncontrollability (disability viewed as fate) and decreased stigmatization (i.e., social distance) as well as, on the other hand, psychological essentialism (disability viewed as an identity) and increased prognostic pessimism. *Social distance* is the most widely used operationalization of individual stigmatizing attitudes “to assess (expected) discriminatory behavior” (Baumann, 2007, p. 132). However, *prognostic pessimism* in the context of inclusive education can be understood in terms of respondents' attitudes toward and efficacy expectations of inclusive educational efforts and settings. This approach builds on Bandura's (1977) concept of *self-efficacy*, which is defined as a belief in one's competence to achieve goals in a certain situation. For instance, teachers' self-efficacy has been repeatedly linked to teaching outcomes (Klassen and Chiu, 2010) as well as their beliefs and attitudes (Dignath et al., 2022) in inclusive educational settings. Therefore, it can be presumed that descriptions of biogenetic causes can help reduce stigmatization, for example, because they lead to the belief that the affected person is not to blame for his or her condition. Simultaneously, such descriptions can also lead to the perception that the person's condition or situation cannot be altered, for instance, by educational efforts or interventions, thereby leading to pessimistic expectations (e.g., reduced efficacy expectations). One study conducted by Zensen and Röhm (2021) examined the ways in which depictions of biogenetic, psychosocial, or bio-psychosocial explanations of ADHD in case vignettes influence student teachers' social distance toward affected children as well as their attitudes toward inclusive education in a 3 × 2 online experiment. Their findings suggest that biogenetic causes (vs. psychosocial causes) decrease student teachers' social distance but not their positive attitudes toward inclusive education. However, a combination of biogenetic and psychosocial causes produced the most positive attitudes. Since the study by those authors did not operationalize student teachers'

efficacy expectations of inclusive education properly, the transferability of effects on prognostic pessimism is highly limited and, hence, deserves further examination.

In light of the present research and the proposed framework, we presume that the depiction of biogenetic causes leads to less social distance (as an indicator of stigmatization) toward children with disabilities but also decreases student and trainee teachers' positive attitudes toward and efficacy expectations (as indicators of prognostic pessimism) of inclusive education compared to cases featuring psychosocial causes:

Hypothesis 1a: Compared to a primary emphasis on psychosocial causes of the depicted disability, highlighting biogenetic causes reduces respondents' social distance toward children with disabilities.

Hypothesis 1b: Compared to a primary emphasis on psychosocial causes of the depicted disability, highlighting biogenetic causes reduces respondents' positive attitudes toward and efficacy expectations of inclusive educational settings.

In accordance with the assumptions of labeling theory (Link and Phelan, 2001) and *priming* (Molden, 2014), certain disability-related labels can unintentionally or unconsciously activate the stereotypical attitudes and intentions associated with that specific label. Due to the heterogeneity of pupils' types of disability in the context of inclusive education, the present study tests the applicability of the adapted mixed-blessings model to *behavior*-related, *communication*-related, and *cognition*-related disability labels. These various labels present different challenges that pertain to teachers' professional competencies, such as pedagogical, didactic, and educational interventions and classroom management (e.g., Blotnicky-Gallant et al., 2015). In the context of this study, *attention deficit hyperactivity disorder* (ADHD) is used as an example of a behavior-related disability label, while *developmental language disorder* (DLD) and *intellectual*

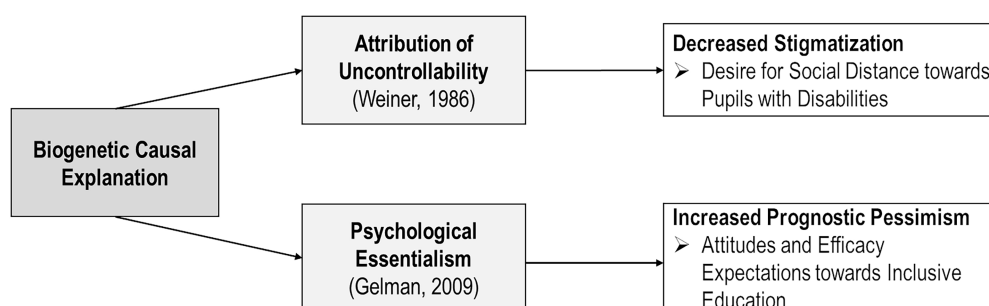


FIGURE 1

Adaptation of the mixed-blessings model to the effect of biogenetic causal explanations on stigmatization, attitudes, and expectations in the context of inclusive education.

disability (ID) represent communication- and cognition-related disability labels, respectively. While stigmatization of children with ID (Scior et al., 2012; e.g., Wilson and Scior, 2015; Scior and Furnham, 2016) and ADHD (e.g., Lebowitz et al., 2016; Zensen and Röhm, 2021) are well documented, little is known regarding stigma-related reactions to children with DLD (Macharey and Von Suchodoletz, 2008). Overall, intellectual disabilities are highly stigmatized due to their invisibility and the severity and controllability that are frequently attributed to them (Miller et al., 2009; Venville et al., 2016), whereas children with ADHD are perceived as noticeable and challenging but do not generally face high levels of stigmatization (Röhm et al., 2018). Regarding the stigma associated with DLD, Bishop (2017) notes that stigmatization can be associated with the specific label but also “that stigmatization is often a reaction to the child’s communication difficulties” (Bishop, 2017, p. 674).

Accordingly, it can be assumed that the depiction of a pupil with ID elicits the most stigmatization and the least efficacy expectations compared to the depictions of students with ADHD or DLD but also that a pupil with ADHD nevertheless faces more stigmatizing reactions than a pupil with DLD:

Hypothesis 2a: A case vignette depicting a pupil with ID is associated with greater social distance as well as fewer positive attitudes toward and efficacy expectations of inclusive education than a case vignette depicting a pupil with ADHD or DLD.

Hypothesis 2b: A case vignette depicting a pupil with ADHD evokes greater social distance as well as fewer positive attitudes toward and efficacy expectations of inclusive education than a case vignette depicting a pupil with DLD.

According to research findings in the context of health communication, the presentation of efficacy-related information (i.e., *efficacy cues*) influences recipients’ attitudes and behavioral intentions toward certain issues, such as vaccinations (Ort and Fahr, 2018). While teachers’ self-efficacy plays an important role in inclusive educational settings (e.g., Klassen and Chiu, 2010) and affects their general attitudes toward this topic (Savolainen et al., 2011), little is known regarding the effect of efficacy cues contained in single-case pupil descriptions on respondents’ attitudes and efficacy expectations in the context of inclusive education. In this regard, it can be assumed that the depiction of a pupil’s support needs that can be easily satisfied and offer a high chance of learning success (a *high-efficacy cue*) evoke more positive attitudes toward and higher efficacy expectations of inclusive education than the depiction of support needs that are more difficult and costly to satisfy and offer hardly any chance of learning success (a *low-efficacy cue*):

Hypothesis 3: Compared to a low-efficacy cue (difficult and costly support needs associated with low expectations of

learning success), a high-efficacy cue (simple support needs associated with high expectations of learning success) increases respondents’ positive attitudes toward and efficacy expectations of inclusive education.

In Germany, the training of teachers is divided into two phases and differentiated in accordance with the teachers’ subsequent type of teaching profession (e.g., elementary school, secondary school, or special education). While student teachers are learning theoretical and didactic basics regarding their teaching profession while studying in bachelor’s and master’s programs at university, trainee teachers are already working in the learning environment of schools. Although both special and general education teachers are expected to work in inclusive educational settings, it is likely that general education teachers have only limited contact with students with disabilities during their training. In both parts of their training, prospective special education teachers encounter content that is specifically adapted to the target group of students with different types of disabilities. In the practical sections of the training, such teachers are required to gain experience with this special target group. Therefore, the preconditions, previous experiences, and efficacy expectations associated with teaching pupils with disabilities are expected to differ between special education and general education teachers as well as between student teachers and trainee teachers. This assumption is supported by the findings of one recent literature review of 71 studies conducted by Wray et al. (2022), who identified teacher education and training as well as teachers’ experiences with people with disability as important factors with respect to teachers’ self-efficacy in the context of inclusive education.

In light of the goals of the present study, the connection between student and trainee teachers’ general confidence in their teaching skills (i.e., their teaching self-efficacy; Savolainen et al., 2011) and their attitudes toward and efficacy expectations of inclusive education in particular deserves further attention. In this regard, Röhm et al. (2018) observed that student teachers with high teaching self-efficacy report more favorable attitudes overall toward children with disabilities than do student teachers with low teaching self-efficacy. Furthermore, teaching self-efficacy served as a moderator in a complex pattern of interaction that involves pupils’ grades and respondents’ sex, whereas “especially bad-graded pupils seem to evoke more positive attitudes in male student teachers than female student teachers with high teaching self-efficacy” (Röhm et al., 2018, 52). Thus, it can be assumed that respondents’ teaching self-efficacy affects their attitudes toward and efficacy expectations of inclusive education as well as their social distance toward children with disabilities in general. In addition, student and trainee teachers with high teaching self-efficacy are presumably more positively disposed toward inclusive education and children with disabilities when they are presented with low-efficacy cues (difficult and costly support needs associated with low expectations of learning success) than are teachers

with low teaching self-efficacy. The following hypotheses are thus proposed:

Hypothesis 4: Student and trainee teachers with high teaching self-efficacy report more positive attitudes toward and higher efficacy expectations of inclusive education as well as less social distance than student and trainee teachers with low teaching self-efficacy.

Hypothesis 5: When low-efficacy cues become salient, student and trainee teachers with high teaching self-efficacy report more positive attitudes toward and higher efficacy expectations of inclusive education as well as less social distance than student and trainee teachers with low teaching self-efficacy.

Although some evidence suggests that female pupils are perceived and evaluated more positively than male pupils (e.g., Steinmayr and Spinath, 2008; Burusic et al., 2012), another study (Röhm et al., 2018) reports that student teachers indicate more stigmatization of and fewer positive attitudes toward female pupils than male pupils. Thus, the role of the sex of the depicted pupil in the process of stigmatization is addressed in the form of the following research question:

Research Question 1: How does the sex of the depicted pupil affect respondents' social distance as well as their attitudes toward and efficacy expectations of inclusive education?

Materials and methods

Design and procedure

In a $3 \times 2 \times 2 \times 2 \times 2$ online experiment, student and trainee teachers were recruited via social media and mailing lists. Each respondent was randomly assigned to one of 48 online survey questionnaires featuring a fictional case vignette depicting a pupil in an inclusive elementary school. The vignettes were experimentally manipulated regarding the pupil's *type of disability* (ADHD vs. DLD vs. ID), *sex* (male vs. female), the *attributed cause* of the disability (biogenetic vs. psychosocial), the pupil's *need for educational support* (low vs. high) and the pupil's *chance of learning success* (low vs. high). Before reading the case vignette, respondents' teaching self-efficacy in terms of a trait was assessed as a potential moderator. Subsequently, their attitudes toward and efficacy expectations of inclusive education as well as their social distance toward children with disabilities were measured as primary dependent variables. Finally, a brief manipulation check was conducted, sociodemographic data were collected, and respondents were thanked for their participation. Participants' consent for and agreement with data collection and processing was obtained at the beginning of the survey by active confirmation in accordance with the EU General Data Protection Regulation

(GDPR). All participants were informed of the context of the study prior to participating and were subsequently debriefed in detail regarding the purpose of the experimental stimulus and the details of the questionnaire. A required sample size of approximately $n = 632$ for the experiment was estimated via an a priori power analysis using G*Power (Faul et al., 2009) to conduct a multivariate analysis of variance (MANOVA) with $f^2 = 0.01$, $p < 0.05$, and a power of 0.80.

Sample

In total, 1,471 student and trainee teachers from various German universities were recruited for the study, of whom $N = 605$ ($M = 25.36$ years; $SD = 5.15$; 96.2% female) completed the full survey (59% dropout). Table 1 shows the distribution of the final sample by student or trainee teacher status and type of teaching profession.

Stimulus material

The stimulus material consisted of a case vignette depicting a pupil in an elementary school featuring an inclusive educational setting (Figure 2; see Table 2 for the English translation). All 48 case vignettes were nearly equal in length, with $M = 129.79$ words ($SD = 1.00$). Each case vignette was illustrated using a neutral image of a class room situation, which was kept constant throughout all experimental conditions.

Experimental manipulations

The pupil's type of disability is labeled ADHD to represent a behavioral disorder, DLD to indicate a communicative disorder, or ID to suggest a cognitive disorder. The pupil's sex is indicated by the name given (male: Jonas; female: Julia) and the use of corresponding pronouns. A genetic predisposition (biogenetic) or a conflict-ridden parental home (psychosocial) are mentioned as the attributed cause of the pupil's disability, which affects the pupil's socioemotional, language, or cognitive development. Symptoms in all conditions are described in terms of noticeable activity in specific brain areas that are responsible for directing and focusing attention (ADHD), language acquisition and

TABLE 1 Sample distribution by student or trainee teacher status and type of teaching profession.

	Student teachers	Trainee teachers	Total
Type of teaching profession			
Elementary school	206	97	303
Secondary school	105	74	179
Vocational school	17	8	25
Special education	68	30	98
Total	396	209	605

Jonas hat eine geistige Behinderung: Ein Fallbeispiel

Der zehnjährige Jonas besucht den inklusiven Unterricht an einer städtischen Grundschule. Er hat Schwierigkeiten, sich dort aufgrund seiner kognitiven Beeinträchtigung zurechtzufinden. Ergebnisse medizinischer und psychologischer Untersuchungen weisen auf biologische Ursachen seiner geistigen Behinderung hin. Demnach kann diese auf eine ungünstige genetische Veranlagung zurückgeführt werden, wodurch Jonas' geistige Entwicklung beeinträchtigt wurde. Außerdem wurden bei Jonas auffällige Aktivitäten in Gehirnregionen festgestellt, die mit der allgemeinen Intelligenz und der Verarbeitung von Informationen in Verbindung stehen.



Im Unterricht ist Jonas auf viel Unterstützung angewiesen. Beim Rechnen benötigt er häufig zusätzliche Instruktionen und Materialien, die die Schwierigkeit der Aufgaben reduzieren sollen und von den Lehrkräften aufwendig vorbereitet werden müssen. Dennoch bestehen gute Chancen, dass Jonas entsprechend seiner individuellen Voraussetzungen gefördert werden kann und beim Lernen viele Erfolgsergebnisse haben wird.

FIGURE 2

Example of stimulus material (see Table 2 for the English translation; manipulations: type of disability: intellectual disability; sex: male; attributed cause: biological; need for support: high; chance of learning success: high; Image Source: [dolgachov/123rf.com](https://doi.org/10.3389/fpsyg.2022.910702)).

processing (DLD), or general intelligence and information processing (ID). The pupil's need for support is illustrated by way of example in the context of arithmetic: he or she requires either a great deal of or little support and either frequent or infrequent additional instructions and materials with respect to structuring (ADHD), visualizing (DLD), or reducing the difficulty of (ID) school tasks. These instructions and materials can be prepared through the expenditure of either low ("quickly") or high ("time-consuming") effort by the teacher. The vignette concludes by highlighting the pupil's prospects of learning success, indicating either a chance for a great deal of success or the risk of hardly any success.

Manipulation check

To estimate the correct recognition of all experimental stimulus manipulations, the respondents were asked whether they remembered (1) which type of disability ("ADHD," "DLD," or "ID"), (2) which sex ("male" or "female"), (3) which cause of disability ("biogenetic" or "psychosocial"), (4) which level of need for educational support ("high and costly" or "low and uncomplicated"), and (5) which chance of learning success ("hardly any chance of success" or "large chance of success") were depicted in the case vignette. Chi-squared tests for each pair of respondents' categorical answers and the corresponding stimulus manipulation showed that all experimental conditions were successfully and unanimously recognized (type of disability: $X^2(4, N = 599) = 1052.01, p < 0.001$; sex: $X^2(1, N = 602) = 594.04, p < 0.001$; attributed cause: $X^2(1, N = 602) = 454.54, p < 0.001$; need for educational support: $X^2(1, N = 599) = 375.76, p < 0.001$; chance of learning success: $X^2(1, N = 598) = 237.21, p < 0.001$).

TABLE 2 Original German stimulus text from Figure 2 and its English translation (manipulations: type of disability: intellectual disability; sex: male; attributed cause: biological; need for support: high; chance of learning success: high).

Original German stimulus text	English translation
Jonas hat eine geistige Behinderung: Ein Fallbeispiel	Jonas has an Intellectual Disability: A Case Study
Der zehnjährige Jonas besucht den inklusive Unterricht an einer städtischen Grundschule. Er hat Schwierigkeiten, sich dort aufgrund seiner kognitiven Beeinträchtigung zurechtzufinden. Ergebnisse medizinischer und psychologischer Untersuchungen weisen auf biologische Ursachen seiner geistigen Behinderung hin. Demnach kann diese auf eine ungünstige genetische Veranlagung zurückgeführt werden, wodurch Jonas' geistige Entwicklung beeinträchtigt wurde. Außerdem wurden bei Jonas auffällige Aktivitäten in Gehirnregionen festgestellt, die mit der allgemeinen Intelligenz und der Verarbeitung von Informationen in Verbindung stehen.	Ten-year-old Jonas attends inclusive classes at a city elementary school. He has difficulties finding his way around there due to his cognitive impairment. The results of medical and psychological examinations point to biological causes of his intellectual disability. These causes can be attributed to an unfavorable genetic predisposition, which has impaired Jonas' mental development. In addition, Jonas has been found to have abnormal activity in brain regions associated with general intelligence and information processing.
Im Unterricht ist Jonas auf viel Unterstützung angewiesen. Beim Rechnen benötigt er häufig zusätzliche Instruktionen und Materialien, die die Schwierigkeit der Aufgaben reduzieren sollen und von den Lehrkräften aufwendig vorbereitet werden müssen. Dennoch bestehen gute Chancen, dass Jonas entsprechend seiner individuellen Voraussetzungen gefördert werden kann und beim Lernen viele Erfolgsergebnisse haben wird.	In class, Jonas relies on a lot of support. When doing arithmetic, he often needs additional instructions and materials to reduce the difficulty of the tasks, which require extensive preparation by the teachers. Nevertheless, there is a good chance that Jonas can be supported according to his individual needs and experience a great deal of success in learning.

Instruments

The following section describes the instruments used in this study. The reliability of each instrument is indicated by both Cronbach's alpha and McDonald's omega, including a standard error and 95% confidence interval based on 1,000 bootstrap samples using Hayes and Coutts (2020) OMEGA macro for SPSS.

Moderator: Teaching self-efficacy as a trait

Teaching self-efficacy as a trait was assessed using the *teacher self-efficacy* scale developed by Schwarzer and Schmitz (1999). This scale measures teachers' expectations of their ability to cope with specific teaching-related situations against the backdrop of their perceived competencies and personality as a teacher based on Bandura's (1977) social cognitive theory. Respondents were asked to indicate their levels of agreement with ten statements pertaining to

an example situation (e.g., “I know that I am able to teach test-relevant content even to problematic pupils”) on a four-point Likert-type scale (0 = “not true”; 3 = “precisely true”). High scores indicate high levels of teaching self-efficacy. The scale’s reliability reached a sufficient Cronbach’s alpha of 0.744 ($SE=0.017$; 95% BaCI[0.707; 0.775]) and McDonald’s omega of 0.742 ($SE=0.018$; 95% BaCI[0.701; 0.775]). To facilitate the inclusion of these scores in the main analysis in the form of a categorical group variable, mean scores were calculated and dummy-coded as either 1 = ‘low teaching self-efficacy’ ($n=314$) or 2 = ‘high teaching self-efficacy’ ($n=291$) as determined by a median split at $Md=2.0$.

Dependent variables

Attitudes toward and efficacy expectations of inclusive education

To operationalize prognostic pessimism in the context of inclusive education, respondents’ attitudes toward and efficacy expectations regarding that topic were assessed using the *perceived self-efficacy in inclusive education* (four items; Cronbach’s alpha = 0.811, $SE=0.013$, 95% BaCI[0.783; 0.834]; McDonald’s omega = 0.816, $SE=0.013$, 95% BaCI[0.789; 0.840]) and *arrangement of inclusive education* (four items; Cronbach’s alpha = 0.743, $SE=0.019$, 95% BaCI[0.700; 0.776]; McDonald’s omega = 0.728, $SE=0.023$, 95% BaCI[0.678; 0.769]) subscales of Bosse and Spörer’s (2014) short scales for inclusive attitudes and self-efficacy of teachers. Participants reported the extent to which they agreed or disagreed with assertions such as “I have the confidence to organize lessons in such a way that children like [Jonas/Julia] can achieve their goals at their own learning pace” (perceived self-efficacy) and “Joint teaching of children with and without disabilities can meet the needs of all children through appropriate methods” (arrangement of inclusive education) on a four-point Likert-type scale (1 = “fully reject”; 4 = “fully agree”). High scores indicate a high perception of self-efficacy and confidence regarding the arrangement of inclusive education, thus representing low levels of prognostic pessimism regarding the assumptions drawn from the mixed-blessings model.

Social distance toward children with a disability

Respondents’ tendency to distance themselves from children with a disability was assessed using eight items taken from the *social distance* subscale of the German adaptation of the *Mental Retardation Attitude Inventory* (MRAI-d; Schabmann and Kreuz, 1999). Participants indicated their levels of agreement or disagreement with statements such as “I would rather not invite a child with a disability to play with the friends of my child who do not have a disability” on a four-point Likert-type scale (1 = “do not agree at all”; 4 = “strongly agree”; Cronbach’s alpha = 0.626, $SE=0.072$, 95% BaCI[0.444; 0.734]; McDonald’s omega = 0.615, $SE=0.070$, 95% BaCI[0.459; 0.729]). High scores indicate a high tendency to engage in social distancing behavior.

Instrument descriptive statistics and missing value analysis

Table 3 shows the means and standard deviations of as well as the intercorrelations among the moderator and the three dependent variables included in this study. A missing value analysis (MVA) indicated that missing values did not occur at random using Little’s MCAR (missing completely at random) test: $X^2=1466.22$, $df=692$, $p<0.001$. Of the total dropout of $n=866$ participants 45% ($n=395$) aborted the survey before or during answering the moderator scale (i.e., teaching self-efficacy) and another 30% ($n=260$) before the stimulus presentation. Further 22% ($n=185$) of the participants dropped out before or during answering the dependent measures, whereas only 3% ($n=26$) of the dropped-out participants incompletely answered one or two dependent measures and did not give any demographic information. All data analyses were conducted using IBM SPSS Version 26.

Results

All hypotheses and the research question were tested by conducting a MANOVA including the five experimental manipulations (1. type of disability; 2. sex; 3. attributed cause of disability; 4. need for educational support; 5. chance of learning success) as well as respondents’ teaching self-efficacy (median split) and student/trainee status as factors for all three dependent variables. Respondents’ sex was not included as a factor to prevent unequal distributions among the experimental and quasi-experimental factors. To ensure sufficient cell sizes with $n>30$ respondents per cell, the MANOVA model was limited to main effects, two-way interactions, and three-way interactions. Concerning potential statistical outliers, the calculation of Mahalanobis distances yielded $n=11$ cases (1.8%) above the chi-squared distribution cutoff value of 16.266 ($df=3$, $p<0.001$). Due to the comparably low number of outliers and the general robustness of MANOVAs against extreme values (cf. Field, 2018), these cases were included in the analyses. To protect subsequent univariate analyses of variance (ANOVAs) against type I error, only effects of the MANOVA with $p\leq 0.01$

TABLE 3 Means, standard deviations, and intercorrelations of the moderator and the three dependent variables.

		<i>M</i>	<i>SD</i>	(2)	(3)	(4)
<i>Moderator</i>						
(1)	Teaching self-efficacy	2.04	0.35	0.52**	0.23**	−0.14**
<i>Dependent variables</i>						
(2)	Perceived self-efficacy	2.80	0.63		0.39**	−0.08*
(3)	Arrangement of inclusive education	3.19	0.62			−0.15**
(4)	Social distance	1.07	0.19			

** $p<0.01$; * $p<0.05$.

TABLE 4 Significant main effects and higher order interactions shown by the MANOVA regarding all five experimental manipulations alongside respondents' teaching self-efficacy and student or trainee status as factors using Pillai's trace.

	<i>V</i>	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>
<i>Main effects</i>					
Attributed cause of disability	0.028	4.986	3	517	=0.002
Respondents' teaching self-efficacy	0.188	39.816	3	517	<0.001
<i>Higher-order interaction of stimulus manipulations</i>					
Type of disability × Chance of learning success	0.032	2.852	6	1,036	=0.009
<i>Higher-order interaction of stimulus manipulations and respondents' characteristics</i>					
Type of disability × Respondents' teaching self-efficacy × Respondents' student/trainee status	0.032	2.803	6	1,036	=0.010

Perceived self-efficacy in inclusive education, arrangement of inclusive education, and social distance were included as dependent variables.

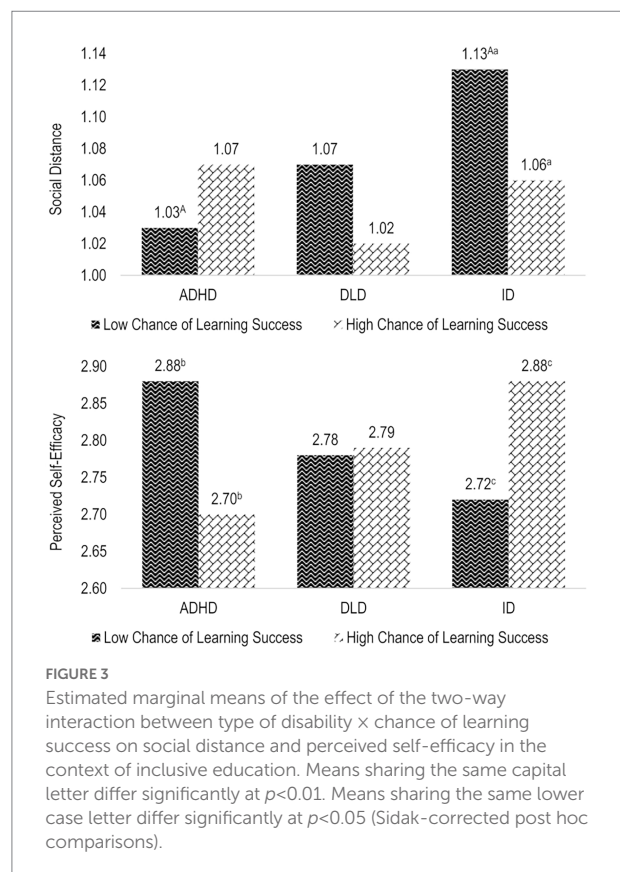
are reported (Field, 2018). Table 4 displays all significant main effects and higher-order interactions uncovered by the MANOVA using Pillai's trace. The significant main effects and higher-order interactions that emerged from subsequent ANOVAs are reported below. Effect sizes are indicated by partial eta-squared (η_p^2) for subsequent ANOVAs based on the output of SPSS (cf. Ellis, 2010; Lakens, 2013; Field, 2018). The significance of differences between the estimated marginal means was determined using Sidak-corrected simple effect post hoc tests.

Main effect of the cause attributed to the disability

A main effect of the cause attributed to the disability became significant with respect to arrangement of inclusive education, $F(1,519) = 12.547$, $p < 0.001$, $\eta_p^2 = 0.024$. Compared to the depiction of a psychosocial cause for the disability ($M = 3.27$; $SE = 0.04$), the depiction of a biogenetic cause led to significantly fewer positive attitudes regarding the arrangement of inclusive education ($M = 3.07$; $SE = 0.04$; $p < 0.001$), a result which indicates higher prognostic pessimism, thus supporting Hypothesis 1b.

Main effects of respondents' teaching self-efficacy

Significant main effects of respondents' teaching self-efficacy emerged with respect to perceived self-efficacy in inclusive education, $F(1,519) = 117.986$, $p < 0.001$, $\eta_p^2 = 0.185$, and arrangement of inclusive education, $F(1,519) = 20.079$, $p < 0.001$, $\eta_p^2 = 0.037$. Respondents in the group with high teaching self-efficacy indicated significantly more perceived self-efficacy in inclusive education ($M = 3.08$; $SE = 0.04$) and



confidence in the arrangement of inclusive education ($M = 3.30$; $SE = 0.04$) after reading the case vignette than respondents with low teaching self-efficacy (perceived self-efficacy: $M = 2.51$, $SE = 0.04$, $p < 0.001$; arrangement of inclusive education: $M = 3.05$; $SE = 0.04$; $p < 0.001$). These findings support Hypothesis 4.

Higher-order interaction of stimulus manipulations

A significant two way interaction between type of disability × chance of learning success appeared with respect to social distance, $F(2,519) = 3.922$, $p = 0.020$, $\eta_p^2 = 0.015$, and perceived self-efficacy in inclusive education, $F(2,519) = 3.809$, $p = 0.023$, $\eta_p^2 = 0.014$. Case vignettes depicting a pupil with ID evoked significantly greater social distance than case vignettes depicting a pupil with ADHD when the ascribed chance of learning success was low (Figure 3). In addition, the depiction of a pupil with ID and a low chance of learning success led to greater social distance and lower perceived self-efficacy than the depiction of a pupil with the same disability and a high chance of learning success. In contrast, depicting a pupil with ADHD and a low chance of learning success led to higher perceived self-efficacy than depicting a pupil with the same disability and a high chance of learning success.

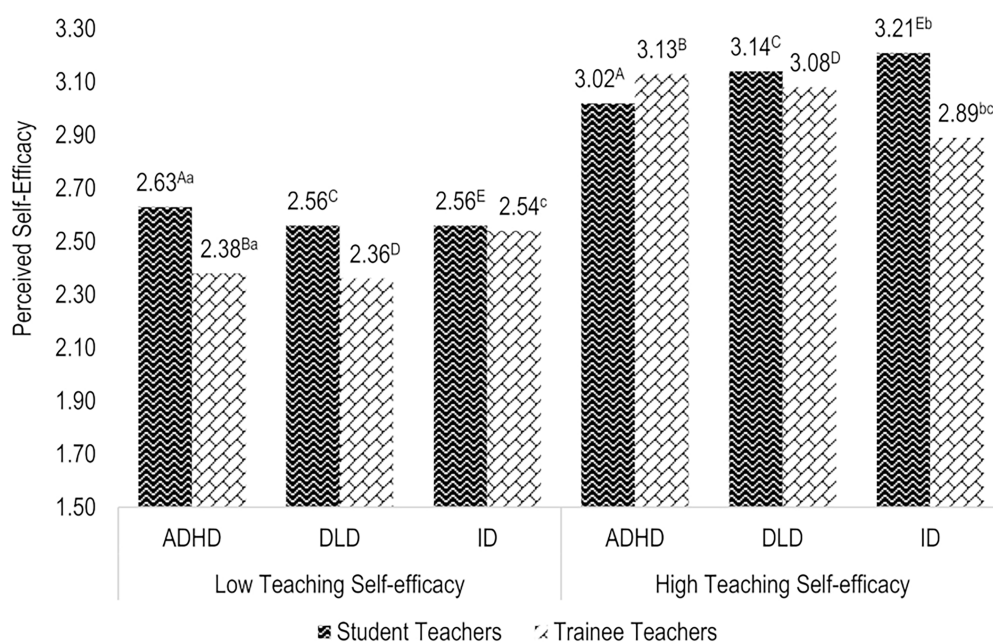


FIGURE 4

Estimated marginal means of the effect of the three-way interaction between type of disability \times respondents' teaching efficacy \times respondents' student/trainee teacher status on perceived self-efficacy in the context of inclusive education. Means sharing the same capital letter differ significantly at $p < 0.01$. Means sharing the same lower case letter differ significantly at $p < 0.05$ (Sidak-corrected post hoc comparisons).

Higher-order interaction of stimulus manipulation and respondents' characteristics

The three way interaction among type of disability \times respondents' teaching self-efficacy \times respondents' student/trainee status became significant with respect to perceived self-efficacy, $F(2,519) = 3.600$, $p = 0.028$, $\eta_p^2 = 0.014$. Figure 4 indicates that trainee teachers with low teaching self-efficacy reported significantly less perceived self-efficacy after reading a case vignette featuring a pupil with ADHD than did student teachers with low teaching self-efficacy. Likewise, trainee teachers with high teaching self-efficacy reported significantly less perceived self-efficacy after reading a case vignette featuring a pupil with ID than did student teachers with high teaching self-efficacy. Moreover, both student and trainee teachers with low teaching self-efficacy indicated significantly less perceived self-efficacy in the context of inclusive education in all disability conditions than did student and trainee teachers with high teaching self-efficacy.

Discussion

The aim of the present study was to shed light on student and trainee teachers' perceptions of pupils with disabilities in the context of inclusive education. In accordance with assumptions drawn from exemplification theory (Zillmann and Brosius, 2000),

respondents' attitudes and expectations regarding this topic were generally influenced by the single-case depictions (i.e., exemplars) that were presented. In particular, the cause attributed to the disability, the depicted type of disability and the chance of learning success led to attitude changes. Respondents' teaching self-efficacy and status as students or trainees emerged as moderators in this regard.

Specifically, the observed effect of an attributed biogenetic cause of the pupil's disability (vs. a psychosocial cause) on decreased positive attitudes regarding the arrangement of inclusive education supports Hypothesis 1b (i.e., that highlighting biogenetic causes reduces respondents' positive attitudes and efficacy expectations compared to cases highlighting psychosocial causes). In line with the assumptions drawn from the mixed-blessings model (Haslam and Kvaale, 2015), emphasizing a biogenetic cause decreased respondents' beliefs regarding the successful design of inclusive education settings for children with the depicted disability, thus indicating increased prognostic pessimism regarding this topic. With respect to the model's underlying assumptions in terms of psychological essentialism (Gelman, 2009), this finding could indicate that a case vignette featuring a biogenetic cause would lead to a perception of the depicted disability as part of the pupil's personality, which would make it unresponsive to educational efforts. However, since this manipulation had no significant effects on either perceived self-efficacy in inclusive education or social distance, no further support was provided for Hypothesis 1b, and no support at all was

found for Hypothesis 1a (i.e., that highlighting biogenetic causes reduces respondents' social distance toward children with disabilities compared to cases highlighting psychosocial causes). Thus, the adoption of the mixed-blessings model to stigma in the context of inclusive education is only partially supported by the present results. Moreover, no interactions between the depicted cause and the pupil's type of disability were found, indicating a lack of disability-specific effects. In contrast to previous findings (Lebowitz et al., 2016; Zensen and Röhm, 2021), it was not the case for any of the three disability labels (ADHD, DLD, or ID) that a suggested biogenetic cause evoked an attribution of uncontrollability resulting in reduced social distance toward affected children. Since intellectual disabilities are generally highly stigmatized due to their alleged severity and controllability (e.g., Venville et al., 2016), it is possible that teachers' perceptions of the controllability or uncontrollability of the pupils' ID became equalized between the biogenetic and the psychosocial condition. Dittrich et al. (2021) describe a similar effect found by their test of the mixed-blessings model regarding persons with schizophrenia. Accordingly, respondents may have been unready to change their original attributions on the basis of the mere presentation of the challenging biogenetic causal explanation. Concerning depictions of pupils with DLD, little is known regarding attributions of controllability. Thus, emphasizing neither a biogenetic nor a psychosocial cause made a difference in terms of the stigmatization of pupils with DLD in our study. It remains unclear, however, why an attributed biogenetic cause did not lead to any differences in respondents' social distance toward affected children in the case of the vignettes featuring a pupil with ADHD.

Although no sole main effects of type of disability emerged, the observed interaction effects with the depicted chance of learning success provide some insights into the ways in which stigmatizing attitudes regarding specific disability labels can be shaped by the prospect that educational efforts will have a positive outcome. In this regard, pupils with ID were most stigmatized and associated with the lowest self-efficacy expectations when their chance of learning success was described as low, a finding which is to some extent in line with Hypothesis 2a (i.e., that a pupil with ID is associated with greater social distance as well as fewer positive attitudes and efficacy expectations than a pupil with ADHD or DLD). However, when the pupil's chances of learning success were high, the ID label did not elicit more stigmatization than ADHD or DLD. Moreover, respondents' perceived self-efficacy regarding a pupil with ID and a high chance of learning success was as high as that concerning a pupil with ADHD and a low chance of learning success. While there were no significant differences whatsoever regarding pupils with DLD and, therefore, no support for Hypothesis 2b (i.e., a pupil with ADHD evokes greater social distance as well as fewer positive attitudes and efficacy expectations than a pupil with DLD), the prospect of a positive or negative learning outcome for pupils with an ADHD label seemed to alter respondents' reactions in the opposite direction compared to cases of an ID label. Case vignettes featuring a pupil with ADHD and a low chance of learning success

evoked the least stigmatization and the most perceived self-efficacy in the context of inclusive education, whereas the combination of an ADHD label and a high chance of learning success resulted in the lowest efficacy expectations. A possible explanation for such a counterintuitive effect pattern could be that, on the one hand, the high degree of stigmatization associated with a low-achieving pupil with ID highlights respondents' general expectations regarding children with ID and reflects the public stigma of individuals with ID (Venville et al., 2016). The prospect of a positive outcome, in contrast, provides an effective efficacy cue that children with ID can benefit from inclusive educational efforts. Accordingly, the reduction in stigma and increase in efficacy expectations partly support the assumptions of Hypothesis 3 (i.e., that a high-efficacy cue increases respondents' positive attitudes and efficacy expectations compared to a low-efficacy cue). Low-achieving children with ADHD, on the other hand, may have evoked respondents' aspirations to support and engage in inclusive education. Depictions of a high chance of successful learning for children with ADHD may have contributed to a lower degree of perceived necessity and ultimately the lower efficacy of an inclusive educational setting. Such pupils may have been perceived as already being successful learners despite their ADHD, leading to the assumption that participants felt less needed or less able to promote the pupils' learning success. However, this situation seems to be the case only for pupils with ADHD, since pupils with ID who were presented as successful learners were associated with high efficacy expectations. This finding may have been due to the cognitive nature of the latter pupils' disability, indicating that their learning success was not perceived as the same as that of other pupils and that it continued to indicate a need for further educational intervention. In summary, the prospect of learning success described moderated the effect of the pupils' type of disability, but not as consistently as presumed by Hypothesis 3. Since the depicted efficacy cues exhibited no other effects, support for this hypothesis is limited to the context of the ID label. Further research should examine how and which efficacy cues are most useful for reducing stigmatization and promoting the efficacy expectations of future teachers.

Student and trainee teachers' general teaching self-efficacy emerged as an important factor with regard to shaping attitudes toward and expectations of inclusive education. As predicted by Hypothesis 4, respondents with high teaching self-efficacy were more confident and positive with respect to this topic than respondents with low teaching self-efficacy. This finding is in line with the conclusions of previous studies (Savolainen et al., 2011; Röhm et al., 2018) investigating the influence of teaching self-efficacy on attitudes toward inclusive education. However, respondents' teaching self-efficacy did not interact with the efficacy cues depicted. Hence, Hypothesis 5, which presumed that if low-efficacy cues are salient, student and trainee teachers with high teaching self-efficacy report more positive attitudes toward and higher efficacy expectations of inclusive education as well as less social distance than student and trainee teachers with low

teaching self-efficacy, must be rejected. Moreover, the observed main effect of teaching self-efficacy was also reflected in a higher-order interaction involving the pupils' type of disability as well as respondents' status as either students or trainees. These effect patterns additionally indicate a difference between student teachers' and trainee teachers' self-efficacy in an inclusive educational setting with respect to pupils with ADHD and ID, respectively, depending on their general teaching self-efficacy as a trait. In the low teaching self-efficacy condition, student teachers felt more equipped to educate pupils with ADHD than trainee teachers. When teaching self-efficacy was high, however, student teachers were more confident of their ability to educate pupils with ID than trainee teachers. This finding could suggest that trainee teachers, who have actual experience with teaching, exhibit reduced self-efficacy expectations with respect to certain teaching constellations and specific pupils. This finding is in line with the results of a recent meta-analysis (Dignath et al., 2022), which reports that "self-efficacy beliefs were found to be higher for preservice than for in-service teachers" (p. 24). Other scholars likewise emphasize the effect of teaching experience on attitudes toward inclusive education (e.g., Avramidis and Kalyva, 2007). Otherwise, student teachers may have overestimated their actual competencies. These findings deserve further investigation in future studies.

Finally, the depiction of the pupil's sex did not make any difference on respondents' reactions, which is in line with the findings of some researchers (e.g., Zensen and Röhm, 2021) but conflicts with those of other scholars (Burusic et al., 2012; Röhm et al., 2018). Nevertheless, there is no consistent answer to the question of whether female or male pupils are more stigmatized or rated more favorably, especially in inclusive educational settings.

Overall, the present study provides valuable insights into the influences and limitations of different causal explanations, various types of disability, and efficacy cues as well as into teacher characteristics with respect to stigma, attitudes, and expectations of inclusive education. With regard to the applicability of the mixed-blessings model (Haslam and Kvaale, 2015) to this context, the findings support only a few of the model's original assumptions and do not indicate any disability-specific effects. Thus, several questions remain unanswered: How and which aspects, symptoms, and causes should be emphasized regarding the three exemplary types of disability referenced by this study to reduce stigmatization and promote educational efforts? Which information, in turn, should not be emphasized or should even be withheld when discussing specific pupils to prevent accidental stigmatization? To what extent is the observed positive effect of a psychosocial causal explanation transferable to other types of disability?

In addition to the effects of the attributed causes, the disability labels used in this study evoked complex patterns of reaction from respondents depending on other influencing factors. With regard to the assumptions drawn from labeling theory (Link and Phelan, 2001), no uniform stereotypical reactions emerged with respect to any of the disability labels. Respondents' reactions were instead influenced by additional contextual information that either

affirmed (e.g., pupils with ID and low chance of learning success) or contradicted (e.g., pupils with ID and high chance of learning success) stereotypical expectations. In addition, teaching self-efficacy played an important role with regard to the effects of the case vignette and produced the largest effect sizes. An integrative model of stigmatization and attitudes in the context of inclusive education, such as the adapted mixed-blessings model, should integrate the impact of student and trainee teachers' individual dispositions and experiences to account for the characteristics of both stigmatizers and the stigmatized (cf. Möhring et al., 2021).

Limitations

The current study examined the factors that shape teachers' attitudes toward inclusive education and social distance toward children with a disability in depth. The interpretation and generalizability of the results is to some extent limited to the context of the German education system and its implementation of inclusive education. Since we collected only explicit self-reported data, a certain degree of social desirability bias must be expected. Due to our use of an online experiment, issues such as a high dropout rate and self-selection bias may further limit the generalizability of our results (Reips, 2002), although the required sample size of approximately $n = 632$ that was estimated a priori was almost met. Although the high dropout rate is not uncommon for online studies (cf. Daikeler et al., 2020) the results from the MCAR test showed that participants did not abort the questionnaire at random (especially at the beginning or directly after the stimulus). This could indicate that participants that are generally less motivated, interested, or even opposed to the topic of inclusive education were to a lesser extent included into our final analyses. Future studies should aim to also include those views by, for example, adopting more research-economically scales, being sensitive to reactance, and providing incentives to guide participants attention und motivation throughout the questionnaire. Since we targeted a rather specific sample, the self-selection bias of student and trainee teachers was limited to their sex and age, characteristics which exhibited similar distributions to those found in previous studies (Röhm et al., 2018; Zensen and Röhm, 2021). Regarding respondents' sex, the present sample reflects the general demographics of the population, particularly that of primary and special education teachers (UNESCO, 2019). However, respondents' sex was not included as a factor in the analysis to prevent unequal distributions among the experimental and quasi-experimental factors. Thus, future studies should focus in more detail on the attitudes of male students and trainee teachers, who are known to be more stigmatizing than females (Kaushik et al., 2016). In addition, researchers should also aim to examine the possible effects of student and trainee teachers' previous experiences with pupils with disability on stigma-related attitudes in more detail. This factor could not be controlled for in the present study due to the relatively low number of student and trainee teachers with special educational teaching backgrounds.

The present study further aimed to limit (and control for) the role of contextual factors such as teaching conditions (e.g., overcrowding) and classroom management (cf. Blotnicky-Gallant et al., 2015) by refraining from describing them as much as possible and therefore ensuring that their salience to the respondents remained low. However, the experimental manipulation of the pupil's depicted need for support and prospects of learning success addressed some of these issues, which, in addition to the type of disability in question, may play an important role regarding teachers' perceptions of and interactions with the pupil.

Since the analyses conducted for this study were strictly controlled with regard to type 1 errors and only MANOVA effects at the 1% level of significance were included into further analyses, effects with $0.01 < p < 0.05$ were excluded, which may also have provided additional insights. Further investigations may employ a more powerful approach to examine the impact of each individual manipulation as well as the possible associated long-term effects. Regarding the operationalization of the constructs related to the mixed-blessings model, the measures applied with regard to prognostic pessimism yielded the hypothesized effects. Future studies should also employ scales that measure factors other than positive attitude dimensions. In addition, there was no direct operationalization of attributions of uncontrollability or essentialist beliefs as mediators as performed in, for instance, the recent study by Dittrich et al. (2021). Thus, interpretations of attributed uncontrollability or, in particular, essentialism relied on the link between the depicted cause and its effects on stigmatization and prognostic pessimism. Furthermore, the wording of the items employed to measure teaching self-efficacy may have biased respondents' responses regarding stigmatization due to statements such as "I know that I am able to teach test-relevant content even to problematic pupils." Additionally, in contrast to previous studies (Röhm et al., 2018; Zensen and Röhm, 2021), the social distance scale used in this study yielded only a comparably low reliability. This scale was also related only to personal social distance toward children with a disability in general instead of to teachers' social distance toward students with a disability. Hence, future studies should aim to operationalize all aspects of the mixed-blessings model in a more detailed, more reliable, and less biased way.

Conclusion

To prevent professional or institutional stigmatization, teacher education and training as well as communication regarding pupils with disabilities require a high degree of sensitivity to disability-specific and efficacy-related cues that can, in one arrangement, promote the inclusion of a certain group but may, in other contexts, lead to unintended reactions or expectations. Supporting future teachers' self-efficacy deserves greater attention with regard to the task of establishing

a more inclusive education system, as do appropriate forms of teacher training that include essentials from anti-stigma communication.

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 for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

AR conceptualized the study, conducted the data analysis, and wrote the first draft. MG, JZ-M, and AR organized and conducted the sample recruitment with contributions. MG, MM, JZ-M, CN, and MH contributed to the design, stimulus material, and operationalization of constructs. All authors contributed to the article and approved the submitted version.

Acknowledgments

We acknowledge the financial support by Deutsche Forschungsgemeinschaft (DFG) and Technische Universität Dortmund/TU Dortmund University as part of the funding program Open Access Costs.

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

- Avramidis, E., and Kalyva, E. (2007). The influence of teaching experience and professional development on Greek teachers' attitudes towards inclusion. *Eur. J. Spec. Needs Educ.* 22, 367–389. doi: 10.1080/08856250701649989
- Avramidis, E., and Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: a review of the literature. *Eur. J. Spec. Needs Educ.* 17, 129–147. doi: 10.1080/08856250210129056
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191
- Baumann, A. E. (2007). Stigmatization, social distance and exclusion because of mental illness: the individual with mental illness as a "stranger". *Int. Rev. Psychiatry* 19, 131–135. doi: 10.1080/09540260701278739
- Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *Int. J. Lang. Commun. Disord.* 52, 671–680. doi: 10.1111/1460-6984.12335
- Blotnick-Gallant, P., Martin, C., McGonnell, M., and Corkum, P. (2015). Nova Scotia teachers' ADHD knowledge, beliefs, and classroom management practices. *Can. J. Sch. Psychol.* 30, 3–21. doi: 10.1177/0829573514542225
- Bosse, S., and Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht [assessment of attitudes and self-efficacy of pre-service teachers towards inclusive education]. *Empirische Sonderpädagogik* 6, 279–299.
- Burusic, J., Babarovic, T., and Seric, M. (2012). Differences in elementary school achievement between girls and boys: does the teacher gender play a role? *Eur. J. Psychol. Educ.* 27, 523–538. doi: 10.1007/s10212-011-0093-2
- Caslin, M. (2021). "They have just given up on me" how pupils labelled with social, emotional and behavioural difficulties (SEBD) experience the process of exclusion from school. *Support Learn.* 36, 116–132. doi: 10.1111/1467-9604.12341
- Cooney, G., Jahoda, A., Gumley, A., and Knott, F. J. (2006). Young people with intellectual disabilities attending mainstream and segregated schooling: perceived stigma, social comparison and future aspirations. *J. Intellect. Disabil. Res.* 50, 432–444. doi: 10.1111/j.1365-2788.2006.00789.x
- Daikeler, J., Bošnjak, M., and Lozar Manfreda, K. (2020). Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *J. Surv. Stat. Methodol.* 8, 513–539. doi: 10.1093/jssam/smz008
- de Boer, A. A., Pijl, S. J., and Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: a review of the literature. *Int. J. Incl. Educ.* 15, 331–353. doi: 10.1080/13603110903030089
- Dignath, C., Rimm-Kaufman, S., van Ewijk, R., and Kunter, M. (2022). Teachers' beliefs about inclusive education and insights on what contributes to those beliefs: a meta-analytical study. *Educ. Psychol. Rev.* doi: 10.1007/s10648-022-09695-0
- Dijkster, A. J. M., and Koomen, W. (2003). Extending Weiner's attribution-emotion model of stigmatization of ill persons. *Basic Appl. Soc. Psychol.* 25, 51–68. doi: 10.1207/S15324834BASP2501_4
- Dittrich, D., Dernbach, K., Speerforck, S., Schindler, S., Häusser, J. A., and Schomerus, G. (2021). Testing the mixed-blessings model: what is the role of essentialism for stigmatizing attitudes towards schizophrenia? *Curr. Psychol.* doi: 10.1007/s12144-021-01653-3
- Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge: Cambridge University Press.
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149
- Field, A. P. (2018). *Discovering Statistics Using IBM SPSS Statistics*. London: Sage.
- Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annu. Rev. Psychol.* 60, 115–140. doi: 10.1146/annurev.psych.59.103006.093659
- Haslam, N., and Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: the mixed-blessings model. *Curr. Dir. Psychol. Sci.* 24, 399–404. doi: 10.1177/0963721415588082
- Hayes, A. F., and Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. *Commun. Methods Measure.* 14, 1–24. doi: 10.1080/19312458.2020.1718629
- Kaushik, A., Kostaki, E., and Kyriakopoulos, M. (2016). The stigma of mental illness in children and adolescents: a systematic review. *Psychiatry Res.* 243, 469–494. doi: 10.1016/j.psychres.2016.04.042
- Klassen, R. M., and Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102, 741–756. doi: 10.1037/a0019237
- KMK (2022). *Sonderpädagogische Förderung in Schulen 2011 bis 2020 [special education in schools from 2011 to 2020]*. Berlin: Kultusministerkonferenz. Available at: https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok231_SoPaeFoe_2020.pdf (Accessed September 01, 2022).
- Kvaale, E. P., Haslam, N., and Gottdiener, W. H. (2013). The "side effects" of medicalization: a meta-analytic review of how biogenetic explanations affect stigma. *Clin. Psychol. Rev.* 33, 782–794. doi: 10.1016/j.cpr.2013.06.002
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863
- Lebowitz, M. S., and Appelbaum, P. S. (2019). Biomedical explanations of psychopathology and their implications for attitudes and beliefs about mental disorders. *Annu. Rev. Clin. Psychol.* 15, 555–577. doi: 10.1146/annurev-clinpsy-050718-095416
- Lebowitz, M. S., Rosenthal, J. E., and Ahn, W.-K. (2016). Effects of biological versus psychosocial explanations on stigmatization of children with ADHD. *J. Atten. Disord.* 20, 240–250. doi: 10.1177/1087054712469255
- Liasidou, A. (2012). Inclusive education and critical pedagogy at the intersections of disability, race, gender and class. *J. Crit. Educ. Policy Stud.* 10, 168–184.
- Link, B. G., and Phelan, J. C. (2001). Conceptualizing stigma. *Annu. Rev. Sociol.* 27, 363–385. doi: 10.1146/annurev.soc.27.1.363
- Macharey, G., and Von Suchodoletz, W. (2008). Perceived stigmatization of children with speech-language impairment and their parents. *Folia Phoniatr. Logop.* 60, 256–263. doi: 10.1159/000151763
- Miller, E., Chen, R., Glover-Graf, N. M., and Kranz, P. (2009). Willingness to engage in personal relationships with persons with disabilities: examining category and severity of disability. *Rehabil. Couns. Bull.* 52, 211–224. doi: 10.1177/0034355209332719
- Möhring, M., Krick, N. K., and Ditzel, A. (2021). The influence of social support and characteristics of the stigmatizers on stigmatizing attitudes towards people with mental illness. *Europ. J. Health Commun.* 2, 66–96. doi: 10.47368/ejhc.2021.204
- Molden, D. C. (2014). Understanding priming effects in social psychology: what is "social priming" and how does it occur? *Soc. Cogn.* 32, 1–11. doi: 10.1521/soco.2014.32.supp.1
- Ort, A., and Fahr, A. (2018). Using efficacy cues in persuasive health communication is more effective than employing threats: an experimental study of a vaccination intervention against Ebola. *Br. J. Health Psychol.* 23, 665–684. doi: 10.1111/bjhp.12310
- Reips, U.-D. (2002). Standards for internet-based experimenting. *Exp. Psychol.* 49, 243–256. doi: 10.1027//1618-3169.49.4.243
- Röhm, A., Schnöring, A., and Hastall, M. R. (2018). Impact of single-case pupil descriptions on student teacher attitudes towards inclusive education. *Learn. Disabil.* 16, 37–58.
- Savolainen, H., Engelbrecht, P., Nel, M., and Malinen, O.-P. (2011). Understanding teachers' attitudes and self-efficacy in inclusive education: implications for pre-service and in-service teacher education. *Eur J. Spec Needs Educ* 27, 51–68. doi: 10.1080/08856257.2011.613603
- Schabmann, A., and Kreuz, A. (1999). Die Erfassung der Einstellungen gegenüber geistig behinderten menschen anhand der deutschsprachigen version des mental retardation attitude inventory-R [measuring attitudes towards people with mental retardation using a German version of the "mental retardation attitude inventory-R"]. *Heilpädagogische Forschung* 25, 174–183.
- Schwarzer, R., and Schmitz, G. S. (1999). "Skala zur Lehrer-Selbstwirksamkeitserwartung (WIRKLEHR), [Scale for teacher-self-efficacy (WIRKLEHR)]" in *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. eds. R. Schwarzer and M. Jerusalem (Berlin: Freie Universität Berlin), 60–61.
- Scior, K., Addai-Davis, J., Kenyon, M., and Sheridan, J. C. (2012). Stigma, public awareness about intellectual disability and attitudes to inclusion among different ethnic groups. *J. Intellect. Disabil. Res.* 57, 1014–1026. doi: 10.1111/j.1365-2788.2012.01597.x
- Scior, K., and Furnham, A. (2016). Causal beliefs about intellectual disability and schizophrenia and their relationship with awareness of the condition and social distance. *Psychiatry Res.* 243, 100–108. doi: 10.1016/j.psychres.2016.06.019
- Steinmayr, R., and Spinath, B. (2008). Sex differences in school achievement: what are the roles of personality and achievement motivation? *Eur. J. Personal.* 22, 185–209. doi: 10.1002/per.676
- UNESCO (2019). *Education: percentage of female teachers by teaching level of education*. Montreal: UNESCO. Available at: <http://data.uis.unesco.org/index.aspx?queryid=178> (Accessed March 31, 2022).
- Venville, A., Mealings, M., Ennals, P., Oates, J., Fossey, E., Douglas, J., et al. (2016). Supporting students with invisible disabilities: a scoping review of postsecondary education for students with mental illness or an acquired brain injury. *Int. J. Disabil. Dev. Educ.* 63, 571–592. doi: 10.1080/1034912X.2016.1153050

- Weiner, B. (1986). *Attributional Theory of Motivation and Emotion*. New York, NY: Springer, doi: 10.1007/978-1-4612-4948-1
- Weiner, B., Perry, R. P., and Magnusson, J. (1988). An attributional analysis of reactions to stigmas. *J. Pers. Soc. Psychol.* 55, 738–748. doi: 10.1037/0022-3514.55.5.738
- Wilson, M. C., and Scior, K. (2015). Implicit attitudes towards people with intellectual disabilities: their relationship with explicit attitudes, social distance, emotions and contact. *PLoS One* 10:e0137902. doi: 10.1371/journal.pone.0137902
- Wray, E., Sharma, U., and Subban, P. (2022). Factors influencing teacher self-efficacy for inclusive education: a systematic literature review. *Teach. Teach. Educ.* 117:103800. doi: 10.1016/j.tate.2022.103800
- Zensen, J., and Röhm, A. (2021). “Nicht-intendierte Stigmatisierung durch Ursachenzuschreibungen am Beispiel ADHS, [non-intentional stigmatization through causal attributions using the example of ADHD]” in *Risiken und Potenziale in der Gesundheitskommunikation: Beiträge zur Jahrestagung der DGPK-Fachgruppe Gesundheitskommunikation 2020*. eds. F. Sukalla and C. Voigt (Leipzig: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e.V.), 146–156.
- Zillmann, D. (2006). Exemplification effects in the promotion of safety and health. *J. Comm.* 56, S221–S237. doi: 10.1111/j.1460-2466.2006.00291.x
- Zillmann, D., and Brosius, H.-B. (2000). *Exemplification in Communication: The Influence of Case Reports on the Perception of Issues*. Mahwah, NJ: Lawrence Erlbaum Associates.



OPEN ACCESS

EDITED BY

Hyemin Han,
University of Alabama,
United States

REVIEWED BY

Jonathan Plucker,
Johns Hopkins University,
United States
Sofia Abreu Mendes,
Lusiada University of Porto, Portugal

*CORRESPONDENCE

Jesús de la Fuente
jdlfuente@unav.es

SPECIALTY SECTION

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

RECEIVED 27 May 2022

ACCEPTED 08 September 2022

PUBLISHED 07 November 2022

CITATION

de la Fuente J, Vera-Martínez MM,
Peralta-Sánchez FJ and
Martínez-Vicente JM (2022) A proposed
protocol for the registration of evidence-
based Educational Psychology programs.
Front. Psychol. 13:954475.
doi: 10.3389/fpsyg.2022.954475

COPYRIGHT

© 2022 de la Fuente, Vera-Martínez,
Peralta-Sánchez and Martínez-Vicente. 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.

A proposed protocol for the registration of evidence-based Educational Psychology programs

Jesús de la Fuente^{1,2*}, Manuel Mariano Vera-Martínez^{3,4},
Francisco Javier Peralta-Sánchez² and
José Manuel Martínez-Vicente²

¹School of Education and Psychology, University of Navarra, Pamplona, Spain, ²School of Psychology, University of Almería, Almería, Spain, ³School of Psychology, University of Granada, Granada, Spain, ⁴General Council for Psychology, Madrid, Spain

The turn to evidence-based interventions is significant for Psychology in general and for Educational Psychology in particular. Although there is a great deal of published evidence for program validation, there is currently no protocol for the evaluation of educational psychology evidence-based intervention programs and there is no General Register of such programs to act as a central information depository. This article has four objectives: (1) To assess the significance of the turn to *Evidence-Based Programs* in the context of today's Knowledge Society and Research and Development, Transfer and Innovation (R&Di) activities; (2) To provide a *Classification of Programs* based on the degree of specialization required for appropriate professional use in line with the requirements of the Code of Ethics in Psychology; (3) To tentatively propose a *Protocol for the Registration of Evidence-Based Educational Psychology Programs*; and (4) Finally, to identify some implications of the protocol. We conclude that there is a need for a *General Register* of validated programs.

KEYWORDS

evidence-based intervention, protocol for evaluating Educational Psychology intervention programs, national database, professional practice, general psychological society

Introduction

The development of Psychology as a science and of the discipline of Educational Psychology has led to a proliferation of conceptual models in each area of study and the development of assessment and intervention tools to implement those models in professional practice. A Google Scholar search for *Psychological Intervention Programs* gave some 360,000 results, *Educational Intervention Programs* produced some 4.5 million publications, and *Educational Psychology Intervention Programs* gave 3 million results.

Those figures demonstrate the large volume of work concerning the creation and use of Programs, properly so called.

The term *Intervention Program* describes a set of components or structured decisions and actions intended to have a particular effect in a given personal or social environment (Berdychevsky et al., 2022; Gijzen et al., 2022). A material proportion of published Psychological Programs have been created by members of the profession and university academics; publishing houses and university knowledge transfer technology companies have generally undertaken publication and sale. The trend is if anything accelerating with, for example, the current promotion by Spain's University System of new processes, products, and services by way of transfer to the wider community and the profession through the Spanish knowledge transfer program, *Six Years of Transfer*, 2012–2018 (Repiso et al., 2019). The European Union has also attributed considerable significance to technology transfer.

However, despite those developments and despite copious attempts to provide a basis for program validation (Sánchez-Meca et al., 2002), there is as yet no general protocol for the assessment of scientific quality or underlying scientific evidence so as to determine whether to endorse a Psychological Intervention Program. However, although much attention has been devoted to the evaluation, validation, and classification of psychological assessment tools by researchers (Moreno et al., 2018; Muñoz and Fonseca-Pedrero, 2019) and practitioners (Commission for Tests Edited in Spain, 2012), the same is not true of *Psychological Intervention Programs* in use in professional practice.

Consequently, this article has four objectives: (1) To assess the significance of the turn to *Evidence-Based Programs* in the context of today's Knowledge Society and Research and Development, Transfer and Innovation (R&Di) activities; (2) To provide a *Classification of Programs* based on the degree of specialization required for appropriate professional use consistent with the requirements of the Code of Ethics in Psychology; (3) To tentatively propose a *Protocol for the Registration of Evidence-Based Educational Psychology Programs* for Educational Psychology Intervention Programs; and (4) Finally, to identify some implications of the protocol. We conclude that there is a need for a *General Register* of validated programs.

The turn to evidence-based programs in the context of the knowledge society of the 21st century

Science in the knowledge society

The *Research, Development, and Innovation (R&Di)* chain as it operates in the Knowledge Society has achieved much, as we can see by considering the links in that chain individually (de la Fuente and Vera-Martínez, 2010; de la Fuente et al., 2012; Martínez-Vicente et al., 2020):

1. Research and the accumulation of scientific knowledge have provided progressively more evidence and led to the development of more accurate scientific models that have been generative of more scientific and more reliable predictions. For example, in Educational Psychology, there is a large body of scientific evidence in specific areas and conceptual models based on that evidence in each of those areas (Harris et al., 2012; de la Fuente and Justicia, 2018).
2. That has in turn led to technological development in the form of innovative assessment and intervention techniques based on that growing scientific knowledge. Strategies and techniques for assessment (Tests) and intervention (Programs) fall into this category: they are technological (in the true sense of the term) developments in a conceptual domain in a given area of knowledge and as such should be developed and applied scientifically. Consequently, the knowledge they are based on—the underlying theoretical model and other principles and assumptions—should be clear.
3. The transfer of such innovations for use in a given population requires the systematic collection and assessment of evidence to identify the results of an intervention. This closes the circle by allowing theoretical prediction and outcomes to be compared and contrasted.
4. *University Knowledge Transfer Departments* are the channel through which those activities are carried out. In essence, the mission of Knowledge Transfer Departments is to encourage research, development and (ultimately) knowledge transfer. They also promote the appropriate use of programs based on robust evidence that allows the effects of interventions to be predicted with a high level of confidence.

For all those reasons, the *Turn to Evidence-Based Programs* has become more prevalent.

The turn to evidence-based programs

The *Turn to Evidence-Based Program* has developed with the notion of scientific evidence in other areas (Hemsley-Brown and Sharp, 2003; Chaves and Ramírez, 2020; See, Walter et al., 2005; See et al., 2019) and in Psychology (de la Fuente et al., 2019b). It is a consequence of scientific maturity and adoption of the principle that psychological interventions developed in the R&Di value chain should be evidence-based (de la Fuente and Vera-Martínez, 2010, 2021; de la Fuente et al., 2019a).

In education, two recent seminal articles by Slavin (2017, 2019) lay out the progress that would be achieved by the level of evidence behind each intervention program being available to the educational system. He posits that such information would lead to: (1) well-founded decision-making by professionals; (2) routine accreditation of evidence of the effects of such programs as they are disseminated; and (3) the adoption of systematic evaluation of programs in the field of education. We would add: (4) a statutory obligation on creators

and businesses that market programs to provide validation information; (5) scientific validation of processes of intervention and consolidation of businesses and knowledge transfer departments in educational organizations of whatever kind, led by people who are qualified to produce empirical evidence and to test such processes (de la Fuente and Zapata, 2012).

The provision of evidence for any treatment or intervention program is a key to moving toward high-quality scientific psychological interventions. In many fields of Health Sciences (Villalbí, 2001) and in educational psychology (de la Fuente, 2020) it is increasingly accepted that interventions should be evidence-based. (For an explanatory video, see <https://www.youtube.com/watch?v=M70bPL63nvk>). The use of evidence-based interventions requires the adoption of criteria for the evaluation and review of intervention programs.

Slavin (2002, 2008, 2013, 2017, 2019) has promoted a national searchable database in the United States (Slavin, 2002, 2019) to contain the evidence and features of each intervention program. Such a database would provide significant support to decision-making in the selection of intervention programs. However, it raises a further, as-yet unresolved problem in relation to a Protocol of Evaluation Criteria for Evidence-Based Educational Psychology Programs (de la Fuente, 2020). That issue is explained below.

Types of programs: Educational, Psychopedagogical/Pedagogical, and Educational Psychology counseling

Although it may seem obvious, words are currently being used indiscriminately in a way that makes it hard to identify the area of study from which Educational Psychology Intervention Programs originate. Because of that lack of clarity, three types of intervention program are being used without their natures and scopes being fully clear.

For these purposes, an Intervention Program can be characterized as a suite of formal decisions, i.e., planned, explicit decisions that are contextualized, systematized, and intended to have specific effects. Those effects may be intended to arise within a family, a school, among teachers or students and concern the processes of development, learning, or teaching in any of those settings that in the aggregate constitute the process of education. We can categorize intervention programs as follows. Table 1 is a tabulation of this classification.

Educational intervention programs

Educational Intervention Programs fall conceptually and functionally under the umbrella of Educational Science, which is the suite of disciplines dedicated to the scientific study of different aspect of education in specific societies and cultures. That group

TABLE 1 Categorization of intervention program.

Program level		Concept	Domain	Focus	Delivered by
1.	Educational	Educational	School	Generalist	Subject teacher
2.	Pedagogical	Psychoped	Psychopedagogy	Specific	Pedagogical or Guidance Counselor
3.	Educational Psych	Educational Psych	Educational Psych	Specific	Educational Psychologist

of disciplines reflects the integration of all the sciences that may focus on a given educational problem, namely: Pedagogy, Comparative Pedagogy, General and Specific Didactics, Sociology, Anthropology, Philosophy, Economics, Theology (where applicable), and Psychology. Consequently, we can use the term Educational Intervention Program for any program that: (1) addresses an educational problem from an educational perspective; (2) falls within the combined field of any of those disciplines; and (3) is capable of being implemented by a professional educator with no specific qualifications beyond a teaching qualification and subject expertise in the curricular subject to which the Program relates. Most subject-specific educational programs come under this heading. Educational intervention programs are generally used by teachers in schools.

Pedagogical intervention programs

This category is made up of professional programs that are closer in concept to interventions as understood in Psychopedagogical (a classification used in Spain), Pedagogical, and Educational Guidance (Aubrey, 1988; Bisquerra et al. 1998). Although, the term “psychopedagogical” arguably merits an article of its own—not this one—it can be taken to mean three things: (1) the program addresses an educational problem from a psychopedagogical perspective, i.e., it includes elements from psychology; (2) the program is a professional guidance program; and (3) such programs tend to have a generalist educational science focus but with some specialist elements from pedagogy or psychology, without necessarily having a specifically educational psychology focus. This category includes many guidance-focused Intervention Programs created, proposed and adopted by Ministries of Education for use in school guidance programs delivered by secondary school teachers. It is school guidance counselors in particular who generally use pedagogical intervention programs.

It should be noted that Psychopedagogy in a sense makes Psychology and the other disciplines subordinate to Pedagogy and Didactics, its core disciplines. These programs are essentially pedagogical in focus, with some ancillary elements or insights from other disciplines, such as Psychology. Thus, such programs are not necessarily based on models, assessment techniques, or

interventions from within Psychology. This can be exemplified by comparing a Health Education Program drawn up as part of the general education of students (Ministry of Health of La Rioja, 2019) and a program that is expressly from within Educational Psychology. The Ministry of Education (Spain) has created a website related to Evidence-Based Education (Ministry of Education, 2020).

Educational Psychology intervention programs

This category is for programs whose conceptual foundations come from Developmental and Educational Psychology, such as the Psychology of Education and related psychological disciplines, namely: Developmental Psychology, Psychology of Learning Difficulties, and the Psychology of Vocational Development. These Programs are generally more specialist. They are specifically based on educational psychology theoretical models and on scientific evidence from each of the disciplines mentioned, such as evidence from *Educational Psychology* (de la Fuente and Justicia, 2018) and other specialist work in the field, whether in Europe or from the APA in the United States (Working Group IAP/APA, 2013). Their application requires: (1) specific educational psychology training or degree; (2) specialist knowledge of the explanatory models and the scientific evidence in the area; and (3) specific training in a particular Intervention Program. The requirements of the Code of Ethics of Psychology for qualifications in Psychology for those who implement psychological tests and techniques and programs for psychological intervention are of particular significance here (Consejo Oficial de la Psicología de España, 2008). Article 19 of the Code reads as follows: “Any Psychology materials, whether related to assessment and intervention or to treatment, may only be used by Psychologists. Psychologists may not provide such materials to unqualified people. Psychologists will manage or, as the case may be, ensure proper custody of documents pertaining to Psychology.” Consequently, these particular programs should be used by Educational Psychologists.

Essential features of a proposed protocol for the registration of evidence-based Educational Psychology programs

The protocol for the registration of evidence-based programs (de la Fuente, 2020) contains the following sections.

Professional context

Professional context

This first level of categorization is the general area of a Program. Whether a program is educational, psychopedagogical or from educational psychology is a threshold issue because the

protocol is, as explained, intended for educational psychology programs.

Intervention area of the program

Beyond that basic information, the area in which a program makes an intervention must be clearly described. In the case of the educational psychology programs for which the Protocol is intended, we need to know whether a given Intervention Program has indeed been designed for an educational psychology program.

Intervention programs to optimize processes of development, teaching, and learning

Interventions in the processes of human development, in programs focused on any aspect of human development (physical, psychomotor, personal or emotional, social-ethical, cognitive, and linguistic). If a Program concerns learning, it will seek to improve the learning process and the skill of learning to learn in any of its conceptual, procedural or attitudinal aspects. Programs to improve teaching also fall into this category.

Intervention programs to address difficulties in the processes of development and learning

This group is for all Intervention Programs specific to the adjustment, treatment and improvement of any difficulties and problems that arise in the macroprocesses of development and learning.

Academic and vocational guidance programs

Specific academic guidance and career-orientation programs are in this group.

Programs of educational psychology research, technology development, and innovation (R&Di)

This category includes all Intervention Programs associated with computer-supported innovation relating to the improvement of a process, product, or service within educational psychology: Apps, Computer Programs, etc.

Specific intervention area of the program

Level 3 is for the specific type of treatment within the broader Level 2 area. For example, an Emotional Intelligence program would be in the specific area of *Optimization of Personal and Social Development*. If a program is intended to support learning under stress conditions, it would be in the specific area of *Optimization of the Learning Process*.

Type of program

It is important to capture the nature of the Program in terms of the level at which intervention is made.

Preventive (primary)

Programs intended to optimize the processes of development or learning in their initial phases before the emergence of educational problems.

Palliative (secondary)

Programs aimed at rechanneling development processes when the first symptoms of maladjustment or difficulty have appeared.

Treatment (tertiary)

Programs aimed at intervening when problems have manifested.

Retrospective (stage 4)

Programs aimed at prevention of the unwanted side effects of prior interventions.

Program details

This section contains:

1. The basic details of a program: its name, the specific problem addressed, year of creation, the authors, and any registered intellectual property rights (IPR).
2. The structure of the Program, the materials for the Program, the publisher (if any), the English language version (if any), and any consents.
3. The details for use and transfer, the person to contact to use the program and any R&Di contracts when a Knowledge Transfer Department is involved.

Details of evidence-based validation

Details of program validation

Basis of the program: Theoretical model and supporting empirical evidence

Integral to an evaluation of an Intervention Program is an explicit statement of the underlying theoretical model and the supporting empirical evidence published in recognized international scientific publications. Such evidence is generally in the form of published peer-reviewed articles or manuals.

There are numerous models in educational psychology that have been confirmed by evidence. Programs can be analyzed by issue, field of intervention and how up to date they are. As a general proposition, reviews of specific research topics and manuals of research review (de la Fuente and Justicia, 2018) describe developments in different subject areas and in relation to different theoretical models.

Validation method

Published peer-reviewed reports of research and impact assessments that set out the effects of the Program and the general procedure for each study in sufficient detail for it to be replicable are also required for full evaluation of a Program. To that end, information is required as follows:

Study participants

Sample: inclusion and exclusion criteria, type of sample.

Instruments

Validity and reliability of instruments used to validate a program, whether part of the program itself or other instruments recognized by the scientific community.

Procedure

The details of the procedure should be sufficiently clear to enable replication in new samples. It is essential to provide Ethics Committee approvals and where applicable the Pre-Clinical Registration of the validation study.

Data design and analysis

The study design (Anguera et al., 1995; Ato et al., 2013) should be stated in sufficient detail and the data analysis should be consistent with the type of validation intended.

Results

The results should be presented clearly and unambiguously relative to the expected effects of Program implementation. This section, depending on the nature of the results submitted and the design, is key to determining the nature of the evidence provided and its strength.

References to publications that provide evidence

The list of published research should be provided. That will enable the sufficiency and adequacy of the evidence to be determined.

Final evaluation

Following Slavin (2019), there are different levels of evidence:

1. Strong evidence: Results are consistent and designs are robust (experimental or quasi-experimental) and well-implemented.
2. Medium evidence: Results are consistent but the design has a medium level of consistency (correlational designs).
3. Weak evidence: Results have consistency but the design is very weak with descriptive data.
4. Evidence non-existent: No results are provided that constitute evidence and the design is very weak.

Proposed protocol for the registration of evidence-based educational psychology intervention programs

Professional context

1. Professional field
 - 1.1 Optimization of processes of development, teaching, and learning (de la Fuente, 2020).
 - 1.2 Management of special educational needs and learning difficulties.
 - 1.3 Academic and career guidance.

- 1.4 Research, Technological Development, and Innovation in Psychopedagogy.
2. Specific field
 - 2.1. Optimization of processes of learning.
 - 2.2. Optimization of processes of development.
 - 2.3. Optimization of processes of teaching.
 - 2.4. Management of special educational needs in development processes.
 - 2.5. Management of special educational needs in learning processes.
 - 2.6. Academic guidance.
 - 2.7. Career guidance.
 - 2.8. Research, Technological Development and Innovation in Psychopedagogy.
3. Subprogram
 - 3.1. Psychomotor Development
 - 3.2. Personal (Emotional) Development
 - 3.3. Social-Ethical Development
 - 3.4. Cognitive Development
 - 3.5. Linguistic Development.
4. Type of program
 - 4.1. Preventive (primary prevention)
 - 4.2. Palliative (secondary prevention)
 - 4.3. Treatment (tertiary prevention)
 - 4.4. Retrospective (Stage 4 prevention)

Program details

1. Name of the Program:
2. Problem addressed:
3. Year of creation:
4. Authors:
5. Intellectual Property Registration:
6. Structure OF Intervention Program:
7. Materials:
8. Publisher (if any):
9. Previous English-language version (if any) and associated consents (if any):
10. Translation into English (if any):
11. Details of use and transfer:
12. Contact:
13. Existence of R&D contract for use (if any):

Details of evidence-based validation

1. Theoretical Basis. Theoretical models on which program is based. Recent evidence:
2. Validation method:
 - 2.1. Participants.
 - 2.2. Instruments.

- 2.3. Procedure.
- 2.4. Data design and analysis.
3. Results obtained: synthesis of results.
4. References to significant publications (peer review system):
5. Evaluation proposal (Slavin, 2019):
 - 5.1. Strong evidence:
 - 5.2. Medium evidence:
 - 5.3. Weak evidence:
 - 5.4. Evidence non-existent:

Exemplar completed protocol for the registration of evidence-based Educational Psychology intervention program: The ALADO program (2012)

Professional context

1. General area
 - 1.1. Optimization of processes of development, teaching, and learning (de la Fuente, 2020).
 - 1.2. Management of special educational needs and learning difficulties.
 - 1.3. Academic and career guidance.
 - 1.4. Research, Technological Development, and Innovation in Psychopedagogy.
2. Specific area
 - 2.1. Optimization of processes of learning.
 - 2.2. Optimization of processes of development.
 - 2.3. Optimization of processes of teaching.
 - 2.4. Management of special educational needs in processes development processes.
 - 2.5. Management of special educational needs in learning processes.
 - 2.6. Academic guidance.
 - 2.7. Career guidance.
 - 2.8. Research, Technological Development, and Innovation in Educational Psychopedagogy.
3. Subprogram
 - 3.1 Psychomotor Development.
 - 3.2 Personal (Emotional) Development.
 - 3.3 Social-Ethical Development.
 - 3.4 Cognitive Development.
 - 3.5 Linguistic Development.
 - 3.6 Language Learning.
 - 3.7 Mathematical Learning.
 - 3.8 Social Learning.
 - 3.9 Personal Learning.
4. Type of program
 - 4.1 Preventive (primary prevention).
 - 4.2 Palliative (secondary prevention).

4.3 Treatment (tertiary prevention).

4.4 Retrospective (Stage 4 prevention).

Program details

Name of the program: ALADO Program to Prevent Alcohol Abuse in Adolescents.

Problem: Prevention of Alcohol Consumption by Adolescents.

Year of Creation: 2012.

Authors: Lead researchers:

- Dr. Jesús Enrique de la Fuente Arias
- Dr. Inmaculada Cubero Talavera

Participating researchers:

- Dr. Francisco Javier Peralta Sánchez
- Dr. María Dolores Sánchez Roda

Intellectual Property:

- Teacher guide: ISBN: 978-84-938657-3-3
- Intervention guide: ISBN: 978-84-938657-2-6

home. It has been developed to be free-standing so that anyone who wants to can register and use the program content. That gives control to the family and allows the activities to be completed in the home.

Can parents participate?

ALADO has a parent-teacher space called Parent-Teacher School with materials that explain and train about the effects of alcohol consumption and the different ways that the problem is addressed at national and regional level. It also offers links to websites and internet portals concerning the prevention of alcohol consumption. Parents and teachers can find valuable reading material for active involvement in the education of their children in a culture based on saying NO to alcohol. By way of example from that material:

- The Guide to Prevention from the National Drug Addiction Plan.
- Alcohol: Clinical Committee reports.
- NIDA prevention program.
- It is always your choice.
- Guidance for SPAN parents and strategies for taking action.

Structure of the program

ALADO is an online program for the prevention of the consumption and abuse of alcohol. It is aimed at young people aged 12–16 and is split into nine modules that progressively educate participants about the undesirable effects of alcohol consumption. The program provides the essential tools to enable young people to reduce or avoid the consumption of alcohol through management of their own behavior.

Objectives of the program

- To make participants aware of the effects of alcohol consumption for adolescents, particularly, on their brains. The program seeks to decisively change attitudes to alcohol and to give participants strategies to regulate their own behavior so that they can avoid excessive alcohol consumption. We offer participants mechanisms to access pertinent information through a fun, attractive online platform, which has been designed and structured to appeal to young people.
- To teach young people that alcohol is a drug that has negative effects on our health.
- To encourage abstinence from alcohol.
- To teach and offer behavioral strategies to truly say NO to alcohol, particularly in social situations where there is peer pressure.

How it works

The program is conceived as a school activity to be carried out in the classroom and includes help and guidance for tutors and teachers. However, ALADO is available online and so can also be used in the

Origins of ALADO

Adolescence is a developmental stage of biological and cognitive maturation, which is characterized by experimental, risky behaviors. That prototypical behavior pattern is perfectly reflected in the unfortunately notorious pattern of heavy weekend binge drinking in which abstinence during the week is followed by abuse at the weekend. Although an adolescent might not always see that pattern of alcohol consumption as risky to their health, science has shown that a habit of intermittent intake in a consumption-relapse pattern is associated with neurobehavioral changes, with a tendency toward increased alcohol consumption and an increased risk of addiction in adulthood. The most recent research has shown that young people are unfortunately less susceptible to alcohol intoxication. In other words, adolescents are more resistant to alcohol's sedative effect and so may drink larger volumes in less time...“an adolescent can take more...” Paradoxically and despite their apparent non-susceptibility to its sedative effects, alcohol affects the adolescent brain more severely than the adult brain, because young brains are still developing.

ALADO at the University of Almería

Concerned at the alarming rise in alcohol abuse in our young people, a group of experts working in neurobiological research and educational psychology at the University of Almería joined forces and knowledge on ALADO, a project to develop an ALcohol prevention program for Adolescents (“with wings” in Spanish).

Financed by the Ministry of Innovation, Science and Business of the Regional Government of Andalusia, the aim of the program is to bring about change in capacity and attitudes concerning excessive alcohol consumption in adolescents. ALADO is a pilot

that seeks to transfer knowledge derived from research to society, to a group as highly vulnerable as adolescents.

Information and education online

This ambitious project aims to develop the first completely free-standing software for secondary school students, parents and teachers to prevent the consumption and abuse of alcohol. The strategy is to give students the necessary information and tools to be able to reflect on the effects of alcohol consumption and to prompt a change in attitudes toward alcohol.

As a pilot project, the expert group in Almería implemented the program in secondary schools to work with students in the 12–16 age range before they had become habituated to the consumption of alcohol. ALADO was implemented in the classroom in curriculum time over a period of 1 year.

Using appropriate evaluation tools, the effectiveness of ALADO was assessed using a pre-post test of the attitudes and behavior of our adolescents before and after their experience of ALADO.

Materials

- www.alado.es (<https://www.miguelturra.es/programa-informatico-pretende-reducir-ingesta-alcohol-adolescentes-wwwaladoes>)
- Printed material: de la Fuente et al. (2012).
- Publisher: Publication associated with the Electronic Journal of Research in Educational Psychology. <http://www.investigacion-psicopedagogica.org/revista/new/index.php>. Education & Psychology R&DI (Spin-Off. University of Almería, Spain).
- Previous English language version: No.
- Translated into English: No.
- Details of use and transfer: Contact: jfuente@ual.es; icubero@ual.es; fjperalta@ual.es; R&D Use Agreement: From OTRI at the University of Almería.

Details of evidence-based validation

Theoretical basis: Theoretical model

Model of Competence to interact with alcohol (de la Fuente et al., 2017). The program consists of a total of 24 lessons that facilitate computer-enabled conceptual, procedural and attitudinal learning.¹

Validation method

Participants

A total of 148 adolescents aged 12–16 took part in three Spanish schools with different social and cultural contexts.

Instruments

- *The Scale for the Evaluation of Facts, Concepts, and Principles concerning Alcohol* (EHCP, in Spanish; Cubero and Sánchez, unpublished).² The scale is made up of 38 items related to the effects of alcohol consumption; psychometric analysis of the scale shows reliability ($\alpha=0.827$) and consistent construct validity, with three factors: knowledge of facts, concepts, and principles concerning alcohol.
- *The SRQ Self-Regulation Questionnaire* in its 21-item abridged form for adolescents (Neal and Carey, 2005) was used. Spanish version of the abridged SRQ, SRQ-21 (Pichardo et al., 2018). Its reliability ($\alpha=0.826$) and validity figures are consistent, with two dimensions: planning and control of actions.
- *The Scale for Assessment of Attitudes to Alcohol* (AAA; Cubero and Sánchez, 2018, unpublished). A total of eight items to evaluate attitudes and values toward alcohol ($\alpha=0.825$).
- *The Adjustment Scale in interaction with alcohol*, which contains four times ($\alpha=0.915$). This scale belongs to the Inventory of Knowledge, Attitudes, and Interaction with Alcohol (Cubero and Sánchez, unpublished).

Procedure

The data-gathering instruments were used during the course of the 2009–2010 academic year as part of the Alado Excellence Project (2007–2010) through an online tool created for the purpose.³ Although the data are old, we consider that given their nature they are still highly relevant to the analysis of the variables assessed.

Consent to cooperation by participating schools and pupils was requested from each School Council, the teachers involved and the families (parents) of participating students. The project was approved by the Bioethical Committee of the University of Almería and by the School Council of each participating school. The students participated voluntarily. Parents were informed in writing. Since the participants in the project were still minors, both the parents and the management of the schools gave written informed consent to the research. Data were protected in a registered archive as required by the Spanish Data Protection Act.

Data design and analysis

We used a quasi-experimental method with repeat measurement and inferential analysis (ANOVA and MANOVA).

¹ www.alado.es

² Cubero, I. and Sánchez, M. C. (2018a). *Scale for the Evaluation of Facts, Concepts, and Principles concerning Alcohol* (EHCP). Manuscript pending publication Almería: University of Almería.

Cubero, I. and Sánchez, M. C. (2018b). *Scale for Assessment of Attitudes to Alcohol* (AAA). Manuscript pending publication. Almería: University of Almería.

³ www.alado.es

Results

The results showed a principal effect for the variable Treatment, the variable Level of Self-Regulation and an interaction effect for Treatment \times Self-Regulation in the conceptual and attitudinal subcompetency for interaction with alcohol. The results are discussed relative to new technology that allows evaluation and intervention in the prevention of alcohol consumption in adolescents. An important implication of this work is the importance of the psychological variable self-regulation. It also supports the suitability of educational psychology interventions in new technology formats in the prevention of alcohol consumption in adolescents as a commercial activity.

References for publications in the evidence report: de la Fuente et al. (2017, 2019) and Marcos-Pérez (2012).

Final evaluation proposal (Slavin, 2019):

Strong evidence:

Medium evidence:

Weak evidence:

Evidence non-existent:

Conclusion and discussion: The future of programs in Educational Psychology

In summary, it seems reasonable following Slavin (2019) to propose the adoption of a *Protocol for the Registration of Evidence-Based Educational Psychology Programs*. The proposal made here is preliminary and could be adapted to other areas of Psychology.

It would also be necessary to create a *National Register of Evidence-Based Educational Psychology Programs*, similar to the register that exists for *Psychological Tests* (Commission for Tests Edited in Spain, 2012). Some existing models in the field of education, e.g., Best-Evidence Encyclopedia; BEE; <http://www.bestevidence.org> and ESSA, <https://www.evidenceforessa.org> could provide guidance. The Register should be based on anonymous peer assessment by researchers and practitioners with accredited experience in the scientific review of Psychological Programs. It is evident that such assessment would give added value to existing Intervention Programs and allow published programs that have no supporting evidence to be distinguished from those that do have such evidence. It could also act as an endorsement in processes of academic and professional accreditation. It would allow users to distinguish among the many

Intervention Programs available (in hard-copy and/or remotely) those that are not evidence-based – and have therefore not been validated – from those that have been. In summary, the proposal made in this article would be a step forward in the scientific improvement of the profession of Psychology in general and Educational Psychology in particular (de la Fuente, 2020).

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 the authors have contributed equally, in the editing and revision of this work. JF and MMV-M made the initial drafting of the protocol. JMM-V and FJP-S made the final corrections and the adaptation of the protocol to the ALADO program. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Project I+D University of Navarra (PGC2018-094672-B-I00; Ministry of Science and Education, Spain) and University of Almería (European Social Fund) Project I+D (AL18-SEJ-DO31-A-FEDER).

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

- Anguera, M. T., Arnau, J., Ato, M., Martínez, R., Pascual, J., and Vallejo, G. (1995). *Métodos de Investigación en Psicología (Research Methods in Psychology)*. Madrid: Síntesis
- Ato, M., López-García, J. J., and Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología (A system of classification of study

designs in psychology) *Anales de Psicología/Annals of Psychology* 29, 1038–1059. doi: 10.6018/analesps.29.3.178511

Aubrey, R. (1988). "A house divided: guidance and counseling" in *Orientación e Intervención Psicopedagógica*. ed. C. Vélaz de Medrano (Málaga: Aljibe: Conceptos, Modelos, programas y Evaluación), 128–135.

- Berdychevsky, L., Stodolska, M., and Shinew, K. J. (2022). The roles of recreation in the prevention, intervention, and rehabilitation programs addressing youth gang involvement and violence. *Leis. Sci.* 44, 343–365. doi: 10.1080/01490400.2019.1571966
- Bisquerra, R., Monescillo, M., and Méndez, J. M. (1998). “Origins and Development of Orientation,” in *Modelos de orientación e intervención psicopedagógica*. ed. R. Bisquerra-Alzina (Barcelona: Praxis), 23–40.
- Chaves, P. C., and Ramírez, A. A. (2020). Políticas Públicas Basadas en Evidencia: una revisión del concepto y sus características (published evidence-based policies: a review of the concept and its features). *Rev. ABRA* 40, 49–76. doi: 10.15359/abra.40-60.2
- Commission for Tests Edited in Spain (2012). Evaluation of tests edited in Spain. Consejo General de la Psicología. Available at: <https://www.cop.es/index.php?page=evaluacion-tests-editados-en-espana>
- Consejo Oficial de la Psicología de España (2008). Código Deontológico de la Psicología [Code of ethics of psychology]. Available at: <https://www.cop.es/pdf/Codigo-Deontologico-Consejo-Adaptacion-Ley-Omnibus.pdf>
- de la Fuente, J. (2020). Protocolo para la Evaluación de los Programa de Intervención Psicoeducativa basados en la evidencia (Protocol for the Evaluation of Evidence-Based Educational Psychology Programs). Colegio Oficial de la Psicología de Madrid [Official College of Psychology of Madrid]: Conference (Spanish). Available at: <https://youtu.be/M70bPL63nvk>
- de la Fuente, J., Cubero, I., Peralta-Sánchez, F. J., Sánchez, M. C., Salmeron, J. L., and Fadda, S. (2019). Characteristics, structure, and effects of an on-line tool for improvement in adolescents’ competency for interaction with alcohol: The e-ALADOTM utility. *Front. Psychol.* 10:127. doi: 10.3389/fpsyg.2019.00127
- de la Fuente, J., Cubero, I., Sánchez-Amate, M. C., Peralta, F. J., Garzón, A., and Fiz Pérez, J. (2017). The Adolescent’s competency for interacting with alcohol as a determinant of intake: the role of self-regulation. *Front. Psychol.* 8:1800. doi: 10.3389/fpsyg.2017.01800
- de la Fuente, J., and Justicia, F. (2018). *Tópicos de Investigación Reciente en Psicología de la Educación [Recent research topics in Educational Psychology]*. Madrid: Gunti-EOS
- de la Fuente, J., Kauffman, D. F., and Diaz-Orueta, U. (Eds.) (2019a). *Psychology, Technological Innovation, and Entrepreneurship*. Lausanne: Frontiers Media SA
- de la Fuente, J., Kauffman, D. F., Diaz-Orueta, U., and Kauffman, Y. (2019b). Adapting the research development and innovation (RD & I) value chain in psychology to educational psychology area. *Front. Psychol.* 9:1188. doi: 10.3389/fpsyg.2018.01188
- de la Fuente, J., Peralta-Sánchez, F. J., Sánchez-Roda, M. D., Cubero-Talavera, I., and Sánchez-Amate, M. D. (2012). *Alado: Guía Docente*. Almería: Education & Psychology
- de la Fuente, J., and Vera-Martínez, M. M. (2010). Educational Psychology: psychology and I&D&I: a strategic action Program for the twenty-first century. *Papel. Psicol.* 31, 162–170.
- de la Fuente, J., and Vera-Martínez, M. (2021). Propuesta de Protocolo para el Registro de Programas Basada en la Evidencia. *Rev. Enginy* 21, 11–30.
- de la Fuente, J., Vera-Martínez, M. M., and Ellawar, M. C. (2012). Contribution from psychology of innovation and entrepreneurship, in today’s knowledge society. *Electron. J. Res. Educ. Psychol.* 10, 942–966.
- de la Fuente, J., and Zapata, L. (2012). *Creación de un departamento de I+D+I en el ámbito educativo (Creation of an R&D+I Department in an Educational Setting)*. Almería: Education & Psychology I+D+I, Spin-Off Universidad de Almería
- Gijzen, M. W., Rasing, S. P., Creemers, D. H., Engels, R. C., and Smit, F. (2022). Effectiveness of school-based preventive programs in suicidal thoughts and behaviors: a meta-analysis. *J. Affect. Disord.* 298, 408–420. doi: 10.1016/j.jad.2021.10.062
- Harris, K. R., Graham, S., and Urdan, T. (Eds.) (2012). *APA Educational Psychology Handbook (Vols. 1, 2 and 3)*. NY: America Psychological Association
- Hemsley-Brown, J., and Sharp, C. (2003). The use of research to improve professional practice: a systematic review of literature. *Oxf. Rev. Educ.* 29, 449–471. doi: 10.1080/0305498032000153025
- Marcos-Pérez, E. (2012). Competencia para la interacción con el alcohol en Adolescentes. Efectos de una e-Intervención [Competence for interaction with alcohol in adolescents. Effects of an e-Intervention]. Tesis Doctoral. Universidad de Almería. Available at: <http://www.diegomarin.net/ual/es/psicologia/472-competencia-de-interaccion-con-el-alcohol-en-adolescentes-efectos-de-una-e-intervencion-9788415487708.html>
- Martínez-Vicente, J. M., de la Fuente, J., Vera-Martínez, M. M., and Aleixandre, M. (2020). *Innovación y Emprendimiento en Psicología [Innovation and Entrepreneurship in Psychology]*. Almería: Servicio de Publicaciones de la Universidad de Almería
- Ministry of Education (2020). Educación guiada por la evidencia (Evidence-guided education). Available at: <https://www.fecyt.es/FECYTedu>
- Ministry of Health of La Rioja (2019) Programas de Educación para la Salud (Health Education Programs). Available at: <https://www.riojasalud.es/rs/docs/PEPS.pdf>
- Moreno, R., Martínez, R. J., and Muñoz, J. (2018). Test item taxonomy based on functional criteria. *Front. Psychol.* 9:1175. doi: 10.3389/fpsyg.2018.01175
- Muñoz, J., and Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test (Ten steps to build a test). *Psicothema* 31, 7–16. doi: 10.7334/psicothema2018.291
- Neal, D. J., and Carey, K. E. (2005). A follow-up psychometric analysis of the self-regulation questionnaire. *Psychol. Add. Behav.* 19:414. doi: 10.1037/0893-164X.19.4.414
- Pichardo, M. C., Cano, F., Garzón-Umerenkova, A., De la Fuente, J., Peralta-Sánchez, F. J., and Amate-Romera, J. (2018). Self-regulation questionnaire (SRQ) in Spanish adolescents: factor structure and rasch analysis. *Front. Psychol.* 9:1370. doi: 10.3389/fpsyg.2018.01370
- Repiso, R., Aguaded, I., Castillo-Esparcia, A., and Montero, J. (2019). Sexenio de transferencia: luces y sombras de un proyecto necesario. Análisis de una muestra de denegaciones en Ciencias de la Información y estudio de la edición de revistas y comisiones de agencia Como mérito (Six years of transfer: light and shadow of a necessary project. Analysis of a sample of rejections in information science and review of the editing of journals and agency commissions as merits). *Anuario Think EPI* 14:e14a02. doi: 10.3145/thinkepi.2020.e14a02
- Sánchez-Meca, J., Boruch, R. F., Petrosino, A., and Rosa-Alcázar, A. I. (2002). La colaboración Campbell y la práctica basada en la evidencia (The Campbell Collaboration and evidence-based practice). *Papel. Psicol.* 28, 1–10.
- See, B. H., Gorard, S., Morris, R., and Siddiqui, N. (2019). Evaluation of the National Literacy Trust’s literacy for life program, project report. Durham University: Evidence Centre for Education, Durham.
- Slavin, R. E. (2002). Evidence-based education policies: transforming educational practice and research. *Educ. Res.* 31, 15–21. doi: 10.3102/0013189X031007015
- Slavin, R. E. (2008). Evidence-based reform in education: what will it take? *Eur. Educ. Res. J.* 7, 124–128. doi: 10.2304/eeerj.2008.7.1.124
- Slavin, R. E. (2013). Overcoming the four barriers to evidence-based education. *Educ. Week* 32, 24–32.
- Slavin, R. E. (2017). Evidence-based reform in education. *J. Educ. Stud. Place. Risk* 22, 178–184. doi: 10.1080/10824669.2017.1334560
- Slavin, R. E. (2019). How evidence-based reform will transform research and practice in education. *Educ. Res.* 55, 21–31. doi: 10.1080/00461520.2019.1611432
- Villalbí, J. R. (2001). Promoción de la salud basada en la evidencia (promotion of evidence-based health). *Rev. Esp. Salud Pública* 275, 489–490.
- Walter, I., Nutley, S., and Davies, H. (2005). What works to promote evidence-based practice? A cross-sector review. *Evid. Policy* 1, 335–363. doi: 10.1332/1744264054851612
- Working Group IAP/APA (2013). IAP/APA evidence-based guidelines for the management of acute pancreatitis. *Pancreatolgy* 13, e1–e5. doi: 10.1016/j.pan.2013.07.063



OPEN ACCESS

EDITED BY

Claudio Longobardi,
University of Turin,
Italy

REVIEWED BY

Gianluca Serafini,
San Martino Hospital (IRCCS), Italy
Michael Musker,
University of South Australia, Australia

*CORRESPONDENCE

Mingyi Zhao
zhao_mingyi@csu.edu.cn
Quan Zhuang
zhuangquansteven@163.com

[†]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 04 July 2022

ACCEPTED 08 November 2022

PUBLISHED 23 November 2022

CITATION

Yang Y, Cao Q, Zhao M and
Zhuang Q (2022) Knowledge mapping of
students' mental health status in the
COVID-19 pandemic: A bibliometric study.
Front. Psychol. 13:985866.
doi: 10.3389/fpsyg.2022.985866

COPYRIGHT

© 2022 Yang, Cao, Zhao and Zhuang. 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.

Knowledge mapping of students' mental health status in the COVID-19 pandemic: A bibliometric study

Yang Yang^{1†}, Qingtai Cao^{2†}, Mingyi Zhao^{1*} and Quan Zhuang^{2,3,4*}

¹Department of Pediatrics, The 3rd Xiangya Hospital, Central South University, Changsha, Hunan, China, ²Transplantation Center, The 3rd Xiangya Hospital, Central South University, Changsha, Hunan, China, ³Department of Educational Affairs, The 3rd Xiangya Hospital, Central South University, Changsha, Hunan, China, ⁴Research Center on Transplantation Medicine of National Health Ministry, Changsha, Hunan, China

Objective: The purpose of this study was to investigate the international scientific output on mental health of students during COVID-19 from 2020 to 2022 through a bibliometric analysis and to explore trend and research hotspots in this field.

Methods: We searched the Web of Science Core Collection for publications and used a variety of software to analyze and visualize the data such as R, CiteSpace, VOSviewer and Scimago.

Results: A total of 2,734 publications were retrieved as of June 4, 2022, published by 3,894 institutions from 120 countries/regions. China and the United States lead in the quantity and quality of publications in this field. According to Bradford's Law, 16 journals are considered core journals in the field. Co-cited references indicate the main psychological problems of students under the epidemic revolve around anxiety, poor sleep and financial difficulty. Their behavior might also be influenced by increased internet and alcohol use.

Conclusion: Mental health of students during COVID-19 is attracting increasing attention. It is identified that the research hotspots in this field continue to revolve around emotional anxiety and unhealthy behaviors. Due to the different troubles faced by different groups under COVID-19, further exploration of the relevant factors specific for students are needed, with a hopeful view to providing ideas for intervention measures.

KEYWORDS

COVID-19, student, mental health, psychology stress, bibliometrics

Introduction

The coronavirus disease 2019 (COVID-19) is an emerging epidemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; Wang D. et al., 2020; Hu et al., 2021). Since 2019, the rapidly spreading COVID-19 pandemic has caused a global health threat (Fang et al., 2022). It is important to note that the COVID-19 pandemic is not only

taking a toll on people's physical health, but also has a negative impact on people's mental health (Zhang and Chen, 2021; Pappa et al., 2022; Zhang et al., 2022).

As a vulnerable group, students are relatively more prone to negative psychological symptoms (Cao et al., 2020; Copeland et al., 2021). Numerous studies have shown that during the COVID-19 pandemic, students' mental health has been severely challenged, regardless of their academic stage. A cross-sectional study using data from the Mental Health Survey of School-aged Children and Adolescents in Guangdong Province, China, showed that the prevalence of self-reported psychological distress among students was relatively high during the COVID-19 pandemic (Qin et al., 2021). A large-scale survey of 746,217 college students revealed that acute stress, anxiety, and depressive symptoms are prevalent during the COVID-19 pandemic, and that multiple epidemiological and psychosocial factors are associated with an increased risk of mental health problems (Ma et al., 2020). The mental health of master's, doctoral and postdoctoral researchers has also been severely negatively impacted by the pandemic (Salim, 2021). This study aims to explore mental health-related conditions of students during the COVID-19 outbreak. We hypothesize that students' mental health has been widely challenged during the COVID-19 pandemic and needs more care and support.

Over the past 2 years, a raft of research has been published on student mental health during the COVID-19 pandemic. However, there is no literature that systematically evaluates the relevant published literature. Bibliometrics uses mathematical and statistical methods to quantitatively analyze a large amount of literature in a specific research field, exploring research aspects and research trends in that field (Contreras-Barraza et al., 2021; Shen et al., 2022). Many researchers have used bibliometrics to evaluate their fields of study (Contreras-Barraza et al., 2021; Darroudi et al., 2021; Peng et al., 2022). However, no specific bibliometric studies have been conducted to date on the knowledge graph of student mental health during the COVID-19 pandemic. We have used bibliometrics to understand what is the current state of research work in this field, and question what are the research priorities and new trends? And further find out new problems that should be addressed and possible directions to solve the problem. Therefore, we used R software, Scimago software, and CiteSpace software to evaluate the literature on psychological stress of students during the COVID-19 period up to June 4, 2022 by bibliometric method, and generated knowledge maps. The results of these analyses can describe the current state of the field and identify new research directions.

Materials and methods

Literature retrieval

The literature retrieval was performed on June 4, 2022, using Web of Science Core Collection (WoSCC) to search for

publications on research on psychological stress among students during COVID-19. The search terms we used were as follows: Topic: (COVID-19 OR SARS-CoV-2 OR 2019-nCov) and (mental health OR psychological stress) and (student OR pupil OR undergraduate OR freshman OR sophomore OR junior OR senior OR graduate OR master OR doctoral student). Among them articles and review articles were included in the study. All the records retrieved from the WoSCC were downloaded independently by two authors (QC and YY). A total of 2,734 documents published from 2020 to 2022 met the search criteria for this study.

Data collection and analysis

We exported a complete data record from WoSCC, including the number of annual publications, countries, journals, citations, impact factor (IF) and Hirsch index (H-index), etc. Firstly, the top 10 disciplines with annual publications were chosen to investigate the major study categories in this field. Accumulative publication, average citations per publication and H-index of top 10 countries with the most publications were analyzed to evaluate the scientific impact of the countries/regions. R version 4.0, a free software environment for statistical computing and graphics, was used to show the above results.

Cooperative networks between countries/regions and institutions were visualized by Scimago software. Co-cited author analysis was conducted, which was further clustered by CiteSpace. CiteSpace V software was used to conduct co-citation analysis of the references and further to group the references into 16 clusters. Next, a timeline view of co-cited references was constructed. The journals were also clustered by CiteSpace algorithm and the core journals were selected according to Bradford's Law. CiteSpace could capture keywords with strong citation bursts and construct visualization maps of all items. A citation burst is a key indicator for identifying emerging trends (Chen et al., 2020). We used double image overlay to emphasize the key words of 2022 in CiteSpace. Conceptual analysis was performed using the R package "Bibliometrix".¹

Results

The overall trend of publications from 2020 to 2022

Information visualization was used to analyze articles on psychological stress among students during COVID-19 from WoSCC database from Jan 1, 2020 to June 4, 2022. The total number of cumulative publications in 2020, 2021, and 2022 was 439, 2,202, and 2,734, respectively (Figure 1A). The

¹ <https://www.bibliometrix.org/home/>

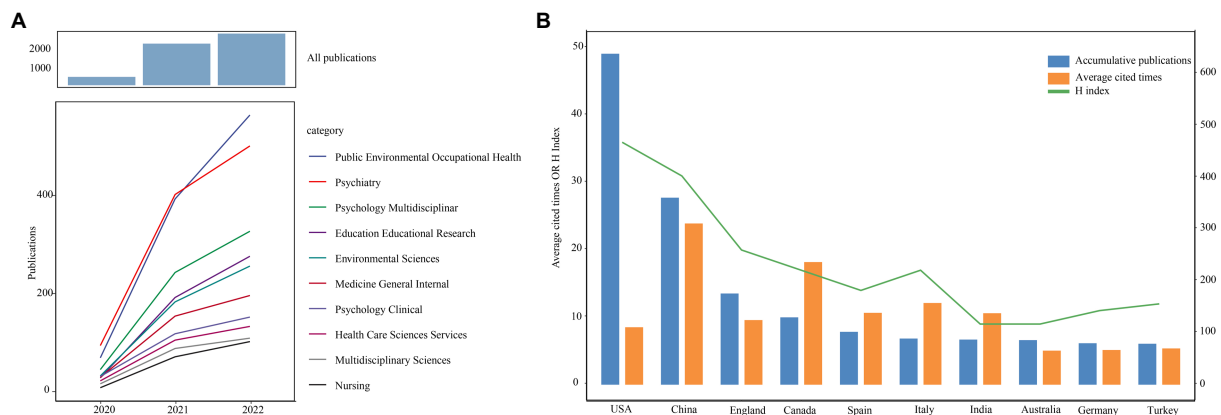


FIGURE 1 The overall distribution of publications. **(A)** The annual accumulative publications were exhibited by year and discipline. **(B)** The total number, average cited time and H-index of publications in countries.

TABLE 1 Top 10 co-cited references in mental health of students in COVID-19.

Rank	Count	Year	Author	Title
1	494	2020	Cao WJ	The psychological impact of the COVID-19 epidemic on college students in China
2	348	2020	Brooks SK	The psychological impact of quarantine and how to reduce it: rapid review of the evidence
3	288	2020	Wang CY	Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China
4	193	2020	Son C	Effects of COVID-19 on College Students' Mental Health in the United States: Interview Survey Study
5	161	2020	Xiong JQ	Impact of COVID-19 pandemic on mental health in the general population: A systematic review
6	151	2020	Qiu JY	A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations
7	147	2020	Holmes EA	Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science
8	147	2020	Sahu P	Closure of Universities Due to Coronavirus Disease 2019 (COVID-19): Impact on Education and Mental Health of Students and Academic Staff
9	146	2020	Odriozola-Gonzalez P	Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a Spanish university
10	145	2020	Wang CY	A longitudinal study on the mental health of general population during the COVID-19 epidemic in China

number of publications in 2021 was around four times than that of 2020, indicating that more attention in this field was obtained after the COVID-19 outbreak and high-quality articles erupted in 2020. Next, the co-cited reference analysis was carried out to explore which articles distinctively contributed to the development of this area. The most 10 co-cited articles were all published in 2020, the first three of which were authored by Cao WJ, Brooks SK and Wang CY, respectively (Table 1).

The disciplines of this area were mainly about “psychology,” “environment” and “medicine.” The disciplines with the most publications are “public environmental occupational health” and “psychiatry,” the annual publication trend of which were consistent to the total publications (Figure 1A).

Country/region publication and collaboration analysis

All publications in the field were distributed among 3,894 institutions from 120 countries/regions, of which the production of the United States ranked the first with 749 documents by far, followed by China (445), the United Kingdom (216), Canada (156) and India (120) (Figure 1B, Table 2). The whole trend of H index was consistent with that of all publications where the United States (40), China (33) and the United Kingdom (21) ranked the first three. With reference to the average citations per articles, different from H index, China (20.82) and Canada (18.28) showed the highest average cited times while the United States (8.6) was relatively lower. Overall,

TABLE 2 The top 10 productive countries/regions related to mental health in students under COVID-19.

Rank	Countries/ Region	Publication	Average per item	H-Index
1	the United States	749	8.6	40
2	China	445	20.82	33
3	The United Kingdom	216	9.09	21
4	Canada	156	18.28	17
5	India	120	8.02	9
6	Italy	119	12.04	19
7	Spain	119	10.5	15
8	Australia	114	9.53	13
9	Turkey	104	4.76	13
10	Germany	98	5.15	11

the United States and China are the more enthusiastic and leading in research in this field.

Next, the inter-country collaborations and inter-agency partnership were investigated. Figure 2A presents the analysis of cooperation between countries. The countries/regions that cooperated closely were clustered together and given the same color. The thicker the line, the closer the cooperation between the two countries. The results showed that China and the United States had the closest cooperation in this area. The collaboration among the United States, Canada and the United Kingdom were also relatively closer, consistent to the publication trend. A network diagram showed partnerships between major institutions (Figure 2B). Circles represent institutions, and lines represent partnerships. The deeper the circle, the more papers the institution published in the field, and the thicker the line, the closer the cooperation between the two institutions. The strongest collaboration relationship was found between The University of Melbourne and Monash University.

Co-cited author analysis

The selected 2,734 publications were produced by 12,889 authors and the top 10 productive authors and co-cited authors were listed in Table 3. Mark D Griffiths headed with 10 documents, followed closely by Teris Cheung, Yutao Xiang and Jingbo Zhao who all published 9 articles. The most frequently co-cited author was Wang CY with 565 co-cited times, followed by Cao WJ (550) and Brooks SK (455). These three authors were also the authors who wrote the most co-cited articles in this field (Table 1), revealing the significant effect of them. The author's co-cited network analysis was visualized in Figure 3. Subsequently, we performed a cluster analysis of co-cited authors (Figure 3B). Most authors were clustered into “#0 Cross-sectional online survey study,” indicating the research in this field mainly focused on online questionnaires and cross-sectional surveys.

Reference co-citation analysis

A total of 41,321 co-cited references were visualized by CiteSpace with time slice set as a year (Figure 4). The references were clustered according to the log-likelihood ratio (LLR) algorithm, and 16 clusters were obtained (Figure 4, Table 4). The largest cluster is “college student” (cluster #0), followed by “mood state” (cluster #1), “poor sleep quality” (cluster #2), and “physical activity” (cluster #3). It is not surprising that college students were the most frequently studied group as it's easier to collect their questionnaires. We further explored the cluster 1 and cluster 2, which both reflected the mental situation of students facing COVID-19. Tang WJ (89), Kaparounaki CK (80), Liu CH (62), Lei L (51) and Cellini N (39) were the most cited authors in “mood state,” and Husky MM (99), Auerbach RP (67), Zhang Y (66), Ammar A (45) and Stanton R (40) were the most cited authors in “poor sleep quality.” It was highly recommended to read these authors' documents if related research needed to be conducted. Besides, “cluster 4: alcohol consumption,” “cluster 7: online learning” and “cluster 8: medical students” were of interest and the related authors were shown in Supplementary Table 1.

Afterwards, we constructed a timeline view of co-cited references to illustrate topic distribution in the field, and also investigated topic trends and interrelationships over time (Figure 5). In the timeline view, we found a sharp increase in the number of documents per cluster since 2019. It is apparent that this resulted from the outbreak of COVID-19.

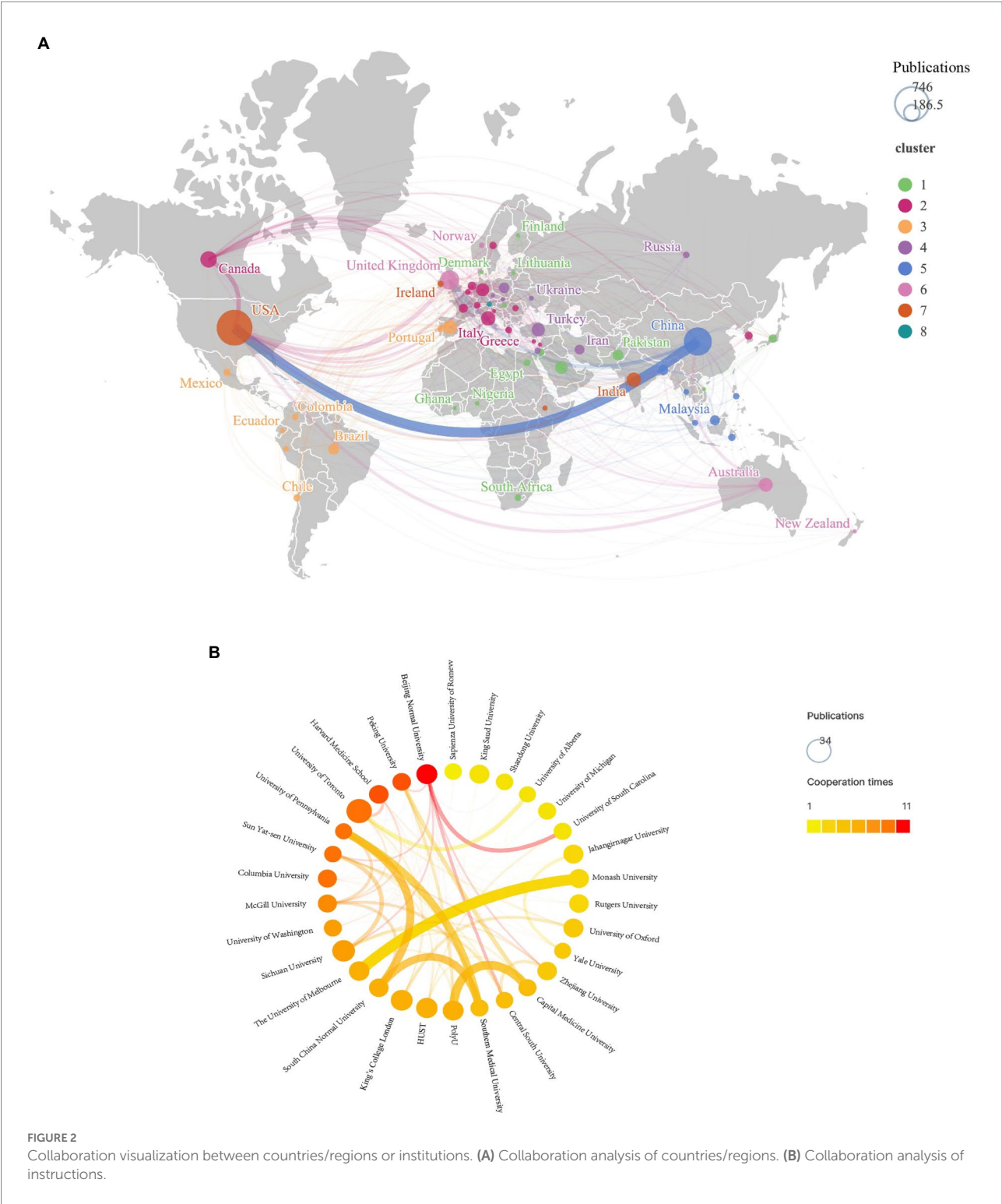
Journal analysis

In total, 927 journals have contributed to the 2,734 publications. According to Bradford's Law, 16 journals were considered core journals in the field (Table 5), among which INT J ENV RES PUB HE contained the largest number of publications (219 publications), followed by FRONT PSYCHOL (158 publications) and FRONT PSYCHIATRY (115 publications). Concerning Impact Factor (IF), INT J MENT HEALTH AD has the highest IF (IF = 11.555) in 2019, followed by PSYCHIAT RES (IF = 11.225) and J AFFECT DISORDERS (IF = 6.533). In the quartile category, 5 of the 16 journals were in Q1 (the top 25% of the IF distribution) of different areas, 9 of 16 were in Q2 and the rest were in Q3.

The most co-cited journal was also INT J ENV RES PUB HE (1,378 citations), followed by PSYCHIAT RES (1,302 citations) and PLOS ONE (1,136 citations). 12 of the 16 co-cited journals were in Q1 and the rest were in Q2. These results showed that INT J ENV RES PUB HE, FRONT PSYCHOL and PLOS ONE had significant contributions in this field due to the higher publications and citations.

Keyword analysis

Citespace was used for keyword contribution analysis and keyword burst analysis from 2019 to 2022. Figure 6A shows the

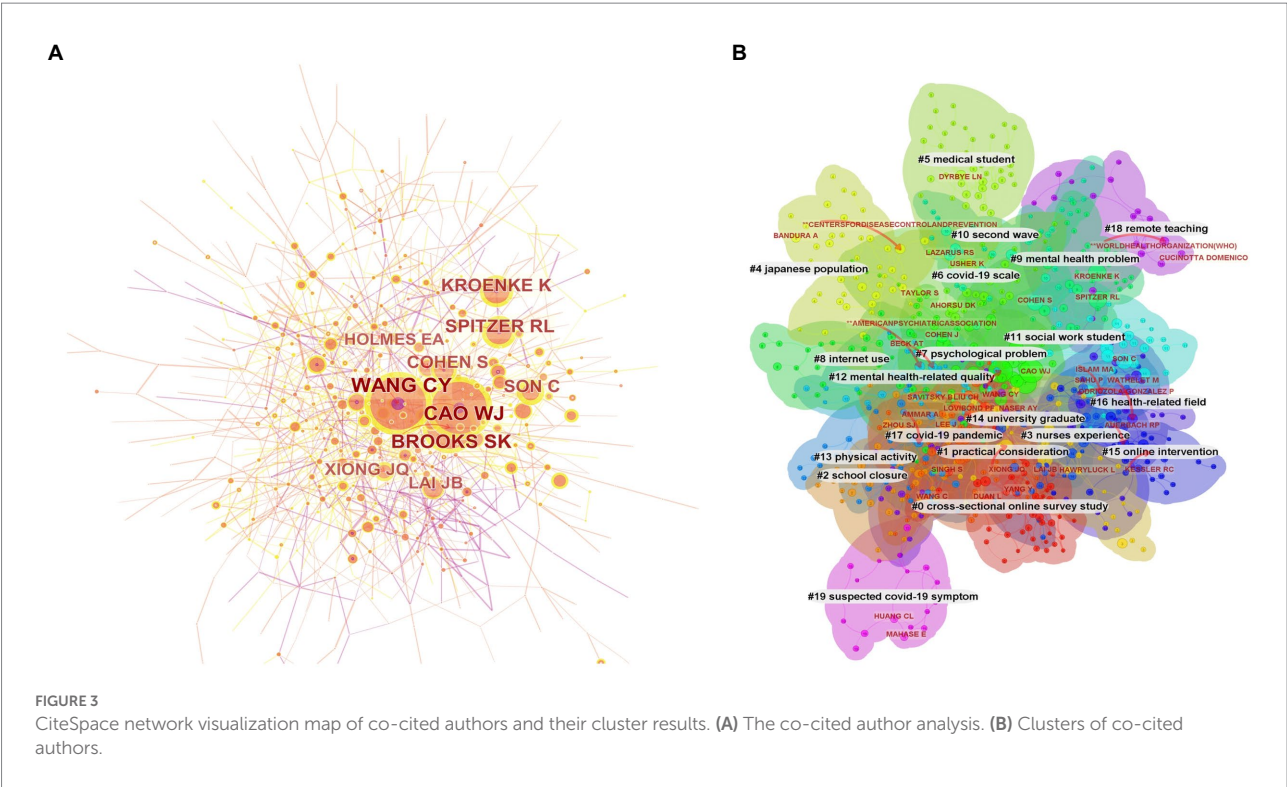


keyword contribution map from 2020 to 2022. Dual image overlay was used to investigate the keyword change trend in this field. The red line in [Figure 6A](#) represents the keyword contribution since 2022. No obvious change, however, was found. The main keywords were still about “mental health,” “anxiety,” and “stress.” Top 10 keywords with the strongest citation bursts were detected to

provide helpful insights to research hotspots in this field ([Figure 6B](#)). All the bursts occurred in 2021 and lasted to 2022, some of which were overlapped with the clusters of co-cited references. Under the circumstance of COVID-19, the mental health issues post-traumatic stress disorder, anxiety and self-esteem might influence students. It was worth investigating how

TABLE 3 Top 10 productive authors and co-cited authors related to ferroptosis.

Citing authors			Cited authors		
Rank	Author	Count	Rank	Co-cited author	Citation
1	Mark D Griffiths	10	1	Wang CY	565
2	Teris Cheung	9	2	Cao WJ	550
3	Yutao Xiang	9	3	Brooks SK	455
4	Jingbo Zhao	9	4	Spitzer RL	295
5	Mohammad Nurunnabi	8	5	Kroenke K	283
6	Kamilah Kamaludin	7	6	Son C	243
7	Heba Bakr Rhoshaim	7	7	Cohen S	221
8	Caruthan Chinna	7	8	Lai JB	190
9	Cezary Kusnierz	6	9	Xiong JQ	178
10	Fangbiao Tao	6	10	Holmes EA	171



COVID-19 influenced students’ behavior on internet and alcohol use.

Conceptual structure analysis

The thematic map developed from the keywords was shown in Figure 7A, to examine the thematic concepts in the research field of students’ mental health during COVID-19, based on the typology of topics. No themes were found in the upper right quadrant. The basic themes in the lower right quadrant included “depression,” “stress” and “anxiety,” indicating the high centrality

but low development degree. The upper left quadrant contained “health,” “students” and “life” with high development but low centrality, which was considered as niche themes. Themes such as “scale,” “validation” and “validity” were emerging or declining themes in the lower left quadrant, meaning the low centrality and low development degree. Moreover, there were also subjects located in the boundary between two quadrants. “Mental-health,” “impact” and “outbreak” were in the boundary between motor themes and niche themes, which had middle centrality and high development degree. Also, “adolescents,” “children” and “risk” were motor themes and basic themes simultaneously with high centrality and middle development degree.

Multiple correspondence analysis (MCA) was used to construct a conceptual structure of this field. All keywords were classified into two clusters in blue and red (Figure 7B). Connection between the two clusters was shown in Figure 7C. The words in blue were mainly associated with COVID-19 and its influence such as “outbreak,” “epidemic” and “psychological impact.” The red cluster represented words related to the influenced population and specific emotions including “students,” “adolescents,” “university students,” “anxiety” and “depression.”

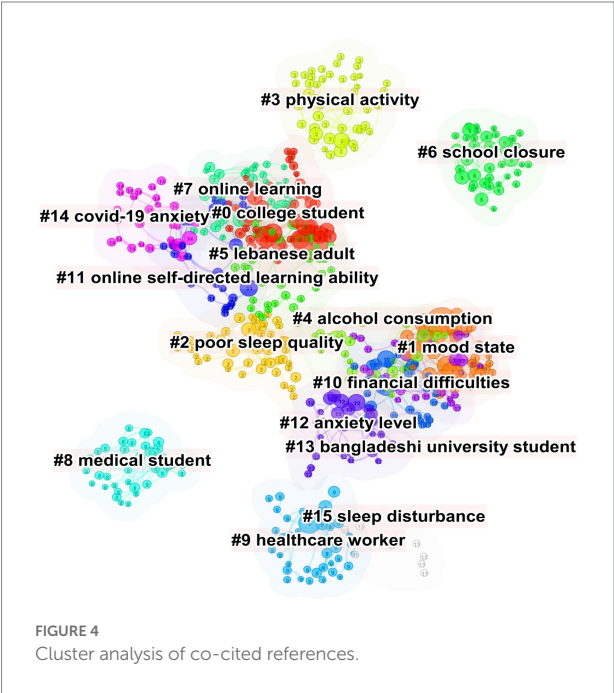


FIGURE 4
Cluster analysis of co-cited references.

Discussion

Overall, this paper aims to investigate the output of international scientific publications on student mental health during the COVID-19 pandemic from 2020 to 2022 through a bibliometric analysis, and to explore research trends and hotspots in the field. We found that China and the United States lead in the quantity and quality of publications in this field. Importantly, 16 journals are considered core journals in the field according to Bradford’s Law. In addition, the co-cited references indicated that the main psychological problems of students under the epidemic were closely related to anxiety, poor sleep and financial difficulties. Their behavior may also be affected, such as increased internet usage and alcohol consumption.

In 2020, with the rapid spread of the COVID-19 pandemic, more and more countries have entered partial or nationwide lockdowns (To et al., 2021). Under this circumstance, the mental health of students during the COVID-19 pandemic received greater attention, and a total of 439 relevant publications were published. Between 2021 and 2022, the number of publications in this field increased dramatically. The cumulative number of related articles published in 2021 was 2,202, and in 2022, the number has risen to 2,734. This dramatic increase may be related to the growing physical, psychological, and social challenges that students faced during the pandemic. In terms of the number of relevant publications, the United States and China have the most publications in this field; China and Canada have the highest average citations; the United States and China have the highest H-index. From this perspective, China and the United States are leading the way in this research area. In addition, country/regional cooperation analysis also indicates that China and the United States have established close cooperation in this field. In

TABLE 4 Major clusters of co-cited references.

Cluster ID	Size	Silhouette	Year Ave.	Label (LLR)
0	51	0.922	2019	College student
1	49	0.839	2019	Mood state
2	49	0.771	2020	Poor sleep quality
3	48	0.858	2019	Physical activity
4	44	0.868	2019	Alcohol consumption
5	43	0.797	2019	Lebanese adult
6	42	0.952	2019	School closure
7	41	0.815	2019	Online learning
8	41	0.908	2019	Medical student
9	38	0.862	2019	Healthcare worker
10	34	0.796	2019	Financial difficulties
11	32	0.875	2019	Online self-directed learning ability
12	31	0.874	2019	Anxiety level
13	31	0.847	2019	Bangladeshi university student
14	28	0.805	2019	COVID-19 anxiety
15	9	0.989	2019	Sleep disturbance

Clusters are referred in terms of the labels selected by log-likelihood ratio test method (LLR).

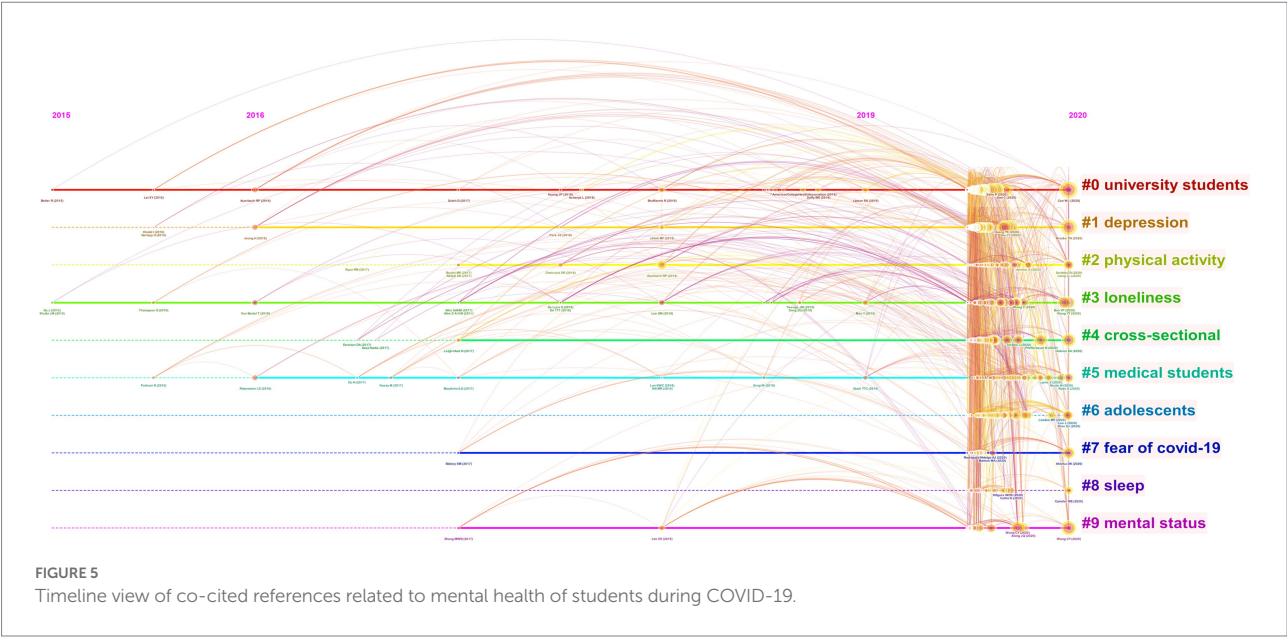


TABLE 5 Core journals related to the research of mental health of students during COVID-19.

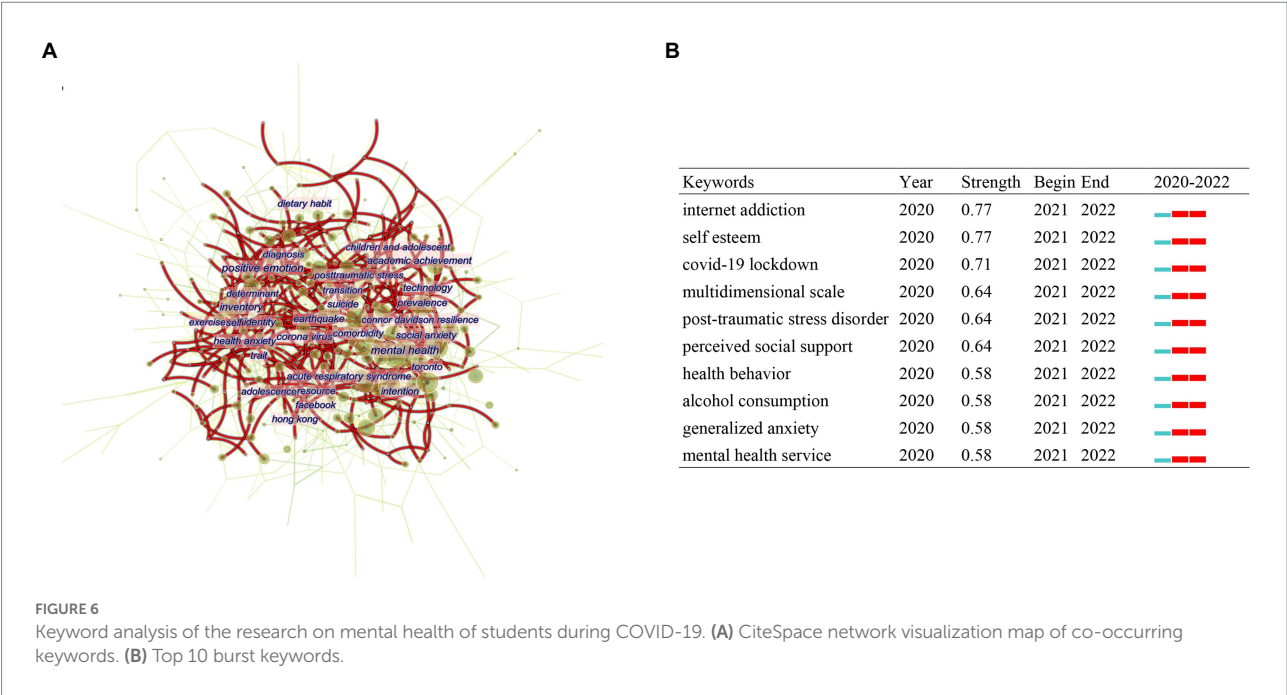
Citing journals					Cited journals				
Rank	Journal	Publications	IF	Quartile	Rank	Journal	Co-cited times	IF	Quartile
1	INT J ENV RES PUB HE	219	4.614	Q2	1	INT J ENV RES PUB HE	1,378	4.614	Q2
2	FRONT PSYCHOL	158	4.232	Q1	2	PSYCHIAT RES	1,302	11.225	Q1
3	FRONT PSYCHIATRY	115	5.435	Q2	3	PLOS ONE	1,136	3.752	Q2
4	JAM COLL HEALTH	58	2.395	Q3	4	J AFFECT DISORDERS	980	6.533	Q1
5	PLOS ONE	50	3.752	Q2	5	LANCET	973	202.731	Q1
6	FRONT PUBLIC HEALTH	48	6.461	Q1	6	FRONT PSYCHOL	861	4.232	Q1
7	J AFFECT DISORDERS	36	6.533	Q1	7	LANCET PSYCHIAT	712	7.35	Q1
8	CURR PSYCHOL	32	2.387	Q3	8	FRONT PSYCHIATRY	623	5.435	Q2
9	BMJ OPEN	29	3.017	Q2	9	J MED INTERNET RES	592	7.093	Q1
10	HEALTHCARE-BASEL	28	0.92	Q2	10	BRAIN BEHAV IMMUN	554	19.227	Q1
11	SUSTAINABILITY-BASEL	28	3.889	Q2	11	PSYCHOL MED	511	10.592	Q1
12	INT J MENT HEALTH AD	27	11.555	Q1	12	ASIAN J PSYCHIATR	495	13.89	Q1
13	BMC PSYCHIATRY	25	4.144	Q2	13	JAMA-JAM MED ASSOC	486	157.335	Q1
14	BMC MED EDUC	22	3.263	Q2	14	BMC PUBLIC HEALTH	456	4.135	Q2
15	HELIYON	22	3.776	Q2	15	JAMA NETW OPEN	424	13.366	Q1
16	PSYCHIAT RES	21	11.225	Q1	16	BMJ-BRIT MED J	408	93.467	Q1

the institutional collaboration analysis, the University of Toronto and Sichuan University have the most publications. In addition, Beijing Normal University, Peking University, and Harvard Medical School have the most cooperation occurrences, which shows that they are very active in inter-institutional collaborations in this field.

In the analysis of co-cited authors, Wang CY, Cao WJ, Brooks SK, Spitzer RL, Kroenke K, Son C, Cohen S, Lai JB, Xiong JQ and Holmes EA contributed the top 10 co-citations in turn. These researchers have made considerable contributions in this field.

According to the major clusters of co-cited references, we found that the hotspots of related research mainly focus on emotional status (mood state, anxiety level, COVID-19 anxiety), living habits (physical activity, alcohol consumption), learning status (school closure, online learning, online self-directed learning ability), sleep status (poor sleep quality, sleep disturbance), economic pressure (financial difficulties), and medical students (medical student, healthcare worker).

In Table 3, the most co-cited reference is “The psychological impact of the COVID-19 epidemic on college students in China”



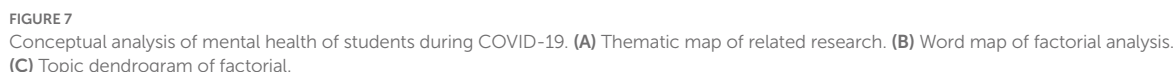
published by Cao WJ in 2020, and the co-citation count is 547. They conducted a sample analysis of college students at Changzhi Medical College and found that economic impact, daily life impact, and delay in academic activities were positively correlated with anxiety symptoms, while social support was negatively correlated with anxiety levels (Cao et al., 2020). At the same time, they made a recommendation for monitoring the mental health of college students during the COVID-19 pandemic. The second most co-cited reference was the review “The psychological impact of quarantine and how to reduce it: rapid review of the evidence.” Its authors, Brooks SK et al. published it in The Lancet in 2020. They examined the psychological effects of quarantine using three electronic databases. They examined the psychological effects of isolation using three electronic databases, most of which reported negative psychological effects (Brooks et al., 2020). Following these two papers is the article “Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China” by Wang et al. Their findings suggest that females, students and people with specific physical symptoms are more likely to be associated with greater psychological impact and higher levels of stress, anxiety and depression (Wang C. et al., 2020). It is worth mentioning that the three authors, Cao WJ, Brooks SK and Wang CY, are also among the top three authors in the co-citation analysis of authors.

Research on student mental health during the COVID-19 pandemic has gained welcome interest from many journal editors. Among them, the 10 journals in Table 4 contributed the most. The top journals by number of publications are INT J ENV RES PUB HE, FRONT PSYCHOL, and FRONT PSYCHIATRY.

We performed keyword analysis and burst keyword analysis using CiteSpace. In the Keyword Contribution Graph, many

keywords have made a large contribution, which shows that the research focus in this field is very diverse. In burst keyword analysis, the top 10 burst keywords in Figure 6 all appeared in 2021, and have maintained a high research interest in 2022. The keywords with the highest burst strength are internet addiction and self-esteem. Through a latent class analysis of Chinese schoolchildren, I-Hua Chen et al. found that during the recovery period of the COVID-19 outbreak, the two groups with a higher level of problematic internet use (PIU) had significantly higher levels of fear of COVID-19 than the one with a lower level (Chen et al., 2021). In addition, a study of residential college students during the COVID-19 lockdown found that the prevalence of PIU was high among residential college students during the COVID-19 lockdown, suggesting that the lockdown policy inevitably had an impact on their social lives. During such stressful events, they are vulnerable to PIU (Xia et al., 2021). Not only in China, many countries have reported the problem of internet addiction among students during the COVID-19 pandemic (Ismail et al., 2021; Kim et al., 2021; Shehata and Abdeldaim, 2021). This reminds us that we should pay attention to students’ Internet addiction during the COVID-19 pandemic. Numerous studies have shown that the negative impact of lockdowns caused by the COVID-19 pandemic on students with low self-esteem is significant. This means that schools and society should provide more support and attention to students with low self-esteem (Testoni et al., 2021; Vall-Roqué et al., 2021).

In addition to the research hotspots in this field, some neglected issues also need to be highlighted. First, suicide is the most serious consequence of negative mental health, but it has not been a research focus. Suicide and self-harm as a result of the direct and indirect effects of the COVID-19 pandemic is a major public health problem, and the social, psychological, and economic consequences of COVID-19 have the potential to



From the results of this study, we found that the sources and influencing factors of students' mental health problems during COVID-19 are extremely complex, which means that a single-approach solution may not be satisfactory. We need multidisciplinary support and multi-institutional collaboration.

In addition, based on our results, it is also important to focus on specific groups of students, including students with addictive habits, students suffering from illness, and students with financial difficulties. It is reasonable to assume that the COVID-19 pandemic exacerbated the plight of these students and significantly damage their mental health. In terms of addictive behaviors, numerous articles have demonstrated the strong relationship between social media addiction, alcohol addiction, drug addiction, smartphone addiction, and other addictive behaviors with COVID-19 and mental health (Sujarwoto et al., 2021; Hu et al., 2022; Lardier et al., 2022; Yehudai et al., 2022). In 2020, a study focusing on children with attention deficit hyperactivity disorder (ADHD) during the COVID-19 pandemic demonstrated that the

COVID-19 pandemic further exacerbated these children's mental health and worsened their behavioral problems (Zhang et al., 2020). There are not many similar studies, but research into the psychological profile of students with disease burden may be a future research trend in this field. There have been many reports on the financial stress of students (Jones et al., 2021; Myhr et al., 2021). However, due to the various factors that cause personal economic stress, such as family, society, region, economy, etc., this direction is still worthy of further research.

Our study has two major limitations. First, we only searched literature data on WoSCC and included only English literature, which may lead to selection bias. Second, authors with the same name and the diversity of keyword expressions can lead to bias.

Conclusion

Since the COVID-19 pandemic, the number of publications related to student mental health has increased dramatically. China and the United States lead in the quantity and quality of publications in this field. This article is the first bibliometric analysis based on the mental health of students during COVID-19 to study the research trends and hot spots. From the detected clusters of citations, students may continue to face difficulties with financial burden, anxiety and poor sleep. Student mental health may be influenced by other factors such as internet and alcohol dependence, which requires further exploration. With the assistance of information visualization, we identified the research hotspots in this field continue to revolve around emotional anxiety and unhealthy behaviors. Further study is needed to investigate the relationship of mental issue and various stressors and the interventions that effectively regulate psychological stress.

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 authors.

References

- Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., et al. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* 395, 912–920. doi: 10.1016/S0140-6736(20)30460-8
- Cao, W., Fang, Z., Hou, G., Han, M., Xu, X., Dong, J., et al. (2020). The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry Res.* 287:112934. doi: 10.1016/j.psychres.2020.112934
- Chaffee, B. W., Cheng, J., Couch, E. T., Hoeft, K. S., and Halpern-Felsher, B. (2021). Adolescents' substance use and physical activity before and during the COVID-19 pandemic. *JAMA Pediatr.* 175, 715–722. doi: 10.1001/jamapediatrics.2021.0541
- Chen, I. H., Chen, C.-Y., Liu, C.-H., Ahorsu, D. K., Griffiths, M. D., Chen, Y.-P., et al. (2021). Internet addiction and psychological distress among Chinese schoolchildren before and during the COVID-19 outbreak: a latent class analysis. *J. Behav. Addict.* 10, 731–746. doi: 10.1556/2006.2021.00052
- Chen, L., Ma, S., Hu, D., Lin, H., Zhu, Y., Chen, K., et al. (2020). Bibliometric study of sodium glucose cotransporter 2 inhibitors in cardiovascular research. *Front. Pharmacol.* 11:561494. doi: 10.3389/fphar.2020.561494
- Contreras-Barraza, N., Madrid-Casaca, H., Salazar-Sepúlveda, G., García-Gordillo, M. Á., Adsuar, J. C., and Vega-Muñoz, A. (2021). Bibliometric analysis of studies on coffee/caffeine and sport. *Nutrients* 13:3234. doi: 10.3390/nu13093234
- Copeland, W. E., McGinnis, E., Bai, Y., Adams, Z., Nardone, H., Devadanam, V., et al. (2021). Impact of COVID-19 pandemic on college student mental health and wellness. *J. Am. Acad. Child Adolesc. Psychiatry* 60, 134–141. doi: 10.1016/j.jaac.2020.08.466
- Darroudi, M., Gholami, M., Rezayi, M., and Khazaei, M. (2021). An overview and bibliometric analysis on the colorectal cancer therapy by magnetic functionalized nanoparticles for the responsive and targeted drug delivery. *J. Nanobiotechnol.* 19:399. doi: 10.1186/s12951-021-01150-6

Author contributions

QZ and MZ designed this study. QZ, MZ, YY, and QC reviewed and revised the manuscript. QC drafted the original manuscript. YY collected data, performed preliminary data analysis, and completed figures and tables. All authors approved the final manuscript submission and agreed to be responsible for all aspects of the work.

Funding

This study was supported by grants from the National Natural Science Foundation of China (81700658 and 82270795) and the Hunan Provincial Natural Science Foundation (2020JJ3058 and 2020JJ4864).

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.2022.985866/full#supplementary-material>

- Duffy, M. E., Twenge, J. M., and Joiner, T. E. (2019). Trends in mood and anxiety symptoms and suicide-related outcomes among U.S. undergraduates, 2007–2018: evidence from two National Surveys. *J. Adolesc. Health* 65, 590–598. doi: 10.1016/j.jadohealth.2019.04.033
- Fang, E., Liu, X., Li, M., Zhang, Z., Song, L., Zhu, B., et al. (2022). Advances in COVID-19 mRNA vaccine development. *Signal Transduct. Target. Ther.* 7:94. doi: 10.1038/s41392-022-00950-y
- Farhoudian, A., Radfar, S. R., Mohaddes Ardabili, H., Rafei, P., Ebrahimi, M., Khojasteh Zonoozi, A., et al. (2021). A global survey on changes in the supply, Price, and use of illicit drugs and alcohol, and related complications during the 2020 COVID-19 pandemic. *Front. Psych.* 12:646206. doi: 10.3389/fpsyg.2021.646206
- Farooq, S., Tunmore, J., Wajid Ali, M., and Ayub, M. (2021). Suicide, self-harm and suicidal ideation during COVID-19: a systematic review. *Psychiatry Res.* 306:114228. doi: 10.1016/j.psychres.2021.114228
- Fuse-Nagase, Y., Marutani, T., Tachikawa, H., Iwami, T., Yamamoto, Y., Moriyama, T., et al. (2021). Increase in suicide rates among undergraduate students in Japanese national universities during the COVID-19 pandemic. *Psychiatry Clin. Neurosci.* 75, 351–352. doi: 10.1111/pcn.13293
- Hou, T.-Y., Mao, X.-F., Dong, W., Cai, W.-P., and Deng, G.-H. (2020). Prevalence of and factors associated with mental health problems and suicidality among senior high school students in rural China during the COVID-19 outbreak. *Asian J. Psychiatr.* 54:102305. doi: 10.1016/j.ajp.2020.102305
- Hu, B., Guo, H., Zhou, P., and Shi, Z.-L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* 19, 141–154. doi: 10.1038/s41579-020-00459-7
- Hu, Q., Liu, Q., and Wang, Z. (2022). Meaning in life as a mediator between interpersonal alienation and smartphone addiction in the context of Covid-19: a three-wave longitudinal study. *Comput. Hum. Behav.* 127:107058. doi: 10.1016/j.chb.2021.107058
- Ismail, N., Tajjudin, A. I., Jaafar, H., Nik Jaafar, N. R., Baharudin, A., and Ibrahim, N. (2021). The relationship between internet addiction, internet gaming and anxiety among medical students in a Malaysian public university during COVID-19 pandemic. *Int. J. Environ. Res. Public Health* 18:11870. doi: 10.3390/ijerph182211870
- Jones, H. E., Manze, M., Ngo, V., Lamberson, P., and Freudenberg, N. (2021). The impact of the COVID-19 pandemic on college Students' health and financial stability in New York City: findings from a population-based sample of City University of New York (CUNY) students. *J. Urban Health* 98, 187–196. doi: 10.1007/s11524-020-00506-x
- Kim, D., Lee, J., and Nam, J. K. (2021). Latent profile of internet and internet game usage among south Korean adolescents during the COVID-19 pandemic. *Front. Psych.* 12:714301. doi: 10.3389/fpsyg.2021.714301
- Lardier, D. T., Zuhl, M. N., Holladay, K. R., Amorim, F. T., Heggenberger, R., and Coakley, K. E. (2022). A latent class analysis of mental health severity and alcohol consumption: Associations with COVID-19-related quarantining, isolation, suicidal ideations, and physical activity. *Int. J. Ment. Health Addict.* 13, 1–24. doi: 10.1007/s11469-021-00722-9
- Ma, Z., Zhao, J., Li, Y., Chen, D., Wang, T., Zhang, Z., et al. (2020). Mental health problems and correlates among 746 217 college students during the coronavirus disease 2019 outbreak in China. *Epidemiol. Psychiatr. Sci.* 29:e181. doi: 10.1017/S2045796020000931
- Marutani, T., Fuse-Nagase, Y., Tachikawa, H., Iwami, T., Yamamoto, Y., Moriyama, T., et al. (2021). Has COVID-19 affected suicides among graduate students in Japan? *Asian J. Psychiatr.* 65:102803. doi: 10.1016/j.ajp.2021.102803
- Myhr, A., Naper, L. R., Samarawickrema, I., and Vesterbekkmo, R. K. (2021). Impact of COVID-19 pandemic lockdown on mental well-being of Norwegian adolescents during the first wave-socioeconomic position and gender differences. *Front. Public Health* 9:717747. doi: 10.3389/fpubh.2021.717747
- Odono, A., Lugo, A., Amerio, A., Borroni, E., Bosetti, C., Carreras, G., et al. (2020). COVID-19 lockdown impact on lifestyle habits of Italian adults. *Acta Biomed* 91, 87–89. doi: 10.23750/abm.v91i9-S.10122
- Otsuka, H., and Anamizu, S. (2019). Japanese university students' difficulty in living and its association with suicidal ideation. *Asian J. Psychiatr.* 43, 50–52. doi: 10.1016/j.ajp.2019.05.004
- Pappa, S., Chen, J., Barnett, J., Chang, A., Dong, R. K., Xu, W., et al. (2022). A systematic review and meta-analysis of the mental health symptoms during the Covid-19 pandemic in Southeast Asia. *Psychiatry Clin. Neurosci.* 76, 41–50. doi: 10.1111/pcn.13306
- Peng, C., Kuang, L., Zhao, J., Ross, A. E., Wang, Z., and Ciolino, J. B. (2022). Bibliometric and visualized analysis of ocular drug delivery from 2001 to 2020. *J. Control. Release* 345, 625–645. doi: 10.1016/j.jconrel.2022.03.031
- Pompili, M., Serafini, G., Innamorati, M., Monteboni, F., Palermo, M., Campi, S., et al. (2012). Car accidents as a method of suicide: a comprehensive overview. *Forensic Sci. Int.* 223, 1–9. doi: 10.1016/j.forsciint.2012.04.012
- Qin, Z., Shi, L., Xue, Y., Lin, H., Zhang, J., Liang, P., et al. (2021). Prevalence and risk factors associated with self-reported psychological distress among children and adolescents during the COVID-19 pandemic in China. *JAMA Netw. Open* 4:e2035487. doi: 10.1001/jamanetworkopen.2020.35487
- Salim, S. (2021). The stress of the COVID-19 pandemic: beyond the data. *Curr. Neuropharmacol.* 19, 1161–1163. doi: 10.2174/1570159X19666210311103136
- Shehata, W. M., and Abdeldaim, D. E. (2021). Internet addiction among medical and non-medical students during COVID-19 pandemic, Tanta University, Egypt. *Environ. Sci. Pollut. Res. Int.* 28, 59945–59952. doi: 10.1007/s11356-021-14961-9
- Shen, J., Shen, H., Ke, L., Chen, J., Dang, X., Liu, B., et al. (2022). Knowledge mapping of immunotherapy for hepatocellular carcinoma: a bibliometric study. *Front. Immunol.* 13:815575. doi: 10.3389/fimmu.2022.815575
- Sujarwoto, J., Saputri, R. A. M., and Yumarni, T. (2021). Social media addiction and mental health among university students during the COVID-19 pandemic in Indonesia. *Int. J. Ment. Health Addict.* 1, 1–15. doi: 10.1007/s11469-021-00582-3
- Testoni, I., Palazzo, L., Ronconi, L., Rossi, G., Ferizoviku, J., and Morales, J. R. P. (2021). The experience of children with a parent suffering from amyotrophic lateral sclerosis during the COVID-19 pandemic. *Sci. Rep.* 11:16046. doi: 10.1038/s41598-021-95338-3
- To, K. K. W., Sridhar, S., Chiu, K. H.-Y., Hung, D. L.-L., Li, X., Hung, I. F.-N., et al. (2021). Lessons learned 1 year after SARS-CoV-2 emergence leading to COVID-19 pandemic. *Emerg. Microbes. Infect.* 10, 507–535. doi: 10.1080/22221751.2021.1898291
- Toprak, S., Cetin, I., Guven, T., Can, G., and Demircan, C. (2011). Self-harm, suicidal ideation and suicide attempts among college students. *Psychiatry Res.* 187, 140–144. doi: 10.1016/j.psychres.2010.09.009
- Vall-Roqué, H., Andrés, A., and Saldaña, C. (2021). The impact of COVID-19 lockdown on social network sites use, body image disturbances and self-esteem among adolescent and young women. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* 110:110293. doi: 10.1016/j.pnpbp.2021.110293
- Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., et al. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 323, 1061–1069. doi: 10.1001/jama.2020.1585
- Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., Ho, C. S., et al. (2020). Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *Int. J. Environ. Res. Public Health* 17:1729. doi: 10.3390/ijerph17051729
- Xia, Y., Fan, Y., Liu, T.-H., and Ma, Z. (2021). Problematic internet use among residential college students during the COVID-19 lockdown: a social network analysis approach. *J. Behav. Addict.* 10, 253–262. doi: 10.1556/2006.2021.00028
- Yehudai, M., Bender, S., Gritsenko, V., Konstantinov, V., Reznik, A., and Isralowitz, R. (2022). COVID-19 fear, mental health, and substance misuse conditions among university social work students in Israel and Russia. *Int. J. Ment. Health Addict.* 20, 316–323. doi: 10.1007/s11469-020-00360-7
- Zhang, S. X., and Chen, J. (2021). Scientific evidence on mental health in key regions under the COVID-19 pandemic – meta-analytical evidence from Africa, Asia, China, Eastern Europe, Latin America, South Asia, Southeast Asia, and Spain. *Eur. J. Psychotraumatol.* 12:2001192. doi: 10.1080/2008198.2021.2001192
- Zhang, S. X., Chen, R. Z., Xu, W., Yin, A., Dong, R. K., Chen, B. Z., et al. (2022). A systematic review and meta-analysis of symptoms of anxiety, depression, and insomnia in Spain in the COVID-19 crisis. *Int. J. Environ. Res. Public Health* 19:1018. doi: 10.3390/ijerph19021018
- Zhang, J., Liu, Y., and Sun, L. (2017). Psychological strain and suicidal ideation: a comparison between Chinese and US college students. *Psychiatry Res.* 255, 256–262. doi: 10.1016/j.psychres.2017.05.046
- Zhang, J., Shuai, L., Yu, H., Wang, Z., Qiu, M., Lu, L., et al. (2020). Acute stress, behavioural symptoms and mood states among school-age children with attention-deficit/hyperactive disorder during the COVID-19 outbreak. *Asian J. Psychiatr.* 51:102077. doi: 10.1016/j.ajp.2020.102077



OPEN ACCESS

EDITED BY

Jesús de la Fuente,
University of Navarra, Spain

REVIEWED BY

Xiaozi Gao,
The Education University of Hong Kong,
Hong Kong SAR, China
Patricia Ortega-Andeane,
National Autonomous University of Mexico,
Mexico

*CORRESPONDENCE

Lidia Infante-Cañete
lidiainfante@uma.es

SPECIALTY SECTION

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

RECEIVED 31 March 2022

ACCEPTED 14 November 2022

PUBLISHED 07 December 2022

CITATION

Infante-Cañete L, Arias-Calero L,
Wallace-Ruiz A, Sánchez-Sánchez AM and
Muñoz-Sánchez Á (2022) One more step in
the study of children's daily stress: The
spillover effect as the transfer of tension in
family and school environments.
Front. Psychol. 13:909928.
doi: 10.3389/fpsyg.2022.909928

COPYRIGHT

© 2022 Infante-Cañete, Arias-Calero,
Wallace-Ruiz, Sánchez-Sánchez and
Muñoz-Sánchez. 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.

One more step in the study of children's daily stress: The spillover effect as the transfer of tension in family and school environments

Lidia Infante-Cañete^{1*}, Lidia Arias-Calero¹, Agustín Wallace-Ruiz²,
Ana María Sánchez-Sánchez¹ and Ángela Muñoz-Sánchez¹

¹Department of Developmental Psychology, University of Málaga, Málaga, Spain, ²Department of Psychobiology and Methodology of Behavioral Sciences, University of Málaga, Málaga, Spain

Introduction: The spillover effect is the psychological overflow due to daily stress in one context and the transfer of its consequences to another close environment. The aim is to explore the spillover effect in conflicts within the family, on the one hand, and school with peers on the other hand, as an inferred measure of daily stress according to the literature.

Method: The study consisted of a sample of 208 6-year-old students and their families. A methodology based on daily report records was used, by means of two ad hoc checklists with simultaneous measurements, for 2 consecutive weeks and 3 academic years, for both family and school contexts. A repeated measures design, together with a nonparametric statistical data analysis with Friedman's test and contrast measures, was used.

Results: Daily stress shows significant differences in the family setting throughout the week ($\chi^2 = 32.44$; $p = 0.000$) and at different times of the day ($\chi^2 = 29.65$; $p = 0.000$). In the school setting, differences were found across the different days of the week ($\chi^2 = 36.96$; $p = 0.000$). Spillover effect has been discovered between conflicts at home in the evening and conflicts at school. At the same time, conflicts at school are related to conflicts at home from Wednesday onward.

Discussion: The results suggest further research on daily stress through the interrelation of the different contexts, as well as the impact that moments of conflict may have on the psychological and emotional development of the child.

KEYWORDS

spillover, daily stress, daily record, family, peers

Introduction

The concept of spillover comes from areas of knowledge such as economics, engineering, physics, or ecology. It refers to those cases when something happening in one context directly affects a different context (Piotrkowski, 1979; Krpan et al., 2019). This very physical concept has been adapted to psychology (Crouter, 1984) and can be understood

as the transfer of mood from one context to another due to excessive situational stress (Staines, 1980). More recently, Pakman (2006) defines the spillover effect as the psychological spillover that takes place when, after a temporary process of tension accumulation, it reaches such an intensity that the equilibrium is broken and the homeostatic mechanisms are saturated, experiencing an emotional overflow because of this cumulative process. The study of the spillover effect is therefore placed within the framework of stress and, more specifically, daily stress. Daily stress refers to the reaction—physical, psychological, and emotional—produced by a series of frustrating and irritating demands which occur daily when interacting with our environment (Kanner et al., 1981). Therefore, it is related to small events, problems, worries, and both high and low frequency setbacks, which are quite predictable and destabilize the physical-emotional wellbeing of the person (Seiffge-Krenke, 2007). The spillover effect implies an addition to the daily stress concept, as it brings in the export of those tensions produced by continuous events in one context to a totally different context, which can manifest themselves at the behavioral, psychological, or emotional level.

Different studies have addressed this transfer of tensions between interrelated contexts, as well as between those agents involved in the contexts. Some of these have focused on the work environment and how work demands affect family relationships (Crouter, 1984; Gerard et al., 2006; Flook and Fuligni, 2008), the way in which the quality of the couple's interaction influences their children (Buckhalt et al., 2009; Nelson et al., 2009; Chung et al., 2011; Kouros et al., 2014), and how the household demands influence work outcomes (Crompton and Lyonette, 2006). Collectively, all these studies confirm the existence of the spillover effect. They also state that contagions between contexts or different agents in the same environment are interrelated, but they can be understood as independent concepts, and should be measured separately (Jacobs and Gerson, 2004; Hill, 2005; Ferrarini, 2006). In this sense, the ecological-systemic perspective (Minuchin, 1982; Bronfenbrenner, 2002) considers that the family microsystem is interdependent on a different microsystem—the school. In turn, both depend on the macrosystem where they are located, which is the residential area or neighborhood. According to Bronfenbrenner (2002), child development is subject to the influence of interconnected, network-like environments.

Both the family context through family member interaction and the school context through peer interaction are subject to multiple situations, of which a high percentage can generate daily stress. Children are not exempt from experiencing such daily stress, as well as the impact on their development and wellbeing (Trianes, 2002; Trianes and Escobar, 2009). Thus, different studies present the family and the school as the main stress-producing contexts (Phelan et al., 1994; De Anda et al., 2000). At the same time, they are essential for child development (Kaufman et al., 2020). In children, social relationships between peers and the conflicts derived from these relationships stand out as typical stressors. More specifically, Oros and Vogel (2005) point to

fighting, being a bully, maintaining dominant and bossy behavior, and being bullied as behaviors related to this type of stressor. Being or feeling isolated from other schoolchildren is another stressor that has also been observed in this population (Salmivalli et al., 2000; Ortega and Monks, 2005). Several authors show that most school conflicts occur during recess, as it is a free environment to interact and relate with peers (Olweus, 1991; Rivers and Smith, 1994; Félix et al., 2008; Filella et al., 2016). Regarding the family environment, tensions within the family members are a negative factor for children's development (Franz and Gross, 2001; El-Sheikh et al., 2006; Sandín, 2008; Buckhalt et al., 2009; Kouros et al., 2014). In particular, at early ages, they can become recipient subjects of parental tension spillover (Greenberger et al., 1994), but they can also be generators by hindering performance at work (Grzywacz et al., 2002; Grzywacz and Carlson, 2007) or parental couple life (Lawrence et al., 2010; Nelson et al., 2014). Consequently, the family's daily routines and practices related to caregiving tasks such as feeding and dressing the kids, brushing their teeth, showering them, potty training them, or asking them to tidy up their things—usually observable activities repeated over time, which respond to physiological needs and are hardly modifiable (Wolin and Bennet, 1984)—would theoretically be linked to promoting the children's wellbeing and health, but may entail both a protective and a destructive value (Palacios and Andrade, 2008; Fiese and Winter, 2010), since they are not exempt from generating conflicts and tensions within the couple or between the couple and the children (Sears et al., 2016). Therefore, dyadic as well as triadic interactions are relevant (Stroud et al., 2011).

Currently, there are barely any studies on the spillover effect at 6 to 8 years, since the literature reviewed in the population under 18 years of age focused either on early childhood or on adolescence and pre-adolescence (Salamon et al., 2011; Stroud et al., 2011; Martin et al., 2017; Davies et al., 2018; Pu and Rodríguez, 2021; Mastrotheodoros et al., 2022). This age range is crucial because from the age of 6, children have to face important changes and events such as the transition from Early Childhood Education to Primary School (Rimm-Kaufman and Pianta, 2000; Schulting et al., 2005; Bulkeley and Fabian, 2006; Fabian and Dunlop, 2006; Dockett and Perry, 2007). A higher degree of autonomy and responsibility is demanded from them by two institutions which are not mutually consistent in their requirements (Oliva and Palacios, 2000; Bronfenbrenner, 2002; Enns et al., 2018), in addition to greater academic demands mediated by stress and anxiety (Ramírez et al., 2016) while they begin to be aware of their emotions and it is a prime age to work on emotional management, due to its impact on learning and academic success (Graham et al., 1984; Palacios and Hidalgo, 1999; Ortiz, 2001; Herrera et al., 2004; Soldevila et al., 2007; Denham et al., 2016; Enns et al., 2018). According to Kinkead-Clark (2015), an essential element with an impact on academic transitions is the quality of educational and parenting practices in the years prior to Primary Education. Moreover, during these early years of Primary Education, a strong and greater socialization

among peers is expected, and the conflict resolution strategies used in Early Childhood Education give way to more complex strategies. Therefore, the first cycle of Primary Education corresponds to the age interval where more conflicts occur (Soldevila et al., 2007; Sidorowicz et al., 2009; Filella et al., 2016).

Having made this point, it is worth tracing how conflicts appearing in the school context when interacting with peers produce a transfer to the family context (and vice versa) in children (Lehman and Repetti, 2007; Nelson et al., 2009; Chung et al., 2011; Du et al., 2018), since there is still not much knowledge about the way this mutual relationship works (Kaufman et al., 2020). A precedent for this can be found in Flook and Fuligni (2008), who examined this spillover effect in adolescents, reporting results that showed a bidirectional influence between daily family spillover and conflictive school relationships. In particular, difficulties in school social relations seem to be closely related to family difficulties, as well as to the child's school adjustment (Repetti et al., 2002). Thus, several studies have shown how behavioral problems in childhood predict parental stress which, in turn, predicts the appearance of behavioral problems in children. It shows that, if the effects are bidirectional, it is possible to reduce behavioral problems in case a decrease in parental stress is achieved (Baker et al., 2005; Crnic et al., 2005; Mackler et al., 2015). The quality of the relationship with peers is also important in relation to the psychological adjustment of children, especially when it is not satisfactory and/or they suffer rejection (Trianes, 2002).

In sum, this study starts from the need to analyze the spillover effect between the family context (defined through the conflictive interactions occurring in the daily parenting activities that parents have with their children) and the peer context through the conflictive interactions that children have with their classroom peers, through the daily report. The daily report methodology allows recording—on a daily basis and in detail—how short-term tensions are transmitted between the parenting process carried out by parents and their children's relationships with their peers (Almeida, 2004; Chung et al., 2011). Daily reports or records are useful when the aim is to catch or capture, whatever the context, the set of micro-processes of tensional spillover at the particular moment in which they take place. Thus, it is possible to analyze the appearance of daily tensions in the family system and the consequent effects on the relationship between children and their peers. On that account, the daily record allows detecting specific episodes of tension both at home and at school, in a short interval of time, from the moment they take place in a given context. This methodology has already been used in recent studies on the spillover effect between parents and adolescents and preadolescents, showing satisfactory results. The daily report has been used in studies such as those of Lehman and Repetti (2007), who evaluated for 5 consecutive days the events that took place at school and at home, or in the work of Flook and Fuligni (2008), in which the evaluation was conducted with adolescents during 2 consecutive weeks. Sherrill et al. (2017) is also another example in the use of daily records. All of these studies report a basal

functioning or behavioral pattern where behavioral differences respond to an objective, non-arbitrary fact across contexts (Lehman and Repetti, 2007; Flook and Fuligni, 2008; Sherrill et al., 2017).

This study has the following objectives:

1. To identify the existence of a possible pattern of conflict behaviors in each context throughout the week. Studies focusing on the spillover effect report a basal functioning in both preadolescents (Sherrill et al., 2017) and adolescents (Lehman and Repetti, 2007; Flook and Fuligni, 2008), where conflict behaviors are repeated and show some kind of interrelation. The aim is to capture the existence of this basal rhythm, as objective and non-arbitrary data, in a smaller population than that studied so far in the literature. In this respect, the demands and requirements that the 6–8 year-old children must face are the change from Early Childhood Education to Primary School, which requires a greater degree of autonomy and responsibility, more academic demands, and a deeper awareness of their emotions, as mentioned above. Therefore, from 6 to 8 year, the child is in the same educational cycle and developmental stage (Ballesteros and García, 2001; Coll et al., 2014) that may provide certain stability in the different contexts, so that it may favor the attainment of the behavioral pattern that supposes that basal rhythm (Gerard et al., 2006).
2. To estimate the prevalence of both parenting conflicts within the family context and peer conflicts within the school context, as measures of context-specific tensions. Studies such as those of Palacios and Andrade (2008); Trianes et al. (2009); Filella et al. (2016) on the school setting, as well as those of Stroud et al. (2011) or Sears et al. (2016) on the family setting, point to behaviors related to stressors. In this sense, finding comparable conflicts in the scientific literature means that our conclusions can be generalizable and replicable, since the sample could not be considered specific or biased by any particular aspect.
3. To analyze the bidirectional relationship between parental conflicts on the subject of children upbringing, and conflicts involving these children and their peers at school, for 3 years. Studies with populations larger than the study sample have discovered a spillover effect in different moments (Lehman and Repetti, 2007; Flook and Fuligni, 2008; Chung et al., 2011; Bai et al., 2017; Sherrill et al., 2017), and it is expected to be able to specify the existing relationship between conflicts in different contexts.

In relation to hypotheses, the following is expected:

Hypothesis 1: Parenting conflicts occurring early in the morning will negatively affect children throughout the day at three points: in the development of interactions with peers at school, in the development of parenting activities in the

afternoon, and in the development of parenting activities in the evening hours.

Hypothesis 2: Conflict situations occurring in the family home follow a characteristic pattern throughout the week, which is systematic in each year evaluated.

Hypothesis 3: Prevalence data on school and family conflicts will be compatible with those found in the scientific literature.

Hypothesis 4: The interaction conflicts between peers in the school context will have a negative effect on daily parenting tasks twice a day: in the afternoon and in the evening.

Hypothesis 5: The spillover effect between family and school conflict situations can be bidirectional.

Materials and methods

Participants

A total of 208 students (104 boys and 104 girls) participated, together with their respective parents. The starting age for schoolchildren is 6 years old. The sample came from four first-year primary school classes from two public schools in the province of Malaga. These were located in a medium-high socioeconomic level area.

Instruments

Daily report on interpersonal conflicts with peers in the school context

It is an *ad hoc* self-report assessing the micro-processes of interpersonal conflicts as they occur at school (see [Appendix 1](#)). Its content is extracted from empirical studies and current theoretical reviews related to the area of interpersonal conflicts in childhood ([Salmivalli et al., 2000](#); [Oros and Vogel, 2005](#); [Ortega and Monks, 2005](#)). It consists of 13 items in a Yes/No response format. In addition, the positive answer is graduated according to frequency: once, 2–4 times, 5 or more times. This report yields two scores—the first one is the total score of the report and ranges from 0 to 13 points, whereas the second one is the number of conflicts the student has been involved in that morning. The second score is the frequency of conflictive events the student has been caught up in during the morning, and ranges from 0 to 39 points. In this study, the total score has been used.

Daily report on parenting conflicts between parents and their children

This is an instrument which has been developed *ad hoc*, in order to assess the micro-processes of interpersonal conflicts

when they occur in the daily routine between parents and children. Specific and updated literature related to stressful family interpersonal conflicts was consulted for its elaboration ([Wolin and Bennet, 1984](#); [Almeida, 2004](#); [Palacios and Andrade, 2008](#); [Fiese and Winter, 2010](#); [Chung et al., 2011](#); [Stroud et al., 2011](#); [Sears et al., 2016](#); [Sherrill et al., 2017](#)); see [Appendix 2](#). For its design, specific and updated literature related to the area of stressful family interpersonal conflicts was consulted. The daily report focuses on the daily tasks. In each one, it is evaluated whether each child performs the main task of that moment without posing any conflict or, on the contrary, presents resistance behaviors, such as being sad, angry, or uncooperative, demanding attention, throwing tantrums, being aggressive, or disobeying. This report yields three scores that refer to morning parenting conflicts (from 0 to 28 points), afternoon parenting conflicts (from 0 to 27 points), and evening parenting conflicts (from 0 to 18). The score is the number of parenting conflicts where the child has been involved. The report also yields a total score ranging from 0 to 73 points—this score corresponds to the number of parenting conflicts throughout the day. It is important to note that, in these scores, items that had a positive wording were excluded, as they did not show conflict. The items referring to “Other problem” offered the families the possibility of describing another situation that differed from those considered.

Procedure

After the selection of the educational centers that showed interest in participating in the research project “*Spillover between parent–child and child–peer. Transmission of bidirectional tensions between parent–child (6–8 years old) and child–peer interactions*,” with project code SEJ2011-7225, the pertinent permissions were requested from the authorities of the Education Department in the province of Malaga, the school board, and the teaching staff of each of the participating centers. Informative meetings were held with both teachers and families with the aim of informing them about the aims of the research. Besides, the signing of a collaboration agreement was asked—specific aspects such as space, timetable, and procedure were specified there. At the same time, the research team was committed to process the data confidentially, make them available to the center’s guidance department and the parents, contribute to the training of parents, and collaborate in other initiatives undertaken by the center during the development of the project. Prior to this study, a pilot study was carried out to test the instruments and data collection.

The *Daily reports on interpersonal conflicts with peers in the school context* were completed individually by means of personal interviews from Monday to Friday during school hours, in the time slot after recess. The instruments were administered by four researchers who told the children about the voluntary nature of participation.

The *Daily Parenting Conflict Reports* were completed individually by the primary caregiver during the last hour of the

day in the same data collection period for their children. The report was collected the morning after its completion in a sealed envelope by the tutor responsible for the class under research.

This procedure was approved by the Ethics Committee of the University of Malaga, within the Excellence Project SEJ-7226, funded by the Junta de Andalucía.

Data analysis

Based on a longitudinal design, the descriptive data of all the study variables and the normality check of their distribution are presented in the first place. Secondly, in order to test the spillover effect, different statistical analyses were performed—Spearman's correlation and Friedman's test. The statistical software package for Social Science (v25) was used for the sake of data analysis.

Results

The analysis of descriptive statistics shows that the selected variables do not follow a normal distribution, which implies the use of non-parametric statistical tests. Previously, a Mann–Whitney *U* test showed that there were no differences between sexes (for morning, afternoon, and evening respectively, $U = 3807.5$, $p = 0.156$; $U = 3732.5$, $p = 0.361$; $U = 3734.0$, $p = 0.126$) or between schools (for morning, afternoon, and evening, respectively, $U = 3461.5$, $p = 0.117$; $U = 3376.5$, $p = 0.102$; $U = 3418.0$, $p = 0.108$). The descriptive statistics of the data (see Table 1) show that the number of conflicts collected during the three consecutive academic years of the research was not very high.

It was also explored which conflict situations were most likely to occur in each context by means of a statistical analysis of frequencies (Table 2). The most common morning conflicts in the family environment are closely related to each other—these fundamentally correspond to four items associated with the refusal to do chores (“S/he does not want to get up,” “S/he does not cooperate with getting washed up,” “S/he does not cooperate when getting dressed” and “S/he does not eat breakfast”). In the afternoon conflicts, something similar happens—there are four items with a higher percentage of occurrence which are associated with aggressive behaviors, expression of negative emotions, and denial of the demand (“S/he appears angry about something,” “S/he disobeys in a particular situation,” “S/he fights with friends or siblings,” “S/he does not want to take care of his/

her personal hygiene”). Finally, in relation to evening conflicts, the same trend continues, with the most prominent items being those related to the expression of sadness and lack of collaboration in the task (“S/he appears sad at dinner,” “S/he appears sad at bedtime,” “S/he does not cooperate at dinner” and “S/he does not cooperate when going to bed”). This table includes those potentially conflictive situations whose percentage exceeds 5%.

On the other hand, the most common conflicts in the school environment also turned out to be mainly four. These situations account for more than 50% of the total (see Table 3), and can be potentially frustrating, since most of them (except for the item “You have been picked on and/or insulted”) are not based so much on direct aggression, but rather on the perception of a conflict in the face of a refusal or reproach from others (“You could not do what you wanted because others decided to do something else,” “They told you to calm down and be quiet,” “You could not borrow what you wanted”).

After having carried out this step, it was tested whether there parenting conflicts between parents and children throughout the day (morning, afternoon, and evening) during all days of the week, in the 3 years of the study. Statistically significant correlations were found between morning-afternoon [$r = 0.434^{**}$; $p < 0.001$], morning-evening [$r = 0.392^{**}$; $p < 0.001$], and afternoon-evening [$r = 0.416^{**}$; $p < 0.001$]. These correlations were the strongest between the times of day closest to each other.

Friedman's test was performed to check if there were statistically significant differences in the number of morning, afternoon, and evening conflicts. Upon finding statistically significant differences, several Wilcoxon *t*-tests with Bonferroni correction were carried out to see between which groups there were differences. The findings show that more conflicts happened in the morning than in the evening, and more conflicts occurred in the afternoon than in the evening. The results of the effect size calculation indicate that the effect size is greater between morning and evening conflicts (Table 4).

It was also tested whether the amounts of conflicts during the days of the week were similar or differed from each other. Using the same statistical tests as in the previous analysis, it was found that Monday is the day with the highest amount of morning conflicts with respect to the rest of the days of the week (Table 5). The results of the effect size calculation indicate that the effect size increases throughout the week.

During the afternoons, there are more conflicts on Mondays than on Thursdays and Fridays, and there are more conflicts on Tuesdays and Wednesdays than on Fridays. In this case, the effect size is very similar, although the same trend as in the previous analysis could be observed (Table 6).

During the evenings, there were no differences in the number of conflicts throughout the week [$\chi^2 = 7.001$; $p > 0.05$].

Regarding interpersonal conflicts with peers in the school context, it was checked whether they changed throughout the week, finding that there were more conflicts on Mondays than on Thursdays and Fridays, and that the number of conflicts was

TABLE 1 Descriptive statistics.

	Average	Median	Mode	Standard deviation
Morning	1.418	0.6	0	2.002
Afternoon	1.440	1	0	1.762
Evening	0.944	0	0	1.440

TABLE 2 Frequency distribution of items in the family conflict daily record.

Morning conflicts item	Percentage	Cumulative percentage
S/he appears cranky and demanding	6.56	40.91
S/he does not cooperate at breakfast	7.63	48.54
S/he does not eat breakfast	7.70	56.24
S/he does not want to get up	13.40	69.64
S/he does not cooperate in getting washed up	13.40	83.04
S/he does not cooperate when getting dressed	16.96	100
Afternoon Conflicts Item	Percentage	Cumulative percentage
S/he throws tantrums over food	6.70	31.33
S/he appears cranky and demanding	9.81	41.14
S/he appears angry about something	10.08	51.22
S/he does not want to take care of their personal hygiene	10.76	61.98
S/he disobeys in a particular situation	18.27	80.24
S/he fights with friends and/or siblings	19.76	100
Evening conflicts item	Percentage	Cumulative percentage
Another problem	5.35	20.93
S/he appears angry because s/he does not want to go to sleep	5.69	26.62
S/he appears cranky and demanding	9.44	36.06
S/he throws tantrums in order not to go to sleep	9.67	45.73
S/he does not want to eat their meal	10.35	56.09
S/he appears sad about something	11.26	67.35
S/he eats their meal without a fuss	14.68	82.03
S/he appears sad	17.97	100

higher on Tuesdays, Wednesdays, and Thursdays than on Fridays. The effect size is larger the farther apart the days of the week are, i.e., results like the previous analyses are obtained (Table 7).

TABLE 3 Frequency distribution of items in the daily record of school conflicts.

School conflicts item	Percentage	Cumulative percentage
Have you had a fight with another child?	3.17	3.17
You have been accused for no reason	5.07	8.24
Someone did not want to be your friend	5.23	13.47
Have you ever had a confrontation with another child without getting into a fight?	5.67	19.14
You have imposed your will on other children	5.96	25.10
You have been attacked by another child (without having started the fight yourself)	6.44	31.54
You have been pushed (stepped on, any covert aggression)	6.98	38.51
You have not been allowed to play	7.36	45.87
You have gotten scolded for no reason	7.75	53.62
You have been told to calm down, to stay still	8.88	62.50
You have been unable to borrow something you wanted	8.93	71.43
They have picked on you (you have been insulted)	13.33	84.77
You have not been able to do what you wanted to do because other children have decided to do something else	15.23	100

Through a previous Spearman correlation analysis, it was found that evening conflicts were related both to morning conflicts the following day [$r=0.403^{**}$; $p<0.001$], and to interpersonal conflicts with peers in the school setting [$r=0.175^{*}$; $p<0.05$].

Also, it was checked whether there was a relationship between interpersonal conflicts at school and conflicts at home in the afternoon and evening during the weekdays, in the 3 years of the study. No relationship was found between school conflicts and family and interpersonal conflicts in the afternoon, but significant correlations were observed from Wednesday onward between school and family conflicts in the evening hours (Table 8).

TABLE 4 Results of Friedman's test and multiple *a posteriori* contrasts of conflicts at home.

Comparison mornings: Afternoons—evenings

Friedman	χ^2	<i>p</i>				
MC-AC-EC	29.65	0.000	**			
T Wilcoxon	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>T</i>	<i>p</i>
Morning conflict–afternoon conflict	6.935	8.445	6.372	6.879	5442.50	0.783
Morning conflict–evening conflict	6.935	8.445	4.616	5.998	3337.50	0.000**
Afternoon conflict–evening conflict	6.372	6.879	4.616	5.998	3209.50	0.000**

p* < 0.05; *p* < 0.01.

Discussion

The use of the concept of spillover linked to the field of psychology is fairly recent (Pakman, 2006) and is connected to the concept of daily stress (Seiffge-Krenke, 2007). The term “spillover” completes the effect that daily stress has on people's lives insofar as it contemplates the transfer of its negative effect to another context, so that one area of life can impact another one. This results in an amplification of responses that may be conflicting, for example, between the family and school context (Repetti and Wang, 2017).

This study delves into the idea of how the conflicts occurring in one context, family or school, are related to the daily difficulties experienced in the other context. In this sense, this study aims to detect the pattern of the spillover effect both on a single day and throughout the week, as well as to analyze the bidirectional relationship between the conflicts experienced by parents in carrying out parenting activities with their children, and conflicts experienced by these children and their peers at school, over a three-year study period. In addition, the most prevalent conflicts in each of the contexts evaluated in the selected sample have been described. Based upon a repeated measures design, nonparametric tests have been used to demonstrate different relationships between the data.

The study findings indicate—in relation to the first hypotheses—that a repetitive pattern is observed throughout the 3 years of the study, with Monday being the day with the highest level of conflict in both contexts. In turn, Monday's potential for conflict also shows its own pattern, with the first moments of the day registering the highest incidence. Friday is the day with the lowest number of conflicts (especially in the afternoon), so the likelihood of conflict decreases as the week progresses and the weekend approaches. Finally, it should be noted that, in general, mornings are more eventful than evenings. Thus, the hypothesis is corroborated, and the first objective of the research is fulfilled—this is very important, since there are no references in the scientific literature about the pattern followed by the family-school spillover effect, so this finding is of great relevance so as to know the specific moments and days which are more prone to spillover.

In relation to the second hypothesis, regarding the prevalence of problematic situations in each environment, although there is not a large number of conflicts in the study

sample, most of those happening in the family context have to do with refusal behaviors on the part of the child to perform tasks, as well as with behaviors of expressing negative emotions in performing the requested behavior. Conflicts with parents as a result of living together have also been considered as a potential area of stress prevalent in reviews such as that of Trianes et al. (2009). Other authors such as Palacios and Andrade (2008) point out that parental practices within the family are linked to positive or negative behavioral responses—at the same time, they are an essential element for understanding problematic behaviors in both childhood and adolescence. The results of this study are consistent with this approach. In the school context, conflicts are more related to non-validation of the child's behavior or direct denial of a child's request by others. This is consistent with the scientific literature reviewed, which includes incidents with classmates, peer rejection, and other problems related to interaction such as feelings of loneliness as frequent sources of stress in children between 6 and 12 years of age (Salmivalli et al., 2000; Ortega and Monks, 2005). However, our results differ from those of Filella et al. (2016), who conclude that physical and verbal aggression were the most frequent types in the study.

For hypothesis #3 the data reveal a significant correlation in conflicts originating at home at different times of the day—the closer they are to each other, the greater the correlation. This means that a conflict happening in the morning is related to the possibility of a conflict occurring in the afternoon—in fact, the morning is the time of the day when conflict is most likely to take place, and the evening is the time of the day when it is least likely to occur. On the other hand, a significant relationship was also observed between the conflicts occurring at home in the evening and the conflicts happening the following morning, both in the family and school contexts (hypothesis #4). Therefore, a spillover effect is observed within the family environment during the different times of the day—the following day is included in this spillover effect—, thus observing a spillover effect between the events occurring at the end of the day both in the family and the school contexts, adding information to the pattern demonstrated for the first objective. These findings are consistent with the existing scientific literature (Greenberger et al., 1994; Cummings and Davies, 2010; Fiese and Winter, 2010) suggesting that parenting tasks can be causal factors of child spillover which affect

TABLE 5 Results of Friedman's test and multiple *a posteriori* contrasts of conflicts at home in the morning during weekdays.

Comparison parents days of the week morning

Friedman	χ^2	p					
M-T-W-Th-F	32.44	0.000	**				
T Wilcoxon	M	SD	M	SD	T	p	
Monday–	1.887	2.313	1.432	1.904	1684.00	0.001**	$r=0.18$
Tuesday							
Monday–	1.887	2.313	1.421	2.206	1580.00	0.000**	$r=0.21$
Wednesday							
Monday–	1.887	2.313	1.258	1.785	1652.00	0.000**	$r=0.22$
Thursday							
Monday–	1.887	2.313	1.202	1.866	1305.50	0.000**	$r=0.25$
Friday							
Tuesday–	1.432	1.904	1.421	2.206	2148.00	0.505	
Wednesday							
Tuesday–	1.432	1.904	1.258	1.785	2086.50	0.066	
Thursday							
Tuesday–	1.432	1.904	1.202	1.866	1567.00	0.011	
Friday							
Wednesday–	1.421	2.206	1.258	1.785	1762.00	0.405	
Thursday							
Wednesday–	1.421	2.206	1.202	1.866	1564.00	0.096	
Friday							
Thursday–	1.258	1.785	1.202	1.866	1816.50	0.436	
Friday							

p* < 0.05; *p* < 0.01.

their development and manifest themselves in the school context in the form of conflicts with peers.

In relation to hypothesis #5, which states that peer interaction conflicts occurring in the school context will negatively affect daily parenting tasks at two times of the day—in afternoon parenting activities and in the development of such tasks in the evening, the spillover effect is only captured from Wednesday onward, through the connection between peer-peer conflicts and evening conflicts at home. This behavioral pattern that shows the spillover effect from Wednesday onward refers to the accumulation process described by Pakman (2006), who states that after an accumulation of tensions there is a rupture of equilibrium, leading to overflow and transference. In the same sense, and described at a more general level, the study by Kaufman et al. (2020) states that tension suffering due to victimization situations with peers is associated with conflicts in the parental-filial relationship.

Finally, the third objective was to analyze a possible bidirectional relationship between conflicts occurring at home and at school, respectively.

The results obtained allow us to state that this bidirectionality exists in the transfer of tensions (sixth hypothesis), although this spillover effect is linked to particular moments and events. The spillover effect between the family and school contexts is reflected daily through the relationship between evening events at home

TABLE 6 Results of Friedman's test and multiple *a posteriori* contrasts of home conflicts in the evenings during weekdays.

Comparison parents weekdays evenings

Friedman	χ^2	p					
M-T-W-Th-F	21.204	0.000	**				
T Wilcoxon	M	SD	M	SD	T	p	
Monday–Tuesday	1.790	2.337	1.601	1.716	1241.00	0.424	
Monday–Wednesday	1.790	2.337	1.587	1.725	1770.50	0.533	
Monday–Thursday	1.790	2.337	1.300	1.711	1211.00	0.004**	$r=0.17$
Monday–Friday	1.790	2.337	1.167	1.515	1256.00	0.002**	$r=0.18$
Tuesday–Wednesday	1.601	1.716	1.587	1.725	1885.50	0.756	
Tuesday–Thursday	1.601	1.716	1.300	1.711	1337.50	0.042	
Tuesday–Friday	1.601	1.716	1.167	1.515	1335.50	0.005**	$r=0.16$
Wednesday–Thursday	1.587	1.725	1.300	1.711	1295.50	0.036	
Wednesday–Friday	1.587	1.725	1.167	1.515	1023.00	0.001**	$r=0.19$
Thursday–Friday	1.300	1.711	1.167	1.515	1525.00	0.235	

p* < 0.05; *p* < 0.01.

and peer-peer tensions in the school context the following day. Nevertheless, the spillover effect between the school and family contexts is only captured from Wednesday onward and through the relationship between interpersonal conflicts between peers and evening conflicts at home. Consequently, the findings of this study are in line with the research of Repetti et al. (2002), which shows how school conflicts have their origin in family conflicts happening at home. It is also related to other publications where a spillover effect between the two contexts is identified through daily log reports (Bai et al., 2017), where positive same-day correlations were reported, and other work where the described bidirectional effect could be ascertained (Crnic et al., 2005; Flook and Fuligni, 2008; Mackler et al., 2015), as a “mirror image” according to Krpan et al. (2019). However, the results found between the school and the family contexts suggest a spillover effect due to a cumulative process of stress (Pakman, 2006)—in this case, school stress, as it was pointed out in the above paragraph. Therefore, the fact that the spillover effect does not take place during the first days of the week, but by the middle of the week, suggests that weekends could be a protective factor against it. Still, this possible relationship remains for further study.

When studying the spillover effect through correlational analysis, statistically significant associations do not always occur at all recorded and evaluated times, but only at particular

TABLE 7 Results of Friedman's test and multiple *a posteriori* contrasts of interpersonal conflicts with peers in the school setting.

Equal days of the week comparison

Friedman	χ^2	<i>p</i>				
M-T-W-Th-F	36.96	0.000	**			
T Wilcoxon	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>T</i>	<i>p</i>
Monday–Tuesday	8.679	6.538	7.989	6.466	5889.50	0.072
Monday–Wednesday	8.679	6.538	7.978	6.605	5886.50	0.071
Monday–Thursday	8.679	6.538	7.527	6.333	5014.50	0.003**
Monday–Friday	8.679	6.538	6.592	5.964	3643.00	0.000**
Tuesday–Wednesday	7.989	6.466	7.978	6.605	6525.50	0.993
Tuesday–Thursday	7.989	6.466	7.527	6.333	5672.00	0.119
Tuesday–Friday	7.989	6.466	6.592	5.964	4055.50	0.000**
Wednesday–Thursday	7.978	6.605	7.527	6.333	5619.00	0.201
Wednesday–Friday	7.978	6.605	6.592	5.964	4437.00	0.000**
Thursday–Friday	7.527	6.333	6.592	5.964	4518.00	0.003**

p* < 0.05; *p* < 0.01.

TABLE 8 Results of Spearman's correlations of interpersonal conflicts and conflicts at home throughout the week.

	Afternoon conflict	Evening conflict
Interpersonal Monday	−0.17	−0.04
Interpersonal Tuesday	0.12	0.18
Interpersonal Wednesday	0.02	0.18*
Interpersonal Thursday	0.02	0.32**
Interpersonal Friday	0.02	0.39**

p* < 0.05; *p* < 0.01.

moments, and the added fact that it is observed for 3 years in a row means that there is a cumulative effect of stress experiences. It should be noted that the longitudinal nature of the study—together with the data collection by daily recording during the different weeks—has undoubtedly allowed to capture the micro-relationships existing between the different moments (Chung et al., 2011; Sears et al., 2016; Sherrill et al., 2017). In other studies, based on questionnaires, it is not possible to detect this cumulative process, but the daily record method allows to do so. As Galizzi and Whitmarsh (2019) indicate, the longitudinal methodology does not identify causality, but it does quantify the strength of the

relationships between the measured behaviors, providing a standardized measure.

The *main finding* of this study refers to the description of the spillover effect between the different moments of parenting within the family context and the spillover effect from the family context to the school context and vice versa at 6–9 years of age. In addition, a characteristic behavioral pattern has been revealed over time—where Mondays are the most difficult days in family living—that morning conflicts are related to later conflicts, and that conflicts at home are related with one another. Evenings are usually the most peaceful time of the day at home, and Fridays are the calmest days of the week. Evening conflicts at home are related to conflicts at school and conflicts at school from Wednesday onward are connected to conflicts at home due to a possible cumulative factor in relation to daily conflicts.

The *strengths* of this study are based on the fact that empirical evidence of the spillover effect is still poorly documented today, and most research addresses it cross-sectionally through questionnaires measuring this phenomenon in an *ad hoc* manner. The spillover effect requires knowledge of the temporal intertwining of tensions and behaviors in a longitudinal design. These are day-by-day studies with a longitudinal approach of small amplitude in terms of the temporal sequence studied with the daily report methodology. As mentioned, in the case of young children, there are barely any studies that investigate spillover in relation to the family and peer systems, since the few studies that exist do so with adolescents, whose chronological stage is completely different. Finally, the methodology based on the daily report, despite its limitations in this study, proves to be a versatile method capable of capturing the course of time.

Limitations and future research

Although the spillover effect has been demonstrated in different times and contexts, this study has certain limitations. One of them was the low incidence of conflicting events obtained with the sample. One possible explanation may be the sociocultural context from which the families come and where the schools are located. This limitation can also be observed in other research such as that of Bai et al. (2017), who report little or few incidents. Another limitation is that the family socioeconomic level has not been specifically evaluated, since this variable was not considered relevant for the study when considering the conclusions of Flook and Fuligni (2008) and Nelson et al. (2009), where they did measure this variable, and it was determined that socioeconomic level does not influence *per se* the spillover effect. Despite these conclusions, it would be convenient to take this into account in future work for greater objectivity.

In future research, it would be interesting to use samples with different conflict incidence rates, and to carry out comparative studies through the relationship between the different

development contexts among various cultural environments. On the other hand, it is essential to continue research with the development of daily recording instruments that facilitate efficient data collection and meet the different criteria established according to criterion and construct validity. This study has described for the first time a behavioral pattern of related conflict behaviors, and it would be interesting to study how variables such as parenting styles, social skills, or coping strategies influence this relationship and pattern.

Implications for psycho-educational intervention

Stress linked to a particular life event has a great impact on the individual's development and on how they relate to their context during a given time—however, everyday stress can have a greater impact on the development of the child or adolescent (Seiffge-Krenke, 2000) since high levels of everyday stress are associated with significant negative consequences of emotional maladjustment and psychopathology (Seiffge-Krenke, 2000; Jose and Ratcliffe, 2004) that are prolonged over time. Consequently, the use of a more global theoretical and methodological perspective where the relationship of several interrelated mycosystems is taken into account can help to learn even more about the stressors with respect to such maladjustment, as well as the type of psychopathology which is more prone to situations and/or specific moments potentially susceptible to conflict, and the intrafamilial bonds affecting the education of children in this age group.

The psychoeducational intervention focused from the spillover perspective on daily stress should contemplate the behavioral pattern described here, as well as the cumulative effect of stressors, which give rise to the spillover of their negative effects. The aim of a psychoeducational intervention should be to reduce daily stressful events by promoting competent behaviors both in the family environment (through appropriate educational styles) and in the school environment, through socioemotional education. Reducing stressors, as well as having appropriate coping strategies, may hamper the cumulative process defining the spillover effect between contexts, thus increasing the individual's well-being. The daily events occurring in the school environment do not produce an overload of stresses on their own, but it is their accumulation over the course of the days that produces a behavioral imbalance in the child, which is transferred to the family, with respect to those days when the spillover effect does not take place.

Conclusion

Stress—both in adult life and in childhood—tends to be related to specific situations, but the findings of this study refer to the interrelation of contexts (Bronfenbrenner, 2002) and the way reactions produced by small events, problems, worries, and setbacks of high frequency, low intensity and high predictability (daily stress)

can be transferred from a particular context to a completely different one, and can manifest themselves at the behavioral level through conflicts in the environment (spillover effect).

In this sense, this research has shown that the spillover effect is present from a very early age, with bidirectionality between family and school contexts, although this spillover effect is linked to particular moments and events. The data also provide a characteristic behavioral pattern indicating the days of greatest conflict and reflecting the process of accumulation that occurs during the week. On the other hand, it points out the usefulness of the longitudinal cut study and the use of reports as the most plausible way to know better the daily experiences lived according to the daily stress in the identification of key moments. Finally, it shows that, if the sources of stress are known—together with the particular moments of conflicts, the stress accumulation process, and how the behavioral transfer of its consequences from one environment to another—it will be easier to prevent and provide coping and stress control strategies to both families and their children. This reduces the adverse effects and contributes to both the psychological and emotional well-being of the person.

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 Universidad de Málaga. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

LI-C, LA-C, AS-S, and AM-S contributed to the conception and design of the study. AW-R organized the database and performed the statistical analysis. LI-C, LA-C, and AW-R wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work has received funding from the Excellence Project by the Junta de Andalucía in 2011–2013: Spillover between parent–child and child–peer conflicts. Transmission of bidirectional tensions between parent–child (6–8 years old) and child–peer daily interactions at school. Consequences for school

performance, parental educational strategies, stress in parents and children, and psychopathology. Project code: SEJ-7226.

Acknowledgments

The authors of this paper would like to thank María Victoria Trianes for her contribution to the research group dedicated to the above-mentioned Excellence Project.

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

- Almeida, D. M. (2004). "Using daily diaries to assess temporal friction between work and family" in *Work-family Challenges for Low Income Parents and Their Children*. eds. A. C. Crouter and Y. A. Booth (Hillsdale, NJ: Lawrence Earlbaum Associates), 127–136.
- Bai, S., Reynolds, B. M., Robles, T. F., and Repetti, R. L. (2017). Daily links between school problems and youth perceptions of interactions with parents: a diary study of school-to-home spillover. *Soc. Dev. (Oxford, England)* 26, 813–830. doi: 10.1111/sode.12229
- Baker, B. L., Blacher, J., and Olsson, M. B. (2005). Preschool children with and without developmental delay. Behavior problems, parents' optimism and well-being. *J. Intellect. Disabil. Res.* 49, 575–590. doi: 10.1111/j.1365-2788.2005.00691.x
- Ballesteros, S., and García, B. (2001). *Procesos Psicológicos Básicos*. Madrid, España: Universitat.
- Bronfenbrenner, U. (2002). *La ecología del desarrollo humano: experimentos naturales y planificados*. Barcelona, España: Paidós Ibérica.
- Buckhalt, J. A., El-Sheikh, M., Keller, P. S., and Kelly, R. J. (2009). Concurrent and longitudinal relationships between children's sleep and cognitive functioning: the moderating role of parent education. *Child Dev.* 80, 875–892. doi: 10.1111/j.1467-8624.2009.01303.x
- Bulkeley, J., and Fabian, H. (2006). Well-being and belonging during early educational transitions. *Int. J. Transit. Child.* 2, 18–31.
- Chung, G. H., Flook, L., and Fuligni, A. (2011). Reciprocal associations between family and peer conflict in adolescents' daily lives. *Child Dev.* 82, 1390–1396. doi: 10.1111/j.1467-8624.2011.01625.x
- Coll, C., Marchesi, Á., and Palacios, J. (2014). *Desarrollo psicológico y educación*. Madrid, España: Alianza Editorial.
- Crnic, K. A., Gaze, C., and Hoffman, C. (2005). Cumulative parenting stress across the preschool period: relations to maternal parenting and child behaviour at age 5. *Infant Child Dev.* 14, 117–132. doi: 10.1002/icd.384
- Crompton, R., and Lyonette, C. (2006). Work-life 'Balance' in Europe. *Acta Sociol.* 49, 379–393. doi: 10.1177/0001699306071680
- Crouter, A. C. (1984). Spillover from family to work: the neglected side of the work-family interface. *Hum. Relat.* 37, 425–441. doi: 10.1177/001872678403700601
- Cummings, E. M., and Davies, P. T. (2010). *Marital Conflict and Children: An Emotional Security Perspective*. Nueva York, Estados Unidos: Guilford Press.
- Davies, P. T., Martin, M. J., and Cummings, E. M. (2018). Interparental conflict and children's social problems: insecurity and friendship affiliation as cascading mediators. *Dev. Psychol.* 54, 83–97. doi: 10.1037/dev0000410
- De Anda, D., Baroni, S., Boskin, L., Buchwald, L., Morgan, J., Ow, J., et al. (2000). Stress, stressors and coping among high school students. *Child Youth Serv. Rev.* 22, 441–463. doi: 10.1016/S0190-7409(00)00096-7
- Denham, S., Ferrier, D., Howarth, G., Herndon, K., and Bassett, H. (2016). Key considerations in assessing young children's emotional competence. *Camb. J. Educ.* 46, 299–317. doi: 10.1080/0305764X.2016.1146659
- Dockett, S., and Perry, B. (2007). *Transitions to School: Perceptions, Expectations, and Experiences*. Sydney: University of NSW Press.
- Du, D., Derks, D., and Bakker, A. B. (2018). Daily spillover from family to work: A test of the work-home resources model. *J. Occup. Health Psychol.* 23, 237–247. doi: 10.1037/ocp0000073
- El-Sheikh, M., Buckhalt, J. A., Mize, J., and Acebo, C. (2006). Marital conflict and disruption of children's sleep. *Child Dev.* 77, 31–43. doi: 10.1111/j.1467-8624.2006.00854.x
- Enns, A., Eldridge, G. D., Montgomery, C., and Gonzalez, V. M. (2018). Perceived stress, coping strategies, and emotional intelligence: A cross-sectional study of university students in helping disciplines. *Nurse Educ. Today* 68, 226–231. doi: 10.1016/j.nedt.2018.06.012
- Fabian, H., and Dunlop, A. W. (2006). Outcomes of good practice in transition processes for children entering primary school. In Paper Commission for the EFA global monitoring report 2007, strong Foundations: Early childhood care and education. Paris: UNESCO.
- Félix, V., Godoy, C., and Martínez, I. (2008). Violencia entre iguales: resultados de un estudio descriptivo de la provincia de Valencia. *Inform. Psicol.* 94, 36–48.
- Ferrarini, T. (2006). *Families, States and Labour Markets*. Cheltenham: Edward Elgar.
- Fiese, B. H., and Winter, M. (2010). "Family routines and rituals" in *The Wiley-Blackwell handbook of family psychology*. eds. J. Bray and M. Stanton (London: Yale University Press), 625–636.
- Filella, G., Ros-Morente, A., Rueda, P., Solano, A., and Minguela, E. (2016). Análisis de conflictos entre iguales en las escuelas de Educación Primaria. *FÉLI Revista de Educación Social* 23, 301–314.
- Flook, L. Y., and Fuligni, A. J. (2008). Family and school spillover in adolescents' daily lives. *Child Dev.* 79, 776–787. doi: 10.1111/j.1467-8624.2008.01157.x
- Franz, D. Z. Y., and Gross, A. M. (2001). Child sociometric status and parent behaviours. *Behav. Modif.* 25, 3–20. doi: 10.1177/0145445501251001
- Galizzi, M. M., and Whitmarsh, L. (2019). How to measure behavioral spillovers: a methodological review and checklist. *Front. Psychol.* 10. doi: 10.3389/fpsyg.2019.00342
- Gerard, J. M., Krishnakumar, A., and Buehler, C. (2006). Marital conflict, parent-child relations, and youth maladjustment: a longitudinal investigation of spillover effects. *J. Fam. Issues* 27, 951–975. doi: 10.1177/0192513X05286020
- Graham, S., Doubleday, C., and Guarino, P. A. (1984). The development of relations between perceived controllability and the emotions of pity, anger, and guilt. *Child Dev.* 55, 561–565. doi: 10.2307/1129967
- Greenberger, E., O'Neill, R. Y., and Nagel, S. K. (1994). Linking workplace and homeplace. Relations between the nature of adults' work and their parenting behaviors'. *Dev. Psychol.* 30, 990–1002. doi: 10.1037/0012-1649.30.6.990
- Grzywacz, J. G., Almeida, D. M., and McDonald, D. A. (2002). Workfamily spillover and daily reports of work and family stress in the adult labor-force. *Fam. Relat.* 51, 28–36. doi: 10.1111/j.1741-3729.2002.00028.x

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.2022.909928/full#supplementary-material>

- Grzywacz, J. G., and Carlson, D. S. (2007). Conceptualizing work-family balance: implications for practice and research. *Adv. Dev. Hum. Resour.* 9, 455–471. doi: 10.1177/1523422307305487
- Herrera, F., Ramírez, M. I., and Roa, J. M. (2004). “El desarrollo emocional, social y moral en la educación primaria (6-12 años)” in *Psicología de la educación y del desarrollo en contextos escolares*. eds. M. V. Trianes and J. A. Gallardo (Coord.) (Madrid: Pirámide), 259–286.
- Hill, J. (2005). Work-family facilitation and conflict, working fathers and mothers, work-family stressors, and support. *J. Fam. Issues* 26, 793–819. doi: 10.1177/0192513X05277542
- Jacobs, J., and Gerson, K. (2004). *The Time Divide: Work, Family and Gender Inequality*. Cambridge, MA: Harvard University Press.
- Jose, P. E., and Ratcliffe, V. (2004). Stressor frequency and perceived intensity as predictors of internalizing symptoms: gender and age differences in adolescence. *N. Z. J. Psychol.* 33, 145–154.
- Kanner, A. D., Coyne, J. C., Schaefer, C., and Lazarus, R. S. (1981). Comparison of two modes of stress measurement: daily hassles and uplifts versus major life events. *J. Behav. Med.* 4, 1–39. doi: 10.1016/0022-3999(67)90010-4
- Kaufman, T. M. L., Kretschmer, T., Huitsing, G., and Veenstra, R. (2020). Caught in a vicious cycle? Explaining bidirectional spillover between parent-child relationships and peer victimization. *Dev. Psychopathol.* 32, 11–20. doi: 10.1017/S0954579418001360
- Kinckad-Clark, Z. (2015). ‘Ready for big school’: making the transition to primary school – a Jamaican perspective. *Int. J. Early Years Educ.* 23, 67–82. doi: 10.1080/09669760.2014.999027
- Kouros, C. D., Papp, L. M., Goeke-Morey, M. C., and Cummings, E. M. (2014). Spillover between marital quality and parent-child relationship quality: parental depressive symptoms as moderators. *J. Fam. Psychol.* 28, 315–325. doi: 10.1037/a0036804
- Krpan, D., Galizzi, M. M., and Dolan, P. (2019). Looking at spillovers in the Mirror: making a case for “behavioral Spillunders”. *Front. Psychol.* 10:1142. doi: 10.3389/fpsyg.2019.01142
- Lawrence, E., Rothman, A. D., Cobb, R. J., and Bradbury, T. N. (2010). “Marital satisfaction across the transition to parenthood: three eras of research” in *Strengthening Couple Relationships for Optimal Child Development: Lessons From Research and Intervention*. eds. M. S. Schulz, M. K. Pruett, P. K. Kerig and R. D. Parke (Washington, DC: American Psychological Association), 97–114.
- Lehman, B. Y., and Repetti, R. (2007). Bad days Don’t end when the school bell rings: the lingering effects of negative school events on Children’s mood, self-esteem, and perceptions of parent-child interaction. *Soc. Dev.* 16, 596–618. doi: 10.1111/j.1467-9507.2007.00398.x
- Mackler, J. S., Kelleher, R. T., Shanahan, L., Calkins, S. D., Keane, S. P., and O’Brien, M. (2015). Parenting stress, parental reactions, and externalizing behavior from ages 4 to 10. *J. Marriage Fam.* 77, 388–406. doi: 10.1111/jomf.12163
- Martin, M. J., Sturge-Apple, M. L., Davies, P. T., Romero, C. V., and Buckholz, A. (2017). A process model of the implications of spillover from coparenting conflicts into the parent-child attachment relationship in adolescence. *Dev. Psychopathol.* 29, 417–431. doi: 10.1017/S0954579417000086
- Mastrotheodoros, S., Papp, L. M., Van der Graaff, J., Deković, M., Meeus, W., and Branje, S. (2022). Explaining heterogeneity of daily conflict spillover in the family: the role of dyadic marital conflict patterns. *Fam. Process* 61, 342–360. doi: 10.1111/famp.12648
- Minuchin, S. (1982). *Familias y Terapia Familiar*. Buenos Aires: Gedisa.
- Nelson, S. K., Kushlev, K. Y., and Lyubomirsky, S. (2014). The pains and pleasures of parenting: when, why, and how is parenthood associated with more or less well-being? *Psychol. Bull.* 140, 846–895. doi: 10.1037/a0035444
- Nelson, J. A., O’Brien, M., Blankson, A. N., Calkins, S. D. Y., and Keane, S. P. (2009). Family stress and parental responses to Children’s negative emotions: tests of the spillover, crossover, and compensatory hypotheses. *J. Fam. Psychol.* 23, 671–679. doi: 10.1037/a0015977
- Oliva, A., and Palacios, J. (2000). *La familia y escuela: padres y profesores en Familia y Desarrollo humano*. Madrid: Alianza.
- Olweus, D. (1991). “Bully/victim problems among schoolchildren: basic facts and effects of a school-based intervention program” in *The Development and Treatment of Childhood Aggression*. eds. D. J. Pepler and Y. K. H. Rubin (Hillsdale, NJ: LEA), 411–448.
- Oros, L. Y., and Vogel, G. (2005). Eventos que generan estrés en la infancia: diferencias por sexo y edad. *Enfoque* 17, 85–101.
- Ortega, R., and Monks, C. (2005). Agresividad injustificada entre preescolares. *Psicothema* 17, 453–458.
- Ortiz, M. J. (2001). “Desarrollo Emocional” in *Desarrollo afectivo y social*. eds. F. López, I. Etxebarria, M. J. Fuentes and M. J. Ortiz (Madrid: Pirámide), 95–121.
- Pakman, M. (2006). Reducción del riesgo en familias multiproblemáticas: la micropolítica de la justicia social en la asistencia en salud mental. *Terapia familiar y de pareja* 42, 698–717.
- Palacios, J. R. Y., and Andrade, P. (2008). Influencia de las prácticas parentales en las conductas problema en adolescentes. *Invest. Univ. Multidiscip.* 7, 7–18.
- Palacios, J., and Hidalgo, V. (1999). “Desarrollo de la personalidad desde los 6 años hasta la adolescencia,” En *Desarrollo psicológico y educación, 1. Psicología evolutiva*. eds. J. Palacios, A. Marchesi y C. Coll (Alianza: Madrid), 355–376.
- Phelan, P., Yu, H. C., and Davidson, A. L. (1994). Navigating the psychosocial pressures of adolescence: the voices and experiences of high school youth. *Am. Educ. Res. J.* 31, 415–447. doi: 10.2307/1163316
- Piotrkowski, C. (1979). *Work and the Family System: A Naturalistic Study of Working-class and Lower-middle-class Families*. New York, NY: Free Press.
- Pu, D. F., and Rodríguez, C. M. (2021). Bidirectional spillover in the family across the transition to parenthood. *Fam. Process* 60, 235–250. doi: 10.1111/famp.12549
- Ramírez, G., Chang, H., Maloney, E. A., Levine, S. C., and Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: the role of problem-solving strategies. *J. Exp. Child Psychol.* 141, 83–100. doi: 10.1016/j.jecp.2015.07.014
- Repetti, R., Taylor, S., and Seeman, T. (2002). Risky family social environments and physical and mental health of the offspring. *Psychol. Bull.* 128, 330–366. doi: 10.1037/0033-2909.128.2.330
- Repetti, R. Y., and Wang, S.-W. (2017). Effects of job stress on family relationships. *Curr. Opin. Psychol.* 13, 15–18. doi: 10.1016/j.copsyc.2016.03.010
- Rimm-Kaufman, S. E., and Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *J. Appl. Dev. Psychol.* 21, 491–511. doi: 10.1016/S0193-3973(00)00051-4
- Rivers, I. Y., and Smith, P. K. (1994). Types of bullying behaviour and their correlates. *Aggress. Behav.* 20, 359–368. doi: 10.1002/1098-2337(1994)20:5<359::AID-AB2480200503>3.0.CO;2-J
- Salamon, R., Johnson, E. I., and Swendsen, J. (2011). Daily life mechanisms of stress spillover among early adolescents experiencing academic difficulty. *Eur. J. Psychol. Educ.* 26, 453–463. doi: 10.1007/s10212-011-0056-7
- Salmivalli, C., Kaukiainen, A., and Lagerspetz, K. (2000). Aggression and sociometric status among peers: do gender and type of aggression matter? *Scand. J. Psychol.* 41, 17–24. doi: 10.1111/1467-9450.00166
- Sandín, B. (2008). “El estrés” in *Manual de psicopatología*. eds. A. Belloch, B. Sandín and Y. F. Ramos (Madrid, España: McGraw-Hill), 3–42.
- Schulting, A. B., Malone, P. S., and Dodge, K. A. (2005). The effect of school-based kindergarten transition policies and practices on child academic outcomes. *Dev. Psychol.* 41, 860–871. doi: 10.1037/0012-1649.41.6.860
- Sears, M. S., Repetti, R. L., Reynolds, B. M., Robles, T. F., and Krull, J. L. (2016). Spillover in the home: the effects of family conflict on parents’ behavior. *J. Marriage Fam.* 78, 127–141. doi: 10.1111/jomf.12265
- Seiffge-Krenke, I. (2000). Causal links between stressful events, coping style, and adolescent symptomatology. *J. Adolesc.* 23, 675–691. doi: 10.1006/jado.2000.0352
- Seiffge-Krenke, I. (2007). Stress, coping, and depression. Ponencia presentada en el Postgrado de Investigación en Procesos Básicos e Intervención en Psicología de la Salud, bienio 2006–2008, Universidad de Málaga, España.
- Sherrill, R. B., Lochman, J. E., DeCoster, J., and Stromeier, S. L. (2017). Desbordamiento entre el conflicto entre padres y el conflicto entre padres e hijos dentro y entre días. *Revista de psicología familiar: JFP: revista de la División de Psicología Familiar de la Asociación Estadounidense de Psicología (División 43)* 31, 900–909. doi: 10.1037/fam0000332
- Sidorowicz, K., Hair, E. C., and Milot, A. (2009). Assessing bullying: A guide for out-of-school time program practitioners (Publication No. 2009-42). Retrieved from the Child Trends website: http://www.childtrends.org/wp-content/uploads/2009/10/child_trends-2009_10_29_rb_assessingbullying.pdf
- Soldevila, A., Filella, G., Ribes, R., and Agulló, M. J. (2007). Una propuesta de contenidos para desarrollar la conciencia y la regulación emocional en Educación Primaria. *Cult. Educ.* 19, 47–59. doi: 10.1174/113564007780191278
- Staines, G. L. (1980). Spillover versus compensation: a review of the literature on the relationship between work and non-work. *Hum. Relat.* 33, 111–129. doi: 10.1177/001872678003300203
- Stroud, C. B., Durbin, C. E., Wilson, S. Y., and Mendelsohn, K. A. (2011). Spillover to triadic and dyadic Systems in Families with young children. *J. Fam. Psychol.* 25:919. doi: 10.1037/a0025443
- Trianes, M. V. (2002). *Estrés en la infancia*. Madrid, España: Narcea.
- Trianes, M. V., Blanca, M. J., Fernández-Baena, F. J., Escobar, M., Maldonado, E. F., and Muñoz, A. M. (2009). Evaluación del estrés infantil: Inventario Infantil de Estresores Cotidianos (IIEC). *Psicothema* 21, 598–603.
- Trianes, M. V. Y., and Escobar, M. (2009). “Estrés cotidiano en la infancia media” in *Educación y menores en riesgo*. ed. F. Etxebarria (Barcelona, España: Sello Editorial), 283–297.
- Wolin, S. J. Y., and Bennet, L. A. (1984). Family rituals. *Fam. Process* 23, 401–420. doi: 10.1111/j.1545-5300.1984.00401.x



OPEN ACCESS

EDITED BY

Douglas F Kauffman,
Medical University of the Americas – Nevis,
United States

REVIEWED BY

David Aparisi,
University of Alicante,
Spain
Andrew Sheldon Franklin,
Norfolk State University,
United States

*CORRESPONDENCE

Feifei Han
✉ feifei.han@acu.edu.au

SPECIALTY SECTION

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

RECEIVED 23 July 2022

ACCEPTED 08 December 2022

PUBLISHED 06 January 2023

CITATION

Han F, Juklová K, Mikoška P and
Novák L (2023) The stability of the twofold
multidimensionality of academic self-
concept: A study of Chinese secondary
school students.
Front. Psychol. 13:1001187.
doi: 10.3389/fpsyg.2022.1001187

COPYRIGHT

© 2023 Han, Juklová, Mikoška and Novák.
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 stability of the twofold multidimensionality of academic self-concept: A study of Chinese secondary school students

Feifei Han^{1,2*}, Kateřina Juklová², Petr Mikoška² and Lukáš Novák³

¹Institute for Learning Sciences and Teacher Education, Australian Catholic University, Brisbane, QLD, Australia, ²Department of Pedagogy and Psychology, Faculty of Education, University of Hradec Králové, Hradec Králové, Czechia, ³Olomouc University Social Health Institute, Palacký University in Olomouc, Olomouc, Czechia

Introduction: The present investigation examined the stability of the twofold multidimensional structure of academic self-concepts (ASCs) in three domains, namely Chinese, math, and general school using four-wave data collected over 2 years among 552 Chinese secondary school students.

Method: Adopting both a within-network and a between-network approach, confirmatory factor analyses (CFAs) and factor correlations were performed in Mplus 8.2.

Results: The within-network results showed that CFA models wherein competence and affect dimensions were conflated generated unacceptable fit. In contrast, the CFAs in which competence and affect were modeled as separate latent factors consistently produced superior fit to the data. The between-network results demonstrated that in the Chinese and math domains and across the four-time waves, the competence components were more strongly related to the achievements in matching domains than the affect components were. Furthermore, both the competence and affect components of ASCs and achievements were positively correlated in the non-matching domains, which were somewhat contradictory to the internal/external frame of reference model predicting zero or negative relations.

Discussion: Such results seem to suggest more involvement in social comparison than in dimensional comparison of Chinese students, which might be attributed to the collectivistic Chinese culture and the common phenomenon of academic social comparisons among Chinese adolescents in schools.

KEYWORDS

academic self-concept, twofold multidimensionality, stability, Chinese secondary school students, internal/external frame of reference model

1. Introduction

Defined as “a person’s perception of himself formed through his experience with his environment and influenced especially by environmental reinforcements and significant others” (Shavelson et al., 1976, 411), self-concept (SC) is widely acknowledged as an important psychological construct, vital in all parts of human life, and it has been researched in a variety of fields, including education, psychology, sport sciences, mental health, and many more (Marsh and Craven, 2006; Onetti-Onetti et al., 2019; Karimova and Csapó, 2021; Henning et al., 2022; Melguizo-Ibáñez et al., 2022). For decades, empirical research has shown that having positive SC is beneficial for many areas of human functioning (Marsh et al., 2017): fighting against anxiety and distress, for strengthening confidence and resilience, and for being adaptive to handle complex problems and situations (Coopersmith, 1960; Harter, 1990; Dodgson and Wood, 1998; Sommer and Baumeister, 2002), and thus enables many human potentials (Marsh and Craven, 2006; Marsh and O’Mara, 2008).

Among various forms of SC, students’ self-perceptions in school curriculum domains (e.g., general school, math, verbal, and science), known as academic self-concept (ASC), have been extensively researched. Research has been conducted to identify factors which may influence ACS and has shown that teacher relatedness (Guay et al., 2019), peer feedback (Simonsmeier et al., 2020), and socioeconomic status (Chevalère et al., 2022) can impact ACS. Furthermore, ASC has been found to be closely associated with various desirable academic outcomes, such as academic achievement, coursework selections, learning motivation and confidence, course and career aspirations, effort expenditure, and use of self-regulated learning strategies (Marsh and Craven, 2006; Marsh and Martin, 2011; Guo et al., 2015; Yang et al., 2016; Kadir et al., 2017; Schneider and Sparfeldt, 2020; Deeba et al., 2022; Maynor et al., 2022). In the past five decades, this knowledge has been constantly expanded and refined by new findings on the multidimensional nature of ASC (Marsh et al., 1999a); the reciprocal relationships between ASC and academic achievement (Marsh and Craven, 2006; Marsh and O’Mara, 2008); the domain-specific relations of ASC with academic achievement (Chiu and Klassen, 2010; Kadir et al., 2017; Arens and Preckel, 2018); and the effects of internal and external comparison processes in the formation of ASCs and explaining the interrelations of ASCs in different domains (Möller et al., 2009; Möller and Marsh, 2013).

An equally important finding is the validity of the twofold structure of ASC, which conceptualizes ASC as being formed by distinct cognitive (based on the student’s self-perceptions of the competence in a domain) and affective (based on the student’s feelings toward a relevant domain) factors (Marsh et al., 1999b). Past research has demonstrated that ASC has both multidimensional and twofold characteristics, known as “the twofold multidimensional structure of ASC” (Rosenberg, 1979; Arens et al., 2011; Gorges and Hollmann, 2019; Smith, 2019). However, most of these studies, whether cross-sectional or longitudinal, either only investigated the twofold structure of ASC

in a single domain or examined it across more domains, but at one time point, hence failed to test the stability of the multidimensional and twofold structure simultaneously. The present study aims to fill this gap by investigating the stability of the twofold multidimensional structure of ASC in three domains—general school, verbal, and math—among Chinese secondary students using three-wave longitudinal data.

2. Theoretical framework

2.1. Theoretical bases of twofold multidimensional ASC

Since the breakthrough in SC research by Shavelson et al.’s review (Shavelson et al., 1976), which refuted the unidimensional conceptualization of SC (Rosenberg, 1979; Marx and Winne, 1978), modern theory has emphasized SC as a multidimensional construct (Arens et al., 2011). While Shavelson et al. proposed a global SC which is formed by a nonacademic SC and an ASC factor, confirmatory factor analyses (CFAs) could not identify a global ASC with students in different age groups, including preadolescents (Marsh, 1990c), adolescents (Marsh, 1990b), and late adolescents and young adults (Marsh and O’Neill, 1984). This resulted in the Marsh/Shavelson model of ASC (Marsh, 1990d), which conceptualizes students’ ASC as being represented by two distinct higher-order factors of verbal and math ASC. Students’ ASCs in various curriculum domains are conceptualized as first-order factors on a continuum, with verbal and math ASCs on the two ends of the continuum and school ASC is in the middle. Therefore, at the one end of the continuum, the first-order verbal ASC has exclusive loadings on the second-order verbal ASC; at the other end of the continuum, the first-order math ASC has exclusive loadings on the second-order math ASC; and in the middle of the continuum, the first-order school ASC has loadings on both the second-order verbal and math ASCs. Such conceptualization has received support from many empirical studies which identified distinct ASCs in various curriculum domains, including general school, verbal, math, physics, art, music, history, and many other school subjects (Marsh et al., 1988; Marsh, 1990a; Arens et al., 2011).

In addition to its multidimensionality, another feature of ASC is its twofold structure, which further makes a distinction between the competence (the cognitive aspect of SC) and affect (the affective aspect of SC) components within each curriculum domain (Marsh et al., 1999b; Arens et al., 2011, 2014b, 2016). Although the competence and affect ASC has been treated as a single factor for some time, the twofold structure of ASC has a sound theoretical basis. For instance, Irwing (1996) provided some theoretical arguments as to why competence and affect should be distinguished. He maintains that essentially an individual tends to adopt different benchmarks when evaluating levels of capabilities and affective status. When judging one’s own competence, a person tends to use others’ qualities as a reference point whereas when evaluating one’s own affect, an internal comparison of feelings between one domain

and another domain is more likely. In line with Irwing's arguments are the assumptions of the internal/external frame of reference model (I/E model) postulated by Marsh (1986), who offers reasons for different relations between ASCs and academic achievements within and across domains. The I/E model similarly assumes the operations of both internal and external comparison, through which the domain-specific ASCs are formed. In an updated I/E model, Möller and Marsh (2013) proposed a dimensional comparison model which considers dimensional (similar to internal) and social (similar to external) comparison as mechanisms establishing interrelations between ASCs and achievements both within and across domains. Both the I/E and dimensional comparison models posit that internal (dimensional) comparison is based on comparing one's own abilities in one domain with own abilities in other domains, whereas external (social) comparison considers an individual's own ability against the perceived ability of his/her classmates. Given the operation of both internal (dimensional) and external (social) comparisons in the processes of forming domain-specific ASCs, despite the high positive correlations among academic achievements across domains, ASCs in dissimilar domains are uncorrelated. Furthermore, compared to significant positive within-domain relations between ASCs and achievements, zero, near zero or negative cross-domain correlations between achievement and academic self-concept are expected (Marsh, 1986, 2014; Marsh et al., 2012; Möller and Marsh, 2013).

Distinguishing the cognitive and affective ASC components and simultaneously assuming that social comparison is more involved in assessing the cognitive (competence) component and dimensional comparison is more present in evaluating the affective (affect) component, it can be expected that the relations between the two components of ASCs and achievements in the corresponding domains will differ from those across academic domains. In particular, while achievement will have positive associations with both competence and affect within a domain, it is expected that the strength of the association between affect and achievement should be weaker than that between competence and achievement, as more dimensional comparison takes place in evaluating affect. Across domains, correlations between the two components of ASC and academic achievement can be expected to be near zero, zero, or negative, and such correlations will also be weaker for affect and achievement than for competence and achievement.

2.2. The within-network approach to testing the twofold multidimensionality of ASC

The co-existence of the multidimensional and twofold nature of ASC is known as “the twofold multidimensional ASC” (Arens et al., 2011; Yang et al., 2016). In the past two decades, many studies have been conducted to empirically test the twofold multidimensionality of ASC adopting both a within-network and a between-network approach. According to Arens et al. (2011), the within-network approach focuses on evaluating the validity of internal structure of

ASC by adopting the exploratory and/or confirmatory factor analyses (CFAs), whereas the between-network approach emphasizes investigating the extent to which the constructs of ASC and other external criteria (constructs) overlap. Hence, the within-network approach examines the internal validity of ASC, whereas the between-network approach investigates the external validity of ASC. Researchers strongly recommend including both the within-network and the between-network approaches for rigorous testing of the construct validity of ASC (Shavelson et al., 1976; Byrne, 1984, 1996; Arens et al., 2011; Arens and Preckel, 2018).

Studies adopting the within-network approach have examined twofold multidimensional ASC with a diverse population. The general findings of these studies provided empirical evidence that the twofold multidimensional structure of ASC is generalizable to students at different stages of schooling, including students in preschools (Arens et al., 2016), primary schools (Marsh et al., 1999b; Arens and Hasselhorn, 2015; Arens and Morin, 2016; Arens and Preckel, 2018; Lohbeck, 2019; Schneider and Sparfeldt, 2020), secondary schools (Marsh et al., 1999b, 2013; Arens et al., 2014a; Kadir et al., 2017), vocational schools (Yang et al., 2016), and universities (Burns et al., 2018). Research has also been carried out to examine if the twofold multidimensional structure of ASCs is also applicable to students from diverse cultures and has found support among students from Western cultures (e.g., German: Arens et al., 2011; Lohbeck and Möller, 2017; Lohbeck, 2019; Schneider and Sparfeldt, 2020; Dutch: Pinxten et al., 2013; French-Canadian: Marsh and Ayotte, 2003; Anglo-Saxon: Marsh et al., 1999b); from Arabic cultures (e.g., Abu-Hilal et al., 2013; Marsh et al., 2013); from Indigenous cultures (e.g., Indigenous Australian: Arens et al., 2014a); and from Eastern countries (e.g., China: Yang et al., 2016; Leung, 2019).

2.3. The between-network approach to testing the twofold multidimensionality of ASC

Not all the studies adopting a within-network approach also include a between-network approach to testing the twofold multidimensionality of ASC. Those studies including the between-network approach also provide some evidence for the validity of the twofold multidimensionality of ASC. Of the different types of the external criteria, the short-term educational outcome—academic achievement—has been the one most frequently used. Studies have reported that consistently in various academic domains, the competence and affect components of ASC have different relations with academic achievement among students at different stages of schooling and in various school domains, including students in primary schools (Arens et al., 2011: general school, verbal, and math; Arens and Preckel, 2018: verbal, math, and physical education; Lohbeck, 2019: verbal and math; Schneider and Sparfeldt, 2020: verbal and math) and in secondary schools (Abu-Hilal et al., 2013: math and science; Arens et al., 2014b: verbal and math;

Kadir et al., 2017: verbal, math, and physics; Marsh et al., 2013: math and science; Karimova and Csapó, 2021: verbal and foreign language).

For instance, with German primary school students, Arens et al. (2011) found that consistently across three domains, namely general school, German, and math, the correlations between competence and achievement were doubled those between affect and achievement in the matching domains. Such results were corroborated by the results in Abu-Hilal et al. (2013) with 2,687 eighth-grade Saudi students in math and science domains that the associations between competence and achievement were more than triple those between affect and achievement in the matching domains. However, such patterns of the different relations between competence-affect and learning outcomes were not found when the external criteria were course aspirations and effort, which were considered as being longer-term learning outcomes. As argued by Yeung et al. (2012) that students' high sense of competence is more likely to have positive relations with the assessment-based outcomes, such as GPAs and test results, students' strong liking tends to be positively associated with longer-term outcomes, such as persistence and motivation in learning the subject.

For example, Marsh et al. (2013) reported that among 8th graders, the relations between affect and course aspirations in the matching domain were more strongly correlated than those between competence and course aspirations in the matching domains. Such results were reasonably consistent across math and science domains and across two cultures (Anglo-Saxon culture: students from the United States, England, Australia, and Scotland and Arab cultures: students from Saudi Arabia, Jordan, Oman, and Egypt).

Using effort as an external criterion, however, the relational patterns are not consistent. With German primary school students, Arens and Hasselhorn (2015) found that consistently in both verbal and math domains, the correlations between affect and effort were stronger than those between competence and effort in the matching domains. Also with German primary school students, however, Lohbeck (2019) reported that the correlations between affect and effort in the corresponding domains were similar to those between competence and effort in the corresponding domains. The two studies adopted different ways of measuring students' perceptions of effort. While Arens and Hasselhorn (2015) measured effort in a general sense, Lohbeck (2019) assessed effort in domain-specific ways. Whether such difference contributed to the inconsistent results found in the two studies requires further investigations in future studies.

Whether adopting within-network and/or between-network approaches, the majority of existing studies are cross-sectional, and, therefore, despite support for the twofold multidimensional structure of ASC, they were not able to test whether such structure was stable across time. Essentially, longitudinal studies are needed in order to show whether the twofold multidimensional structure of ASC can be clearly established over time.

2.4. The stability of twofold multidimensional ASC

To date, little research has been carried out to examine the stability (longitudinal design) of the twofold multidimensional structure of ASC. The longitudinal design which examined the validity of separating the competence and affect aspects of ASC have predominantly been held in a single domain (general school: Han, 2019; math: Pinxten et al., 2014; Arens et al., 2016; psychology: Burns et al., 2018). While these studies confirmed the stability of separating competence and affect in these specific domains, it failed to address the twofold structure and multidimensional structure of ASC simultaneously.

Adopting both a within- and a between-network approach, Pinxten et al. (2014) collected five-wave data to examine: (1) the structure of Dutch primary school students' math ASC in terms of a single factor conflating competence and affect or a clear distinction between competence and affect and (2) the relations between math ASC (competence-affect) on the one hand and math achievement and math effort expenditure on the other hand. The results of the within-network analysis demonstrated that a two-factor model (separating competence and affect) was superior to a one-factor model (conflating both competence and affect components). The results of the between-network approach showed that the relations between competence-affect and achievement had reversed pattern with the relations between competence-affect and effort. Specifically, reasonably across time, the relations between competence and achievement were stronger than those between affect and achievement. In contrast, the associations between competence and effort were weaker than those between affect and effort. The study provided support for the stability of the twofold structure of math self-concept and the stability of the relations between competence-affect and short-term (i.e., achievement) and long-term (i.e., effort) educational outcomes in math domain.

In terms of investigation of the stability of ASC in multiple domains, Marsh et al.'s (1999b) study is the only one that examined whether separating competence and affect was stable across a number of domains. As the participants in Marsh et al. are Australian primary school students, there is a lack of investigation of the stability of the twofold multidimensionality of ASC with older students (e.g., secondary school students) or in other cultures (e.g., Eastern cultures). Another limitation of Marsh et al.'s study was that it did not examine if competence-affect had different relations with academic outcomes (e.g., academic achievement), while cross-sectional studies in multiple domains have reported that the competence ASC had a higher correlation with academic achievement than the affect ASC not only within the matching domain but also across domains (Arens et al., 2011). The longitudinal studies in a single domain also reported that the different relations between competence-affect and academic achievement were consistent across time waves (Pinxten et al., 2013; Arens et al., 2016; Han, 2019). However, the single-domain longitudinal studies failed to investigate the longitudinal patterns of competence-affect and academic achievement across domains.

2.5. Present investigation

To fill the above-mentioned gaps in the literature, the aims of the present investigation were twofold. First, we aimed to extend Marsh et al. (1999a) examination of the stability of the twofold multidimensional ASC to a much less researched population—students in mainland China. As the students in Marsh et al. were primary school students, we would target an older population—secondary school students. Second, Marsh et al.'s study did not include students' academic achievement and thus it was not able to comprehensively test the construct validity of ASC by including external criteria. While both cross-sectional and longitudinal studies reported that competence had stronger association with achievement than affect did in the matching domain, the present study would investigate whether the differential associations between competence and affect in relation to academic achievement in the matching domain were consistent across domains and across time among Chinese secondary school students. To fulfill the two research aims, the present study aimed to investigate the stability of the twofold multidimensionality of ASC with Chinese secondary school students in the general school, verbal and math domains in a four-wave dataset collected over 2 years by adopting both a within-network and a between-network approach.

According to the literature review, the following three hypotheses were formulated, one for the within-network approach, and two for the between-network approach:

H1: Across four waves, Chinese secondary school students' ASCs would consist of competence and affect factors rather than a single factor conflating competence and affect in all the three domains (the within-network approach).

H2: Across four waves, self-perceptions of competence would be more strongly associated with achievements than self-perceptions of affect within both the Chinese and math domains (the within-network approach).

H3: Across four waves, the relations between competence and affect components and achievements would be zero or negative in non-matching domains.

3. Materials and methods

3.1. Participants and data collection procedure

The study was conducted in a junior secondary school in mainland China. The reason for using one school was all the students would sit the same examinations, allowing their academic

performance to be compared. The recruitment of the participants targeted grade 7 students (the first year of junior secondary school). Because the study was designed as a longitudinal study, lasting for 2 years. Chinese junior students in the last year of junior secondary school face pressure of attending the high school matriculation examination. This meant that when students would complete the last wave of data collection, they would be in grade 8. There were 630 students in grade 7 and all of them were invited to participate on a voluntary basis. Altogether 552 students (268 girls: 48.6%; and 284 boys: 51.4%) participated in the study, accounting for 87.6% of the total population.

Ethical procedures were strictly followed. Before the data collection, written consent was obtained from all the participants and their parents. T1 data collection was done approximately toward the end of the first term in grade 7, and T2 data collection was done approximately toward the end of the second term in grade 7. T3 and T4 data collection was done toward the end of the first and second terms, respectively, in grade 8. The ages of the participants were approximately between 13 and 15 years old during the first round of data collection (T1) and were between 15 and 17 years old during the last round of the data collection (T4). The vast majority of students stayed for the whole duration of 3 years of junior high school; therefore, missing data amounted to only 6.7%. To deal with this missing data, we adopted the full information maximum likelihood (FIML) estimation in Mplus. According to Graham (2009), full FIML is generally considered to be a better approach to dealing with missing data than the traditional approaches, such as listwise deletion and pairwise deletion.

3.2. Measures

3.2.1. ASC

We used the Chinese version of SDQ I to measure the participants' ASC in general school, math, and Chinese. The psychometric properties of the Chinese version of SDQ I were validated with Chinese primary and secondary students in mainland China (Watkins et al., 1997). As our participants were secondary school students in mainland China, it was considered appropriate to use this instrument. We used all the original items for the general school ASC and math ASC scales in the Chinese version of SDQ I, which corresponded with the items in the original English SDQ I. However, we modified the items in the reading ASC scale as past research has suggested that the reading ASC was only a sub-facet of the verbal ASC and represents specific sets of verbal ASC skills (Yeung et al., 2012; Arens et al., 2014b). Thus, the word "reading" in the reading ASC in the Chinese version of SDQ I was replaced by the word "Chinese" to present students' verbal ASC in their native language. For instance, the item "I am good at reading" was changed into "I am good at Chinese."

There were 10 for each domain and the wording of items was paralleled in the three domains. In each domain, five items captured students' self-perceptions of capability in terms of their performance (i.e., the cognitive aspect of ASC), and another five

items assessed students' self-appraisal of their enjoyment and liking of the subject (i.e., the affective aspect of ASC). Of the five items, four items were positively worded and one was negatively worded item, which was useful as the negatively worded items helped tease out any possible response bias (Arens et al., 2011). All the items were placed on a 5-point Likert scale, with the anchors representing the following: 1 = false, 2 = mostly false, 3 = sometimes false, sometimes true, 4 = mostly true, and 5 = true.

3.2.2. Academic achievement

As the students were from the same school and the same grade, we used students' performance in the final-term Chinese and math exams as indicators of their academic achievement. Due to the strict policies in the school, the exams had high standards and all the students were required to take part in the exams. We obtained their scores directly from the school's administrative system.

3.3. Data analysis

Data analyses were conducted using the within-network approach and the between-network approach, both of which performed a series of CFAs, which aim to test pre-specified models based on theory and previous literature (Brown, 2015). Mplus version 8.2 was used to analyze the data. We followed the general procedures proposed by Kline (2015) and Joreskog and Sorbom (2005) to evaluate the fit of CFAs, and considered the following goodness-of-fit indices as primary indicators of model fit: the Tucker-Lewis Index (TLI) (Tucker and Lewis, 1973), the Comparative Fit Index (CFI) (Bentler, 1990), and the root mean square error of approximation (RMSEA) (Browne and Cudeck, 1992). According to Bentler (1990) and Hu and Bentler (1999), a TLI or CFI value higher than 0.90 is generally considered an acceptable fit to the data. With regard to the RMSEA value, Browne and Cudeck (1992) suggest that RMSEA values below 0.06 are indicative of a good fit between the hypothesized model and the observed data. In addition to these fit statistics, we also consulted the two other criteria recommended by Joreskog and Sorbom (2005) in order to evaluate the appropriateness of the models. The first criterion was that the factor loadings of items in corresponding scales should be greater than 0.30. Secondly, the correlations between the two factors should be smaller than 0.90 in order to ensure the two scales are clearly distinguishable from each other.

3.3.1. Within-network approach to CFAs

The within-network approach to CFA tests the internal structure of ASC. We first constructed two models separately for each domain: models 1 and 2 for ASC in Chinese, models 3 and 4 for ASC in math, and models 5 and 6 for ASC in general school. In each domain, the first models (models 1, 3, and 5) were one-factor CFAs of ASC, in which all the ten items were conflated to represent a single latent ASC factor. The second models in each domain (models 2, 4, and 6) were two-factor CFAs of ASC, with five items forming a competence factor of competence and the other five items forming an affect factor.

Following these separate CFAs in each domain, we then constructed CFAs for the three domains together in order to examine the twofold multidimensional structure of ASC. Model 7 comprised three separate factors for ASCs in Chinese, math, and general school respectively, but ignored the distinction between competence and affect in each domain, whereas model 8 was established to form a six-factor model which consisted of three domains and separated competence and affect in each domain. For all the models, the error terms of the same items were correlated as each item was used repeatedly for each wave of data collection. In addition, we also included correlated uniquenesses in models 7 and 8, as parallel worded items were used across Chinese, math, and the general school domains.

3.3.2. Measurement invariance

Before conducting between-network approaches to CFAs, we also conducted measurement invariance test to examine if the latent construct stayed the same over time (Möller et al., 2011). The invariance tests were conducted by increasingly constraining parts of the measurement structure. We followed Brown's (2015) recommended procedure for performing the invariance tests in a stepwise manner from loose to tight. The invariance tests involved evaluating three levels of restricted models: (1) a configural model (10A), which tested whether the factor structures were identical across years. (2) the metric model (10B), which tested whether the factor loadings were equal; and (3) the scalar model (10C), which tested whether the intercepts were equal. For model comparison, we adopted the following guidelines: a decrease in fit of less than 0.010 in CFI (Cheung and Rensvold, 2002; Chen, 2007) and an increase in fit of less than 0.015 in RMSEA (Chen, 2007) would support the more parsimonious model.

3.3.3. Between-network approaches to CFAs

For between-network approaches, we integrated academic achievement in Chinese and math as external validity criteria into the three final retained CFA models in the within-network approaches. Similar to the procedure adopted in our within-network approach analyses, we also constructed separate CFAs for each domain first (model 11 for Chinese and model 12 for math respectively). If the one-factor models (10 items conflating competence and affect) had a better fit, we would add achievement into the one-factor models, and if the two-factor models (5 items representing competence and 5 items representing affect) produced better fit, we would add achievement to the two-factor models. When Chinese, math, and general school considered together (model 13), if the three-factor model had a better fit, we would add Chinese and math achievement results into the three-factor model; and if the six-factor model had a better fit, we would add Chinese and math achievement results into the six-factor model. As the achievement scores were single-item indicators, the measurement error of scores was fixed as 0.95.

4. Results

4.1. Results of the within-network approach to CFAs

Models 1 to 10 represented the within-network approach to CFAs and their results are displayed in the upper half of Table 1, which summarizes the values of Chi-square, degree of freedom, and the fit statistics of the within-network CFA models. The results of the CFAs conducted for the three domains separately were as follows: Chinese: model 1: χ^2 (674) = 2927.83, CFI = 0.83, TLI = 0.81, RMSEA = 0.07; math: model 3: χ^2 (674) = 2788.58, CFI = 0.87, TLI = 0.85, RMSEA = 0.07; general school: model 5: χ^2 (674) = 3629.01, CFI = 0.78, TLI = 0.75, RMSEA = 0.08; and these were the results for the three domains simultaneously: model 7: χ^2 (6654) = 16229.77, CFI = 0.82, TLI = 0.81, RMSEA = 0.06. All showed that the one-factor CFIA models of ASC failed to yield acceptable fit, even when the negatively worded items were deleted (model 8: χ^2 (4158) = 10360.87, CFI = 0.86, TLI = 0.85, RMSEA = 0.05).

In contrast, the two-factor CFAs of self-concept wherein competence and affect represented distinct latent factors generated good fit when Chinese, math, and general school were examined separately (Chinese: model 2: χ^2 (652) = 1374.77, CFI = 0.95, TLI = 0.94, RMSEA = 0.04; math: model 4: χ^2 (652) = 1217.13, CFI = 0.97, TLI = 0.96, RMSEA = 0.04; general: model 6: χ^2 (652) = 1549.95, CFI = 0.93, TLI = 0.92, RMSEA = 0.05); and when the three domains were jointly considered (model 9: χ^2 (6444) = 10865.93, CFI = 0.92, TLI = 0.91, RMSEA = 0.03). On the basis of model 9, we then eliminated negatively worded items of each domain to construct model 10, and the fit statistics showed even better fit (model 10: χ^2 (3948) = 6283.48, CFI = 0.95, TLI = 0.94, RMSEA = 0.03).

Apart from the fit statistics, we also checked the factor loadings in these models, and the results show that the loadings of items for their corresponding factors were all above 0.65. As model 10 which tested the twofold multidimensional structure of ASC in four-time waves by eliminating negatively worded items produced better fit, we used model 12 to calculate the factor correlations, which were all below 0.86. Furthermore, we also calculated Cronbach's alpha coefficients for all the competence and affect scales in the three domains across four-time waves, which showed that all the values had fairly good reliability, as all the α s were above 0.70 (see Table 2). The descriptive statistics, including *Ms* and *SDs* of competence and affect in the three domains across four-time waves are also displayed in Table 2. The factor loadings, factor correlations, and values of scale reliability together with the fit statistics demonstrated that the twofold multidimensional structure of ASC in the Chinese, math, and general school domains was stable across 3 years for Chinese secondary school students.

4.2. Results of measurement invariance tests

As model 10 produced a better fit, we thus examined the twofold multidimensional structure of ASC in four-time waves without negative items. The results of the invariance tests are presented in Table 3. Our configural model (without constraining of factor loadings) produced good fit: χ^2 (4092) = 1549.95, CFI = 0.933, TLI = 0.925, RMSEA = 0.034. Our metric model which constrained all factor loadings of the same latent construct to be equal over time resulted in a non-substantial decline in model fit: χ^2 (4146) = 7180.903, CFI = 0.932, TLI = 0.925, RMSEA = 0.034. The results suggest that the factor loadings of the twofold multidimensional structure of ASC, which separated competence and affect in Chinese, math, and general school, were equal over four-time points. Similarly, our scalar model also had a non-substantial decline in model fit: χ^2 (4200) = 7485.158, CFI = 0.927, TLI = 0.920, RMSEA = 0.035, supporting the invariance intercepts across time.

4.3. Results of correlations between competence and affect

The factor correlations are displayed in Tables 4–7 for T1–T4, respectively. In general, all the cross-sectional correlations between competence and affect in each domain were positive. However, competence and affect were more highly correlated in Chinese (T1–T4: r s = 0.57, 0.62, 0.59, and 0.56) and math (T1–T4: r s = 0.60, 0.63, 0.56, and 0.64) than in general school (T1–T4: r s = 0.38, 0.42, 0.42, and 0.37).

We also compared the relations of competence and affect across domains. For the competence aspect, we observed that consistently across four-time waves, the strength of the correlations between general school competence and Chinese competence and math competence were much stronger (general school and Chinese: r s = 0.64, 0.65, 0.69, and 0.67; general school and math: r s = 0.58, 0.61, 0.67, and 0.62) than those between Chinese competence and math competence (r s = 0.17, 0.25, 0.30, and 0.31). Likewise, we found stronger associations between general school affect and Chinese affect and general school affect and math affect (general school and Chinese: r s = 0.63, 0.71, 0.71, and 0.68; general school and math: r s = 0.57, 0.59, 0.66, and 0.66) than the association between Chinese affect and math affect (r s = 0.32, 0.42, 0.50, and 0.48).

Comparing the correlations between Chinese and math competence and those between Chinese and math affect, we found that one's liking of Chinese and math are more related than one's perceptions of ability in Chinese and math. Moreover, from a developmental perspective, the correlations also displayed an increase between Chinese and math competence and between Chinese and math affect, suggesting that feeling competent in the two subjects and feeling enjoyment while learning them became more related in grade 8 than in grade 7.

TABLE 1 Goodness of fit of CFA models.

Within-network approach						
Model	Model description	χ^2	df	CFI	TLI	RMSEA
1	One-factor CFA model for Chinese	2927.83	674	0.83	0.81	0.07
2	Two-factor CFA model for Chinese	1374.77	652	0.95	0.94	0.04
3	One-factor CFA model for math	2788.58	674	0.87	0.85	0.07
4	Two-factor CFA model for math	1217.13	652	0.97	0.96	0.04
5	One-factor CFA model for general	3629.01	674	0.78	0.75	0.08
6	Two-factor CFA model for general	1549.45	652	0.93	0.92	0.05
7	Three-factor CFA model for Chinese, math, and general (correlated)	16229.77	6,654	0.82	0.81	0.06
8	Three-factor CFA model for Chinese, math, and general (correlated, elimination of negative worded items)	10360.87	4,158	0.86	0.85	0.05
9	Six-factor CFA model for Chinese, math, and general (correlated)	10865.93	6,444	0.92	0.91	0.03
10	Six-factor CFA model for Chinese, math, and general (correlated, elimination of negative worded items)	6283.48	3,948	0.95	0.94	0.03
Between-network approach						
11	Two-factor CFA model for Chinese + Chinese achievement	1541.41	780	0.95	0.94	0.04
12	Two-factor CFA model for math + math achievement	1513.13	780	0.96	0.95	0.04
13	Integration of Chinese and math achievements into model 10	7159.90	4,524	0.95	0.94	0.03

CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation.

TABLE 2 Descriptive statistics and reliability of scales.

	ChiCOM	ChiAFF	MatCOM	MatAFF	GeCOM	GeAFF
	<i>M (SD)</i> alpha	<i>M (SD)</i> alpha	<i>M (SD)</i> alpha	<i>M (SD)</i> alpha	<i>M (SD)</i> alpha	<i>M (SD)</i> alpha
T1	3.31 (0.70) 0.88	3.75 (0.79) 0.88	3.46 (0.75) 0.87	3.87 (0.79) 0.82	3.33 (0.61) 0.80	3.91 (0.73) 0.77
T2	3.36 (0.72) 0.90	3.90 (0.77) 0.81	3.50 (0.82) 0.92	4.01 (0.77) 0.83	3.34 (0.69) 0.87	3.99 (0.78) 0.82
T3	3.32 (0.73) 0.89	3.85 (0.79) 0.85	3.44 (0.83) 0.92	3.93 (0.82) 0.84	3.31 (0.75) 0.89	3.92 (0.83) 0.84
T4	3.34 (0.74) 0.87	3.65 (0.86) 0.81	3.39 (0.86) 0.91	3.83 (0.87) 0.94	3.21 (0.84) 0.83	3.69 (1.00) 0.83

ChiCOM, Chinese competence scale; ChiAFF, Chinese affect scale; MatCOM, math competence scale; MatAFF, math affect scale; GeCOM, general school competence scale; GeAFF, general school affect scale.

TABLE 3 Results of the measurement invariance tests of model 10.

model	model description	χ^2 (df)	CFI	Δ CFI	TLI	RMSEA	Δ RMSEA
10A	configural	1549.95 (4029)	0.933	---	0.925	0.034	---
10B	metric	7180.903 (4146)	0.932	-0.001	0.925	0.034	0.000
10C	scalar	7485.158 (4200)	0.927	-0.006	0.920	0.035	+0.001

TABLE 4 Factor correlations at T1.

	ChiCOM	ChiAFF	MatCOM	MatAFF	GeCOM	GeAFF	ChiACH
ChiAFF	0.57**	---	---	---	---	---	---
MatCOM	0.17**	0.05	---	---	---	---	---
MatAFF	0.14**	0.32**	0.60**	---	---	---	---
GeCOM	0.54**	0.39**	0.58**	0.39**	---	---	---
GeAFF	0.35**	0.63**	0.28**	0.57**	0.38**	---	---
ChiACH	0.30**	0.10*	0.11*	0.09*	0.27**	0.10*	---
MatACH	-0.04	-0.11*	0.52**	0.33**	0.26**	0.11*	0.40**

ChiCOM, Chinese competence scale; ChiAFF, Chinese affect scale; MatCOM, math competence scale; MatAFF, math affect scale; GeCOM, general school competence scale; GeAFF, general school affect scale; ChiACH, Chinese achievement; MatACH, math achievement. * $p < 0.05$, ** $p < 0.01$.

4.4. The results of the between-network approach to CFAs

Because the within-network approach to CFAs showed better fit results when competence and affect were considered separately, we established our between-network CFAs on the basis of the two-factor ASC. Similar to the within-network approach to CFAs, we also applied the between-network approach to CFAs separately for each domain first. On the basis of model 2, we constructed model 11 by adding Chinese achievement scores in three waves. Similarly, on the basis of model 4, we constructed model 12 by adding math achievement scores in three waves. Finally, on the basis of model 10, we constructed model 13 by adding both Chinese and math achievement scores in four waves. The results showed that the three models all produced good fit: model 11: χ^2 (780) = 1541.41, TLI = 0.95, TLI = 0.94, and RMSEA = 0.04; model 12: χ^2 (780) = 1513.13, TLI = 0.96, TLI = 0.95, and RMSEA = 0.04; and model 13: χ^2 (4524) = 7159.90, TLI = 0.95, TLI = 0.94, and RMSEA = 0.03. These results demonstrated that the stability of the twofold multidimensional structure of ASC was also valid when applying the between-network approach.

The factor correlations between competence/affect and the academic achievement are presented in Tables 3–6, which showed that in both the Chinese and math domains, the correlations between competence and achievement (Chinese T1–T4: r_s = 0.30, 0.41, 0.30 and 0.30; and math T1 to T4: r_s = 0.52, 0.57, 0.43, and 0.45) were higher than those between affect and achievement (Chinese T1–T4: r_s = 0.10, 0.16, 0.09, and 0.16; math T1–T4: r_s = 0.32, 0.30, 0.29, and 0.25). These results indicated that self-evaluation of competence is more closely associated with one's cognitive performance than the self-appraisal of affect in the matching domain.

Across domains, the results showed that the achievement scores had much weaker association with perceptions of competence and affect in non-matching domains. Chinese achievement had either weak or non-significant correlations with math competence and math affect in the four-time waves (Chinese achievement and math competence T1 to T4: r_s = 0.11, 0.20, 0.19, and 0.14; Chinese achievement and math affect T1 to T4: r_s = 0.09, 0.10, 0.10, and 0.07). Similarly, the associations between math achievement and Chinese competence were non-significant for T1 (r = -0.04, p = 0.43) and weak but positive for T2 to T4 (r_s = 0.09, 0.16, and 0.13). The correlations between math achievement and Chinese affect were negative for T1 (r = -0.11) and non-significant for T2 to T4 (r = -0.02, p = 0.73; r = 0.07, p = 0.09; r = 0.06, p = 0.14). The relations between the achievement scores in Chinese and math and general school competence and affect appeared to be weaker compared to those of achievement and competence and affect in the matching domains (Chinese achievement and general school competence T1–T4: r_s = 0.27, 0.36, 0.26, and 0.36; Chinese achievement and general school affect T1–T4: r_s = 0.10, 0.12, 0.08 (p = 0.06), and 0.08 (p = 0.06); math achievement and general school competence T1 to T4: r_s = 0.26, 0.37, 0.34 and 0.33; math achievement and general school affect T1–T4: r_s = 0.11, 0.09, 0.17, and 0.09).

5. Discussion

The present study investigated the stability of the twofold multidimensionality ASC construct in three domains, namely Chinese and math, and general school by adopting both a within-network and a between-network approach from four-wave data among Chinese secondary school students. Our CFA results

TABLE 5 Factor correlations at T2.

	ChiCOM	ChiAFF	MatCOM	MatAFF	GeCOM	GeAFF	ChiACH
ChiAFF	0.62**	---	---	---	---	---	---
MatCOM	0.25**	0.11**	---	---	---	---	---
MatAFF	0.23**	0.42**	0.63**	---	---	---	---
GeCOM	0.65**	0.39**	0.61**	0.44**	---	---	---
GeAFF	0.46**	0.71**	0.29**	0.59**	0.42**	---	---
ChiACH	0.41**	0.16**	0.20**	0.10*	0.36**	0.12**	---
MatACH	0.09*	−0.02	0.57**	0.30**	0.37**	0.09*	0.60**

ChiCOM, Chinese competence scale; ChiAFF, Chinese affect scale; MatCOM, math competence scale; MatAFF, math affect scale; GeCOM, general school competence scale; GeAFF, general school affect scale; ChiACH, Chinese achievement; MatACH, math achievement. * $p < 0.05$, ** $p < 0.01$.

TABLE 6 Factor correlations at T3.

	ChiCOM	ChiAFF	MatCOM	MatAFF	GeCOM	GeAFF	ChiACH
ChiAFF	0.59**	---	---	---	---	---	---
MatCOM	0.30**	0.16**	---	---	---	---	---
MatAFF	0.19**	0.50**	0.56**	---	---	---	---
GeCOM	0.69**	0.42**	0.67**	0.41**	---	---	---
GeAFF	0.38**	0.71**	0.33**	0.66**	0.42**	---	---
ChiACH	0.30**	0.09*	0.19**	0.10*	0.26**	0.08	---
MatACH	0.16**	0.07	0.43**	0.29**	0.34**	0.17**	0.62**

ChiCOM, Chinese competence scale; ChiAFF, Chinese affect scale; MatCOM, math competence scale; MatAFF, math affect scale; GeCOM, general school competence scale; GeAFF, general school affect scale; ChiACH, Chinese achievement; MatACH, math achievement. * $p < 0.05$, ** $p < 0.01$.

showed that whether the three domains were considered individually or together, the one-factor models (1, 3, 5, 7, and 8) failed to produce appropriate fit, whereas the two-factor models (2, 4, 6, 9, and 10) yielded superior fit, supporting the assumption that competence and affect should be conceptualized as two separate constructs among Chinese secondary school students.

5.1. The within-network evidence of the stability of the twofold multidimensional structure of ASC

Consistently with our first hypotheses, our study found that the stability of the twofold multidimensional structure of ASC was also valid with Chinese secondary school students. We found that both of the CFAs (model 7 and 8) which conflated competence and affect in the three academic domains generated non-fit, even when the negative items were eliminated (model 8), whereas both of the CFAs (model 9 and 10) which examined the stability of the validity of the twofold structure of ASC in the three domains produced good fit, even when the negative items were included (model 9). These results were consistent with the longitudinal study by Marsh et al. (1999b) with Australian primary school students. More recently, Leung (2019) also found strong support for the twofold multidimensional structure of the ASC among Chinese students, despite that Leung's study was cross-sectional and used SDQ II

(appropriate for the high school students). Our study together with studies by Leung and Yang et al. (2016) seemed to indicate that the twofold multidimensional structure of ASC was applicable for Chinese students who were above primary school levels.

Similar to the findings reported in previous cross-sectional studies (Arens et al., 2011; Abu-Hilal et al., 2013; Leung, 2019), we also observed that the competence scales were positively related to the affect scales within the Chinese, math, and general school domains, and such patterns were stable across the four-time points. Also aligned with previous cross-sectional studies (Arens et al., 2011; Yang et al., 2016; Schneider and Sparfeldt, 2020), we observed that the correlations between competence and affect in Chinese (r s ranging from 0.56 to 0.62) and math (r s ranging from 0.56 to 0.64) were consistently higher than those in general school (r s ranging from 0.37 to 0.42). These results may indicate that the relations between self-perceptions of cognitive ability and affective status in specific subjects are stronger than the relations between competence and affect in school learning in general and support the domain-specific nature of ASC.

The cross-domain correlations of the competence and affect components found in our study somewhat contradict the domain specificity of both the competence and affect components in Arens et al. (2011), as the cross-domain correlations of both the competence and affect components were negligible in their study. Our study only showed the domain specificity of competence components. The cross-domain relations between the affect

TABLE 7 Factor correlations at T4.

	ChiCOM	ChiAFF	MatCOM	MatAFF	GeCOM	GeAFF	ChiACH
ChiAFF	0.56**	---	---	---	---	---	---
MatCOM	0.31**	0.19**	---	---	---	---	---
MatAFF	0.27**	0.48**	0.64**	---	---	---	---
GeCOM	0.67**	0.37**	0.62**	0.40**	---	---	---
GeAFF	0.40**	0.68**	0.36**	0.66**	0.37**	---	---
ChiACH	0.36**	0.16**	0.14**	0.07	0.30**	0.08	---
MatACH	0.13**	0.06	0.45**	0.25**	0.34**	0.09*	0.57**

ChiCOM, Chinese competence scale; ChiAFF, Chinese affect scale; MatCOM, math competence scale; MatAFF, math affect scale; GeCOM, general school competence scale; GeAFF, general school affect scale; ChiACH, Chinese achievement; MatACH, math achievement. * $p < 0.05$, ** $p < 0.01$.

components were moderate (r s ranging from 0.32 to 0.50), suggesting that students' liking of the two dissimilar school subjects (Chinese and math) was not as pronounced as their own evaluation of their competence in these two subjects. One possible explanation could be that Chinese and math subjects, along with English, are the most highly valued subjects in the Chinese secondary school system, as the examination scores in these three subjects are given more weights than other subjects in the High School Entrance Examination. The affect developed in these two subjects, therefore, might not be based solely on students' personal interests, but on the importance students attached to them.

5.2. The between-network evidence of the stability of the twofold multidimensional structure of ASC

Confirming the second hypothesis, our results showed consistently across the four-time waves that self-perceptions of competence were more strongly associated with achievement than self-perceptions of affect within both the Chinese and math domains. These results corroborated the findings of previous cross-sectional studies (Arens et al., 2011; Abu-Hilal et al., 2013; Lohbeck, 2019; Schneider and Sparfeldt, 2020) and longitudinal studies in a single domain (Pinxten et al., 2014; Arens et al., 2016; Han, 2019). The results also supported the assumptions made by Irwing (1996) that evaluation of one's competence involves more external (social) comparison, whereas evaluation of one's affect involves more internal (dimensional) comparison.

In our study across the Chinese and math domains, the relations between competence and achievement were much stronger for matching domains than those for non-matching domains in all the four waves, confirming the domain-specific nature of academic self-concept and the I/E model and the dimensional comparison model of self-concept (Marsh, 1986; Möller and Marsh, 2013). The magnitude of correlations between affect and achievement was also stronger within one domain than across domains for both Chinese and math learning except for wave three where we observed that the correlation between Chinese affect and Chinese achievement ($r = 0.09$) was slightly

weaker than math affect and Chinese achievement ($r = 0.10$). Comparing the associations between achievement scores and affect in both matching and non-matching domains, the features of domain specificity which were normally found in the relations between achievement and competence were also identified in the relations between achievement and affect.

Stronger relations between competence and achievement than between affect and achievement were also observed between the Chinese and math achievement scores and the two components of ASC in general school. We found that for all four-time waves, Chinese and math achievement scores had stronger correlations with general school competence than with general school affect. Furthermore, the associations between general school competence and achievement scores in Chinese were similar to those between general school competence and achievement scores in math; this was also the case for the associations between general school affect and achievement scores in Chinese and math (except for wave three). These relational patterns between domain-specific achievements and general school competence and affect corroborated the findings in cross-sectional studies (Arens et al., 2011) and provided longitudinal evidence of the claim postulated in the Marsh/Shavelson model (Marsh, 1990d) that general school ASC incorporates both the verbal and math domains.

Although the relations of both components of ASC with achievements in non-matching domains were weaker than those found in the corresponding domains, they were neither zero nor negative at most time points, which did not seem to align with the predictions of the I/E model, and somewhat contradicted with our third hypothesis. A number of cross-sectional studies with younger learners (primary school students) also reported no zero or negative cross-domain relations between ASCs and achievements, which were inconsistent with the predictions of the I/E model (both competence and affect: Arens et al., 2011; Lohbeck, 2019; only competence: Ehm et al., 2014; Lohbeck and Möller, 2017). For example, both Lohbeck and Möller (2017) and Ehm et al. (2014) studies found positive and significant cross-domain relations in the sample of German 1st to 3rd graders. A number of possible explanations for such results were offered by these researchers, such as younger children's less developed cognitive abilities with regard to dimensional comparison (Harter, 1999); younger learners were

not as familiar with the grading system as they were at the beginning of their schooling (Lohbeck and Möller, 2017) or both contrast and assimilation processes were in operation (Möller and Marsh, 2013), in particular if the domains were less distinct (Ehm et al., 2014).

These explanations, however, might not be applicable to our study, as our participants were secondary students, and the verbal and math subjects were significantly different. One of the possible explanations is that our participants might engage in social comparison more than dimensional comparison, which might be attributable to three reasons: (1) the Chinese culture; (2) the developmental stage of the participants; and (3) the educational context of the participants. For the first reason, the Chinese culture is a collectivistic culture (Ho and Chiu, 1994; White and Lehman, 2005), and research has shown that people with higher level of collectivism tend to have a higher desire for social comparison (Chung and Mallery, 1999). White and Lehman (2005) argue that “those from Eastern and Western cultural contexts differ in the extent to which they construe the self as interconnected with or distinct from those around them.” As the proposal of the I/E model is largely based on earlier evidence garnered primarily from Western contexts, it may not be applicable to the Chinese culture. However, to confirm this conjecture, more empirical research evidence is required. Secondly, from the developmental perspective, our participants were in their secondary school students, who were in a period of re-orientating to the new learning contexts, being more sensitive to the social contexts and peer influence, and a significant increase in social interactions (Brown and Larson, 2009; Maehr and Zusho, 2009; Molloy et al., 2011; Crone and Dahl, 2012). Hence, our participants might have heightened engagement in social comparison. Last but not least, the predominance of social comparison over dimensional comparison might also be related to the Chinese educational context. In Chinese schools, academic social comparisons are a common phenomenon, in particular among Chinese adolescents (Fu et al., 2018; Li et al., 2021). In such an educational context, our participants might use more social comparison than dimensional comparison to rate their self-concepts in Chinese and math.

Other possible explanations of the unexpected results of the cross-domain correlations between both components of ASC and achievements could be some mediators, which had not been examined in our study. Recent studies have identified a range of factors could mediate between ASC and achievement, such as parental expectations (Phillipson and Phillipson, 2017), maternal control (Lu et al., 2017), and cultural norms (Wentzel et al., 2021; Vu et al., 2022). Hence, in order to fully understand the correlations between ASCs and achievements in non-matching domains, future studies should be designed by incorporating possible mediators. Despite these partial inconsistencies concerning the predictions of the I/E model, in general, our findings regarding the between-network approach also provided support for the validity of the twofold multidimensional nature of ASC in these groups of students across time.

5.3. Limitations and suggestions for future research

Some limitations of the study need to be pointed out so that they can be addressed in future studies. First, our sample size was not large considering that our study was longitudinal and involved many constructs. Moreover, the participants were recruited from the same secondary school. Although there were advantages of recruiting the students from the same school, such as their academic achievement results being comparable as the students sat the same examinations, this limited the generalization of the study results. Therefore, future studies should have a large sample size and recruit participants from different schools so that the sample would be more representative, which will provide stronger support for the study results. Second, our study only demonstrated the stability of the twofold multidimensional structure of ASC with Chinese secondary school students, future research should extend such longitudinal investigation into Chinese students of other age groups and students in other cultures. Another limitation is that our study did not include other important factors in learning, such as learning engagement (Henning et al., 2022), academic enjoyment (Liu et al., 2022), study interest (Fryer and Ainley, 2019), and achievement emotions (Forsblom et al., 2022). Including these important elements may help reveal complex relations (e.g., mediating relations) between ASC, learning outcomes, and these factors.

5.4. Conclusion

The current study examined the stability of the twofold multidimensional structure of ASC with Chinese secondary school students. It filled the gap in the literature where there was a lack of research examining the separation of competence and affect in multiple domains using both a within-network and a between-network approach. While the majority of past studies on the validity of the twofold multidimensional structure of ASC were conducted with learners from Western and Arab cultures, our study provided additional evidence in this area by including a less researched population – Chinese learners. Our results demonstrate that both the multidimensional and twofold nature of academic self-concept obtained from cross-sectional studies with Western samples is applicable to our sample of Chinese secondary students.

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 Human Research Ethics Committee of the University of Sydney. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

FH contributed substantially to the conception of the work, the acquisition, analysis, and interpretation of the data. FH, KJ, PM, and LN drafted the work and revised it critically for important intellectual content, approved the final version of the paper to be published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

This study was supported by the International mobilities for research activities of the University of Hradec Králové II, CZ.02.2.69/0.0/0.0/18_053/0017841; Faculty of Education, the

University of Hradec Králové as part of the projects: Individual and contextual predictors of the effect of supporting academic autonomy in future teachers (number: 2101); the Competition for Postdoctoral Job Positions (number: 2201); and Excessive use of social networks and emotional intelligence (number: 2110).

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

- Abu-Hilal, M. M., Abdelfattah, F. A., Alshumrani, S. A., Abduljabbar, A. S., and Marsh, H. W. (2013). Construct validity of self-concept in TIMSS's student background questionnaire: a test of separation and conflation of cognitive and affective dimensions of self-concept among Saudi eighth graders. *Eur. J. Psychol. Educ.* 28, 1201–1220. doi: 10.1007/s10212-012-0162-1
- Arens, A. K., Bodkin-Andrews, G., Craven, R. G., and Yeung, A. S. (2014a). Self-concept of indigenous and non-indigenous Australian students: competence and affect components and relations to achievement. *Learn. Individ. Differ.* 32, 93–103. doi: 10.1016/j.lindif.2014.03.019
- Arens, A. K., and Hasselhorn, M. (2015). Differentiation of competence and affect self-perceptions in elementary school students: extending empirical evidence. *Eur. J. Psychol. Educ.* 30, 405–419. doi: 10.1007/s10212-015-0247-8
- Arens, A. K., Marsh, H. W., Craven, R. G., Yeung, A. S., Randhawa, E., and Hasselhorn, M. (2016). Math self-concept in preschool children: structure, achievement relations, and generalizability across gender. *Early Child. Res. Q.* 36, 391–403. doi: 10.1016/j.ecresq.2015.12.024
- Arens, A. K., and Morin, A. J. S. (2016). Examination of the structure and grade-related differentiation of multidimensional self-concept instruments for children using ESEM. *J. Exp. Educ.* 84, 330–355. doi: 10.1080/00220973.2014.999187
- Arens, A. K., and Preckel, F. (2018). Testing the internal/external frame of reference model with elementary school children: extension to physical ability and intrinsic value. *Contemp. Educ. Psychol.* 54, 199–211. doi: 10.1016/j.cedpsych.2018.06.003
- Arens, A. K., Yeung, A. S., Craven, R. G., and Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: domain specificity and separation between competence and affect components. *J. Educ. Psychol.* 103, 970–981. doi: 10.1037/a0025047
- Arens, A. K., Yeung, A. S., and Hasselhorn, M. (2014b). Native language self-concept and reading self-concept: same or different? *J. Exp. Educ.* 82, 229–252. doi: 10.1080/00220973.2013.813362
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull.* 107, 238–246. doi: 10.1037/0033-2909.107.2.238
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. 2nd Edn. New York, NY, USA: The Guilford Press.
- Brown, B. B., and Larson, J. (2009). "Peer relationships in adolescence," in *Handbook of Adolescent Psychology*. (Hoboken, NY, USA: John Wiley & Sons, Ltd)
- Browne, M. W., and Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociol. Methods Res.* 21, 230–258. doi: 10.1177/0049124192021002005
- Burns, R. A., Crisp, D. A., and Burns, R. B. (2018). Competence and affect dimensions of self-concept among higher education students: a factorial validation study of an academic subject-specific self-concept. *Eur. J. Psychol. Educ.* 33, 649–663. doi: 10.1007/s10212-018-0369-x
- Byrne, B. M. (1984). The general/academic self-concept nomological network: a review of construct validation research. *Rev. Educ. Res.* 54, 427–456. doi: 10.2307/1170455
- Byrne, B. M. (1996). "Academic self-concept: its structure, measurement, and relation to academic achievement," in *Handbook of Self-Concept: Developmental, Social, and Clinical Considerations*. ed. B. A. Bracken (Oxford, England: John Wiley & Sons), 287–316.
- 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
- Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5
- Chevalère, J., Cazenave, L., Wollast, R., Berthon, M., Martinez, R., Mazenod, V., et al. (2022). The influence of socioeconomic status, working memory and academic self-concept on academic achievement. *Eur. J. Psychol. Educ.* doi: 10.1007/s10212-022-00599-9
- Chiu, M. M., and Klassen, R. M. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: cultural differences among fifteen-year-olds in 34 countries. *Learn. Instr.* 20, 2–17. doi: 10.1016/j.learninstruc.2008.11.002
- Chung, T., and Mallory, P. (1999). Social comparison, individualism-collectivism, and self-esteem in China and the United States. *Curr. Psychol.* 18, 340–352. doi: 10.1007/s12144-999-1008-0
- Coopersmith, S. (1960). Self-esteem and need achievement as determinants of selective recall and repetition. *J. Abnorm. Soc. Psychol.* 60, 310–317. doi: 10.1037/h0042888
- Crone, E. A., and Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nat. Rev. Neurosci.* 13, 636–650. doi: 10.1038/nrn3313
- Deeba, F. D., Ullah, S., and Saleem, A. (2022). An analysis of students' academic self-concept in English as predictor of their academic performance. *Pak. J. Hum. Soc. Sci.* 10, 403–415. doi: 10.52131/pjhss.2022.1001.0209

- Dodgson, P. G., and Wood, J. V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. *J. Pers. Soc. Psychol.* 75, 178–197. doi: 10.1037/0022-3514.75.1.178
- Ehm, J.-H., Lindberg, S., and Hasselhorn, M. (2014). Reading, writing, and math self-concept in elementary school children: influence of dimensional comparison processes. *Eur. J. Psychol. Educ.* 29, 277–294. doi: 10.1007/s10212-013-0198-x
- Forsblom, L., Pekrun, R., Loderer, K., and Peixoto, F. (2022). Cognitive appraisals, achievement emotions, and students' math achievement: a longitudinal analysis. *J. Educ. Psychol.* 114, 346–367. doi: 10.1037/edu0000671
- Fryer, L. K., and Ainley, M. (2019). Supporting interest in a study domain: a longitudinal test of the interplay between interest, utility-value, and competence beliefs. *Learn. Instr.* 60, 252–262. doi: 10.1016/j.learninstruc.2017.11.002
- Fu, R., Chen, X., Liu, J., and Li, D. (2018). Relations between social comparison orientation and adjustment in Chinese adolescents: moderating effects of initial adjustment status. *Int. J. Psychol.* 53, 133–141. doi: 10.1002/ijop.12278
- Gorges, J., and Hollmann, J. (2019). The structure of academic self-concept when facing novel learning content: multidimensionality, hierarchy, and change. *Eur. J. Psychol.* 15, 491–508. doi: 10.5964/ejop.v15i3.1716
- Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annu. Rev. Psychol.* 60, 549–576. doi: 10.1146/annurev.psych.58.110405.085530
- Guay, F., Stupnisky, R., Boivin, M., Japel, C., and Dionne, G. (2019). Teachers' relatedness with students as a predictor of students' intrinsic motivation, self-concept, and reading achievement. *Early Child. Res. Q.* 48, 215–225. doi: 10.1016/j.ecresq.2019.03.005
- Guo, J., Marsh, H. W., Morin, A. J. S., Parker, P. D., and Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: a longitudinal study. *Am. Educ. Res. J.* 52, 371–402. doi: 10.3102/0002831214565786
- Han, F. (2019). Longitudinal relations between school self-concept and academic achievement. *Rev. Psicodidáctica (English ed.)* 24, 95–102. doi: 10.1016/j.pscico.2019.03.002
- Harter, S. (1990). "Processes underlying adolescent self-concept formation," in *From Childhood to Adolescence: A Transitional Period? Advances in Adolescent Development: An Annual Book Series*. Vol. 2. eds. R. Montemayor, G. R. Adams, and T. P. Gullotta (Thousand Oaks, CA, USA: Sage Publications, Inc.), 205–239.
- Harter, S. (1999). *The Construction of the Self: A Developmental Perspective*. New York, NY, USA: Guilford Press.
- Henning, L., Tietjens, M., and Dreiskämper, D. (2022). The multidimensionality of the physical fitness self-concept: a recommendation to consider competence and affect components in childhood. *Sport Exerc. Perform. Psychol.* 11, 79–92. doi: 10.1037/spy0000281
- Ho, D. Y.-F., and Chiu, C.-Y. (1994). "Component ideas of individualism, collectivism, and social organization: an application in the study of Chinese culture," in *Individualism and Collectivism: Theory, Method, and Applications, Cross-Cultural Research and Methodology Series*, Vol. 18. eds. U. Kim, H. C. Triandis, Ç. Kâitçibai, S.-C. Choi, and G. Yoon (Thousand Oaks, CA, USA: Sage Publications, Inc.), 137–156.
- 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
- Irwing, P. (1996). Evaluative and affective dimensions of self-concept: a test of construct validity using structural equations modeling. *Psychol. Rep.* 79, 1127–1138. doi: 10.2466/pr0.1996.79.3f.1127
- Joreskog, K., and Sorbom, D. (2005). LISREL 8.72. Chicago, IL, USA: Scientific Software International.
- Kadir, M. S., Yeung, A. S., and Diallo, T. M. O. (2017). Simultaneous testing of four decades of academic self-concept models. *Contemp. Educ. Psychol.* 51, 429–446. doi: 10.1016/j.cedpsych.2017.09.008
- Karimova, K., and Csapó, B. (2021). Cognitive and affective components of verbal self-concepts and internal/external frame of reference within the multidimensional verbal domain. *SAGE Open* 11:215824402110186. doi: 10.1177/21582440211018683
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. New York, NY, USA: Guilford Press.
- Leung, K. C. (2019). Competence-affect separation on the math and verbal scales of the self-description questionnaire II in Chinese sample. *J. Psychoeduc. Assess.* 37, 119–124. doi: 10.1177/0734282917723481
- Li, J., Zhang, N., Yao, M., Xing, H., and Liu, H. (2021). Academic social comparison and depression in Chinese adolescents: the mediating role of basic psychological needs satisfaction. *Sch. Ment. Heal.* 13, 719–729. doi: 10.1007/s12310-021-09436-8
- Liu, Y., Wang, Y., Liu, R.-D., Ding, Y., Wang, J., and Mu, X. (2022). How classroom environment influences academic enjoyment in mathematics among Chinese middle school students: moderated mediation effect of academic self-concept and academic achievement. *Psychol. Res. Behav. Manag.* 15, 2035–2048. doi: 10.2147/PRBM.S371092
- Lohbeck, A. (2019). Social and dimensional comparison effects on academic self-concepts and self-perceptions of effort in elementary school children. *Educ. Psychol.* 39, 133–150. doi: 10.1080/01443410.2018.1527018
- Lohbeck, A., and Möller, J. (2017). Social and dimensional comparison effects on math and reading self-concepts of elementary school children. *Learn. Individ. Differ.* 54, 73–81. doi: 10.1016/j.lindif.2017.01.013
- Lu, M., Walsh, K., White, S., and Shield, P. (2017). The associations between perceived maternal psychological control and academic performance and academic self-concept in Chinese adolescents: the mediating role of basic psychological needs. *J. Child Fam. Stud.* 26, 1285–1297. doi: 10.1007/s10826-016-0651-y
- Maehr, M. L., and Zusho, A. (2009). "Achievement goal theory: the past, present, and future," in *Handbook of Motivation at School Educational Psychology Handbook Series*. eds. K. R. Wentzel and D. B. Miele (New York, NY, USA: Routledge/Taylor & Francis Group), 77–104.
- Marsh, H. W. (1986). Verbal and math self-concepts: an internal/external frame of reference model. *Am. Educ. Res. J.* 23, 129–149. doi: 10.3102/00028312023001129
- Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: theoretical and empirical justification. *Educ. Psychol. Rev.* 2, 77–172. doi: 10.1007/BF01322177
- Marsh, H. W. (1990b). Causal ordering of academic self-concept and academic achievement: a multiwave, longitudinal panel analysis. *J. Educ. Psychol.* 82, 646–656. doi: 10.1037/0022-0663.82.4.646
- Marsh, H. W. (1990c). Confirmatory factor analysis of multitrait-multimethod data: the construct validation of multidimensional self-concept responses. *J. Pers.* 58, 661–692. doi: 10.1111/j.1467-6494.1990.tb00249.x
- Marsh, H. W. (1990d). The structure of academic self-concept: the Marsh/Shavelson model. *J. Educ. Psychol.* 82, 623–636. doi: 10.1037/0022-0663.82.4.623
- Marsh, H. W. (2014). "Academic self-concept: theory, measurement, and research" in *Psychological Perspectives on the Self*, vol. 4. ed. J. Suls (New York, NY, USA: Psychology Press), 71–110.
- Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., et al. (2013). Factorial, convergent, and discriminant validity of timss math and science motivation measures: a comparison of Arab and Anglo-Saxon countries. *J. Educ. Psychol.* 105, 108–128. doi: 10.1037/a0029907
- Marsh, H. W., and Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *J. Educ. Psychol.* 95, 687–706. doi: 10.1037/0022-0663.95.4.687
- Marsh, H. W., Byrne, B. M., and Shavelson, R. J. (1988). A multifaceted academic self-concept: its hierarchical structure and its relation to academic achievement. *J. Educ. Psychol.* 80, 366–380. doi: 10.1037/0022-0663.80.3.366
- Marsh, H. W., Byrne, B. M., and Yeung, A. S. (1999a). Causal ordering of academic self-concept and achievement: reanalysis of a pioneering study and. *Educ. Psychol.* 34, 155–167. doi: 10.1207/s15326985ep3403_2
- Marsh, H. W., and Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1, 133–163. doi: 10.1111/j.1745-6916.2006.00010.x
- Marsh, H. W., Craven, R., and Debus, R. (1999b). Separation of competency and affect components of multiple dimensions of academic self-concept: a developmental perspective. *Merrill Palmer Q* 45, 567–601.
- Marsh, H. W., and Martin, A. J. (2011). Academic self-concept and academic achievement: relations and causal ordering. *Br. J. Educ. Psychol.* 81, 59–77. doi: 10.1348/000709910X503501
- Marsh, H. W., Martin, A. J., Yeung, A. S., and Craven, R. G. (2017). "Competence self-perceptions," in *Handbook of Competence and Motivation: Theory and Application*. 2nd Edn. eds. A. J. Elliot, C. S. Dweck, and D. S. Yeager (New York, NY, USA: The Guilford Press), 85–115.
- Marsh, H. W., and O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. *Personal. Soc. Psychol. Bull.* 34, 542–552. doi: 10.1177/0146167207312313
- Marsh, H. W., and O'Neill, R. (1984). Self description questionnaire III: the construct validity of multidimensional self-concept ratings by late adolescents. *J. Educ. Meas.* 21, 153–174. doi: 10.1111/j.1745-3984.1984.tb00227.x
- Marsh, H. W., Xu, M., and Martin, A. J. (2012). "Self-concept: a synergy of theory, method, and application," in *APA Educational Psychology Handbook, Vol 1: Theories, Constructs, and Critical Issues APA Handbooks in Psychology*. eds. K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, and J. Sweller (Washington, DC, USA: American Psychological Association), 427–458.

- Marx, R. W., and Winne, P. H. (1978). Construct interpretations of three self-concept inventories. *Am. Educ. Res. J.* 15, 99–109. doi: 10.3102/00028312015001099
- Maynor, L., Gálvez-Peralta, M., Barrickman, A., Hanif, A., and Baugh, G. (2022). Perceived stress, academic self-concept, and coping mechanisms among pharmacy students following a curricular revision. *Curr. Pharm. Teach. Learn.* 14, 159–165. doi: 10.1016/j.cptl.2021.11.013
- Melguizo-Ibáñez, E., Zurita-Ortega, F., Ubago-Jiménez, J. L., López-Gutiérrez, C. J., and González-Valero, G. (2022). An explanatory model of the relationships between sport motivation, anxiety and physical and social self-concept in educational sciences students. *Curr. Psychol.* doi: 10.1007/s12144-022-02778-9
- Möller, J., and Marsh, H. W. (2013). Dimensional comparison theory. *Psychol. Rev.* 120, 544–560. doi: 10.1037/a0032459
- Möller, J., Pohlmann, B., Köller, O., and Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Rev. Educ. Res.* 79, 1129–1167. doi: 10.3102/0034654309337522
- Möller, J., Retelsdorf, J., Köller, O., and Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: an integration of models of relations between academic achievement and self-concept. *Am. Educ. Res. J.* 48, 1315–1346. doi: 10.3102/00028312111419649
- Molloy, L. E., Gest, S. D., and Rulison, K. L. (2011). Peer influences on academic motivation: exploring multiple methods of assessing youths' Most "influential" peer relationships. *J. Early Adolesc.* 31, 13–40. doi: 10.1177/0272431610384487
- Onetti-Onetti, W., Chinchilla-Minguet, J. L., Martins, F. M. L., and Castillo-Rodríguez, A. (2019). Self-concept and physical activity: differences between high school and university students in Spain and Portugal. *Front. Psychol.* 10:1333. doi: 10.3389/fpsyg.2019.01333
- Phillipson, S., and Phillipson, S. N. (2017). Generalizability in the mediation effects of parental expectations on Children's cognitive ability and self-concept. *J. Child Fam. Stud.* 26, 3388–3400. doi: 10.1007/s10826-017-0836-z
- Pinxten, M., De Fraine, B., Van Damme, J., and D'Haenens, E. (2013). Student achievement and academic self-concept among secondary students in Flanders: gender and changes over time. *Irish Educ. Stud.* 32, 157–178. doi: 10.1080/03323315.2012.749058
- Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., and Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *Br. J. Educ. Psychol.* 84, 152–174. doi: 10.1111/bjep.12028
- Rosenberg, M. (1979). *Conceiving the self*. New York, NY: Basic books.
- Schneider, R., and Sparfeldt, J. R. (2020). Academic competence and affect self-concepts in elementary school students: social and dimensional comparisons. *Soc. Psychol. Educ.* 23, 233–257. doi: 10.1007/s11218-019-09532-3
- Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Rev. Educ. Res.* 46, 407–441. doi: 10.2307/1170010
- Simonsmeier, B. A., Peiffer, H., Flaig, M., and Schneider, M. (2020). Peer feedback improves students' academic self-concept in higher education. *Res. High. Educ.* 61, 706–724. doi: 10.1007/s11162-020-09591-y
- Smith, K. H. (2019). The multi-dimensionality of academic self-concept. *Educ. Pract. Theory* 41, 71–81. doi: 10.7459/ept/41.1.05
- Sommer, K. L., and Baumeister, R. F. (2002). Self-evaluation, persistence, and performance following implicit rejection: the role of trait self-esteem. *Personal. Soc. Psychol. Bull.* 28, 926–938. doi: 10.1177/014616720208007006
- Tucker, L. R., and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38, 1–10. doi: 10.1007/BF02291170
- Vu, T., Magis-Weinberg, L., Jansen, B. R. J., van Atteveldt, N., Janssen, T. W. P., Lee, N. C., et al. (2022). Motivation-achievement cycles in learning: a literature review and research agenda. *Educ. Psychol. Rev.* 34, 39–71. doi: 10.1007/s10648-021-09616-7
- Watkins, D., Dong, Q., and Xia, Y. (1997). Age and gender differences in the self-esteem of Chinese children. *J. Soc. Psychol.* 137, 374–379. doi: 10.1080/00224549709595448
- Wentzel, K. R., Jablansky, S., and Scalise, N. R. (2021). Peer social acceptance and academic achievement: a meta-analytic study. *J. Educ. Psychol.* 113, 157–180. doi: 10.1037/edu0000468
- White, K., and Lehman, D. R. (2005). Culture and social comparison seeking: the role of self-motives. *Personal. Soc. Psychol. Bull.* 31, 232–242. doi: 10.1177/0146167204271326
- Yang, L., Arens, A. K., and Watkins, D. A. (2016). Testing the twofold multidimensionality of academic self-concept: a study with Chinese vocational students. *Educ. Psychol.* 36, 1651–1669. doi: 10.1080/01443410.2014.995597
- Yeung, A. S., Craven, R. G., and Kaur, G. (2012). Mastery goal, value and self-concept: what do they predict? *Educ. Res.* 54, 469–482. doi: 10.1080/00131881.2012.734728



OPEN ACCESS

EDITED BY

Claudio Longobardi,
University of Turin, Italy

REVIEWED BY

Louis S. Nadelson,
University of Central Arkansas, United States
Majid Murad,
Jiangsu University, China

*CORRESPONDENCE

Junfeng Zhang
✉ junfeng.zhang@nju.edu.cn

SPECIALTY SECTION

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

RECEIVED 27 May 2022

ACCEPTED 18 January 2023

PUBLISHED 20 February 2023

CITATION

Sun J, Shi J and Zhang J (2023) From
entrepreneurship education to entrepreneurial
intention: Mindset, motivation, and prior
exposure. *Front. Psychol.* 14:954118.
doi: 10.3389/fpsyg.2023.954118

COPYRIGHT

© 2023 Sun, Shi 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 entrepreneurship education to entrepreneurial intention: Mindset, motivation, and prior exposure

Junhua Sun¹, Jingyi Shi² and Junfeng Zhang^{1*}

¹School of Education, Nanjing University, Nanjing, China, ²Department of Mathematics and Statistics, Mississippi State University, Mississippi State, MS, United States

We studied the relationships between entrepreneurship education and entrepreneurial intention among college students, with a focus on the mediating role of an entrepreneurial mindset as well as the moderating roles of learning motivation and prior entrepreneurial exposure. More than 90,000 students from 100 colleges or universities participated in the investigation, the data were subjected to structural equation modeling with Mplus. The results indicated that entrepreneurship education (curriculum attendance and extracurricular activity) significantly enhanced the entrepreneurial mindset of students, which, in turn, strengthened their entrepreneurial intention. In terms of learning, intrinsic motivation positively moderated the relationships between curriculum attendance and entrepreneurial intention/mindset, whereas extrinsic motivation moderated it negatively. Entrepreneurial exposure positively moderated the correlation between extracurricular activity and academic performance. Implications concerning the adjustment of entrepreneurship education to the entrepreneurial climate are discussed.

KEYWORDS

entrepreneurship education, entrepreneurial intention, mindset, learning motivation, entrepreneurial exposure

Introduction

One of the highly debated topics in higher education that has attracted increasing attention in recent years relates to entrepreneurship education for college students (Santos et al., 2019; Cui et al., 2021; Wang et al., 2021). It has been of serious concern to administrators and course developers, researchers, and policymakers (Rauch and Hulsink, 2015), given the increasing importance of entrepreneurship in generating innovation and fueling economic growth (Sutter et al., 2019). Education in this field has the potential to boost the learning of college students and promote the acquisition of entrepreneurial knowledge, skills, and behavior (Cui et al., 2021), thereby enabling them to reach a high level of entrepreneurship (Jack and Anderson, 1999).

The aim of including entrepreneurship within higher education is to promote entrepreneurial intent and behavior among college students, and the entrepreneurial mindset has been deemed potential in terms of correlation (Solevik et al., 2013; Pfeifer et al., 2016; Nabi et al., 2018; Cui et al., 2021). Rather than reflecting the *status quo*, it has been conceived of as a frame of mind for approaching problems, implementing innovations, finding solutions, sharing ideas, and making change happen, based on a spectrum associated with business ownership (Nadelson et al., 2018). It could be shaped through entrepreneurship education, not only in reflecting the thinking of entrepreneurs but also in enabling others to think and act like entrepreneurs. However, research on any link between entrepreneurship education and an entrepreneurial mindset, as well as on

the role of the mindset in predicting entrepreneurial intention or behavior, is in the early stages (Krueger, 2015; Cui et al., 2021).

A meta-analysis carried out by Bae et al. (2014) revealed a significant but small correlation between entrepreneurship education and entrepreneurial intention, indicating the need to control for other variables that might affect the effectiveness of the education. Although learning motivation has been shown to affect the experiences of college students (National Survey of Student Engagement, NSSE, see Kuh, 2001), exactly how it benefits from entrepreneurship learning remains unclear (Hytti et al., 2010). In addition to curriculum attendance and extracurricular activity, prior entrepreneurial exposure has also been shown to boost entrepreneurial intention (Chlosta et al., 2012), such as by changing attitudes (Krueger et al., 2000; Zapkau et al., 2015). It would, therefore, be useful to assess learning motivation and entrepreneurial exposure as contextual factors of entrepreneurship education.

The following aspects are addressed in the current study. First, we investigate the impact of entrepreneurship education on the entrepreneurial mindset and the entrepreneurial intentions of college students. Despite the increasing number of empirical studies exploring entrepreneurship education, the entrepreneurial mindset, and entrepreneurial intention, a gap remains concerning how entrepreneurship education affects the two latter simultaneously. Thus far, results appear to be heterogeneous (Cui et al., 2021), thereby leading to a lack of generality (Wang et al., 2021). Second, we examine the role of an entrepreneurial mindset in mediating the relationship between entrepreneurship education and entrepreneurial intention. Given the strengthening focus on the entrepreneurial mindset in validating entrepreneurship (Daspit et al., 2021), one should naturally acknowledge its power in explaining the mechanism by which entrepreneurship education affects entrepreneurial behaviors. Finally, we explore the role of learning motivation and entrepreneurial exposure in moderating the impact of entrepreneurship education on mindset and intention.

Theoretical foundation and research hypotheses

Entrepreneurship education and entrepreneurial intentions

Entrepreneurship education aims to develop students' entrepreneurial intentions (Li and Wu, 2019). As the optimal predictor of entrepreneurial behavior (Krueger et al., 2000), entrepreneurial intention has been highlighted in investigations of its relationship with entrepreneurship education (Zhang and Huang, 2021). Empirical studies have identified the mechanisms by means of which entrepreneurship education promotes entrepreneurial intentions (Nabi et al., 2018). First, *via* its courses and programs, it enables students to enhance their entrepreneurial knowledge, skills, attitudes, and even personal qualities (Wu et al., 2022). Second, incorporating field studies, internships, and extracurricular activities could give students entrepreneurial experience and constructive ideas. Finally, the process of learning could support the building of motivation and commercial networks, which could encourage student involvement (Egan et al., 2017).

Accordingly, our first hypothesis builds on findings showing how entrepreneurship education correlates with entrepreneurial behaviors (Ni and Ye, 2018) as follows:

H1: Entrepreneurship education relates positively to entrepreneurial intention.

The mediating role of an entrepreneurial mindset

An entrepreneurial mindset reflects the capability to identify and exploit opportunities in the entrepreneurial field (Davis et al., 2016). It has been suggested that the surrounding environment paves the way for shaping mindsets (Zhang, 2022) through training or learning (Schmidt and Ford, 2003), thereby supporting the role of entrepreneurship education (Solevik et al., 2013; Cui et al., 2021). As we understand it, this boosting functions in two ways: it creates an entrepreneurial climate at school, and it provides entrepreneur-related experience (Fayolle and Gailly, 2015).

Empirical evidence has shown that an entrepreneurial mindset is closely linked to individuals' entrepreneurial behavior (Lindberg et al., 2017), orienting behavioral patterns toward activities and outcomes related to entrepreneurship (Fayolle and Liñán, 2014). Thus, education has the potential to shape the mindset, which then predicts entrepreneurial intention. Accordingly, we posit that the entrepreneurial mindset could mediate between entrepreneurship education and entrepreneurial intention, as follows:

H2: Entrepreneurship education relates positively to the entrepreneurial mindset.

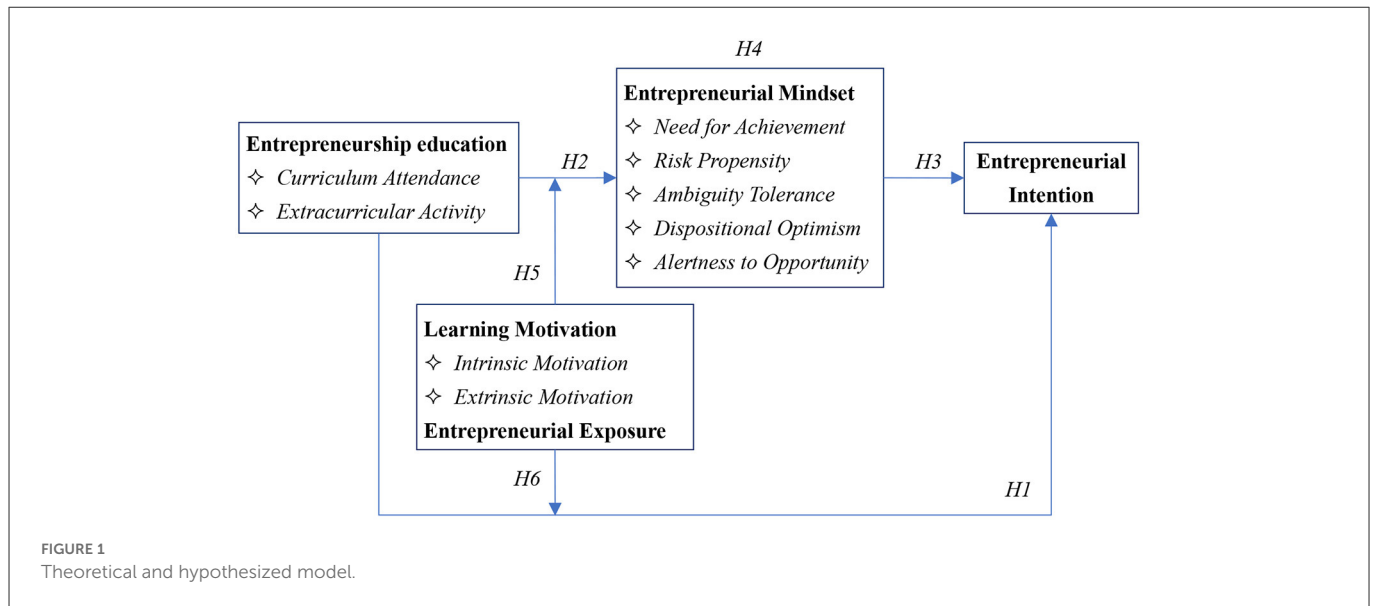
H3: An entrepreneurship mindset relates positively to entrepreneurial intention.

H4: An entrepreneurial mindset mediates the relationship between entrepreneurship education and entrepreneurial intention.

The moderating role of learning motivation

Learning motivation is defined as the psychological will to drive and maintain behavioral patterns (Woolfolk, 2021) that may be generated by learning experiences and rewards. Motivation may be intrinsic or extrinsic (Biggs, 1987), the former referring to an internal intention in individuals to seek and overcome challenges, whereas the latter relates to external benefits, such as obtaining credits, avoiding punishment, and pleasing someone (Woolfolk, 2021). Among those with intrinsic motivation, the behavior itself gives internal and psychological pleasure without external reason or reward (Hytti et al., 2010), whereas extrinsic motivation relies on the environment to reap the rewards and thus is linked to negative emotions and maladaptive behaviors, at least to some extent (Vallerand et al., 1992).

Equally noteworthy is that extrinsic motivation has been reported to encourage individuals to try again, thereby enabling them to complete tasks that do not interest them (Woolfolk, 2021). In sum, according to cognitive theories, intrinsic motivation is more valuable than extrinsic, which is nevertheless necessary for



maintaining motivation in general. Accordingly, we put forward the following hypotheses:

H5a: Intrinsic learning motivation positively moderates the relationship between entrepreneurship education and entrepreneurial mindset.

H5b: Extrinsic learning motivation negatively moderates the relationship between entrepreneurship education and entrepreneurial mindset.

H6a: Intrinsic learning motivation positively moderates the relationship between entrepreneurship education and entrepreneurial intention.

H6b: Extrinsic learning motivation negatively moderates the relationship between entrepreneurship education and entrepreneurial intention.

The moderating role of prior entrepreneurial exposure

Nadelson et al. (2018) refer to “entrepreneurship on a spectrum, recognizing the contextual nature and psychological development associated with entrepreneurial thinking” (p. 114). This spectrum ranges from low to high in characteristics such as visionary thinking, creativity, self-regulation, risk-taking, resilience, and tolerance. According to Botha (2020, p. 2), “Aspiring entrepreneurs are more likely to start new businesses when they learn from existing entrepreneurs in the form of role models,” namely, in response to entrepreneurial exposure. Examples of entrepreneurial exposure include the experience of running a business or being an employer and contact with the family business or other entrepreneurial role models (Krueger, 2015). College students with entrepreneurial experience appear to have a higher level of entrepreneurial intention (Zapkau et al., 2015).

In sum, previous entrepreneurial exposure has the potential to moderate entrepreneurship education among college students.

TABLE 1 Distribution of the validated sample.

Variable	Classification	N of samples	Proportion (%)
Gender	Males	39,564	42.87
	Females	52,720	57.13
Age	20 or less	74,848	81.18
	21 or more	17,353	18.81
Grade	Grade 2018	7,170	7.77
	Grade 2019	16,243	17.60
	Grade 2020	28,208	30.57
	Grade 2021	40,663	44.06
Major	Humanities & social science	18,661	20.22
	Economics & management	18,050	19.56
	Science & engineering	55,573	60.22
Educational level	Junior college	38,915	42.17
	Bachelor's degree	53,369	57.83

N = 92,284.

First, as an entrepreneurial learning experience (Sommarström et al., 2017), it could have a synergistic effect with other learning behaviors in enhancing academic performance; second, it could make those concerned more inclined to run a business in the future (Soria-Barreto et al., 2017), thereby boosting intrinsic motivation and adaptive entrepreneurial behaviors. Accordingly, we propose the following hypotheses:

H5c: Prior entrepreneurial exposure positively moderates the relationship between entrepreneurship education and an entrepreneurial mindset.

H6c: Prior entrepreneurial exposure positively moderates the relationship between entrepreneurship education and entrepreneurial intention.

TABLE 2 The measurement items of the main variables.

Variable	Measurement Item	Loading
Entrepreneurial intention	(1) I am interested in setting up my own business	0.932
	(2) I have considered setting up my own business	0.964
	(3) I am preparing to set up my own business	0.970
	(4) I will try my best to set up my own business	0.934
	(5) I will set up my own business as soon as possible	0.924
Need for achievement	(1) I always do my best whether I am alone or with someone	0.940
	(2) I always try hard to improve on my past performance	0.944
	(3) I enjoy working toward clear and challenging goals	0.943
	(4) In general, I try to make every minute count	0.937
	(5) I often put pressure on myself to achieve as much as I can	0.911
Risk propensity	(1) I like to take chances, although I may fail	0.935
	(2) I like waiting until things have been tested before I try them out	0.913
	(3) I seek new experiences even if their outcomes may be risky	0.950
Ambiguity tolerance	(1) If I am uncertain about the responsibilities involved in a task, I get very anxious	0.894
	(2) It disturbs me when I am unable to follow another person's train of thought	0.839
	(3) Before doing any important task I must know how long it will take	0.930
	(4) A good task is one in which what is to be done and how it is to be done are always clear	0.927
Dispositional optimism	(1) In uncertain times, I would expect the best	0.913
	(2) I am always optimistic about my future	0.926
	(3) Overall, I expect more good things to happen to me than bad	0.934
Alertness to opportunity	(1) I have frequent interactions with others to acquire new information	0.921
	(2) I am keen on looking for information	0.910
	(3) I can recognize links between seemingly unrelated pieces of information	0.937
	(4) I can distinguish between profitable and non-profitable opportunities	0.936
Intrinsic motivation	(1) I am interested in studying entrepreneurship	0.984
	(2) I would study entrepreneurship even if I did not have to	0.982
	(3) Studying entrepreneurship is not useless because one day I may be an entrepreneur myself	0.983

(Continued)

TABLE 2 (Continued)

Variable	Measurement Item	Loading
Extrinsic motivation	(1) Because getting a good diploma is important, one should get good grades in entrepreneurship courses	0.962
	(2) Students of entrepreneurship must manage the courses and exams effectively	0.986
	(3) You have to study entrepreneurship to get a good job	0.987
Extracurricular activity	(1) Entrepreneurship clubs	0.894
	(2) Design competition	0.893
	(3) Listening to entrepreneurs' talks	0.869
	(4) Visit to an enterprise or an internship	0.875
	(5) Face-to-face communication with an entrepreneur	0.939
	(6) Conferences or workshops related to entrepreneurship	0.915
	(7) Business simulators or games	0.892
	(8) Entrepreneurial incubation project	0.954
	(9) Entrepreneurial activity involving resourcing or networking	0.959
	(10) Entrepreneurial spirit and values transmitted by the university or colleges	0.828

Figure 1 depicts our theory-based and hypothesis-based model.

Materials and methods

Participants and data collection

The participants were from more than 100 institutions of higher education in 28 provinces of China. The survey was carried out from December 2021 to January 2022 in the form of an online questionnaire; the participants' consent was obtained, and their information was anonymous. A pilot study was carried out in 2019 to adjust and revise the scale when more than 30,000 college students from 29 institutions of higher education completed the questionnaire. After eliminating invalid data, such as screening questions and unverified information from institutions, and given the limited filling-in time, we included a total of 92,284 questionnaires in the final data, representing a response rate of 75.11%. There were some missing values in a few variables; thus, the final sample for the hypothesis testing was 91510 (see Table 1, for the distribution of the validated sample).

Measures

Entrepreneurial intention

Based on a five-item questionnaire rated on a six-point Likert scale, this instrument was adopted from Chen et al. (1998). College students were asked to indicate how interested they were in and to what extent they were prepared to set up businesses in the future:

TABLE 3 Reliability, validity, correlation, and descriptive statistics.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Intention	(0.945)											
2. Mind_na	0.588	(0.935)										
3. Mind_rp	0.661	0.860	(0.933)									
4. Mind_at	0.540	0.837	0.777	(0.898)								
5. Mind_do	0.576	0.846	0.816	0.787	(0.924)							
6. Mind_ao	0.641	0.857	0.865	0.807	0.855	(0.926)						
7. Mindset	0.681	–	–	–	–	–	–					
8. Mov_i	0.411	0.349	0.369	0.333	0.338	0.372	0.393	(0.983)				
9. Mov_e	0.357	0.326	0.335	0.318	0.314	0.343	0.363	0.969	(0.978)			
10. Exposure	0.084	0.060	0.073	0.061	0.057	0.076	0.076	0.117	0.115	–		
11. Extra	0.533	0.368	0.431	0.364	0.366	0.431	0.448	0.460	0.431	0.136	(0.903)	
12. Curriculum	0.219	0.152	0.173	0.155	0.146	0.177	0.184	0.897	0.916	0.119	0.348	–
Min	1	1	1	1	1	1	1	0	0	0	0	0
Max	6	6	6	6	6	6	6	6	6	1	6	1
Mean	3.954	4.555	4.388	4.481	4.522	4.426	4.436	3.266	3.373	0.259	3.728	0.710
SD	1.460	1.088	1.164	1.099	1.123	1.125	1.050	2.326	2.353	0.438	2.251	0.454
α	0.976	0.972	0.952	0.944	0.946	0.960	–	0.989	0.986	–	0.978	–
CR	0.977	0.972	0.953	0.943	0.946	0.960	–	0.989	0.985	–	0.978	–
AVE	0.893	0.874	0.870	0.807	0.855	0.858	–	0.966	0.957	–	0.815	–
N of items	5	5	3	4	3	4	19	3	3	1	10	1

Intention, entrepreneurial intention; Mind_na, need for achievement; Mind_rp, risk propensity; Mind_at, ambiguity tolerance; Mind_do, dispositional optimism; Mind_ao, alertness to opportunity; Mindset, entrepreneurial mindset; Mov_i, intrinsic motivation; Move, extrinsic motivation; Exposure, prior entrepreneurial exposure; Extra, extracurricular activity; Curriculum, curriculum attendance. The italic values in parenthesis on the diagonal are the square roots of the AVE, and the values below the hypotenuse of the triangle are correlations among the variables. N = 91,510. All correlations are significant at the 0.01 level.

the responses were graded on a scale ranging from 1 to 6, with six signifying the highest level.

Entrepreneurship education

Aimed at capturing participants' learning experiences of entrepreneurship, this instrument measured curriculum attendance and extracurricular activity. The measure of curriculum attendance, adopted from [Sieger et al. \(2014\)](#) and [Cui et al. \(2021\)](#), was based on one multiple-choice question concerning the number of entrepreneurial courses the college students had taken (0 for non-attendance and 1 for attendance). The measure of extracurricular activity (adopted from [Arranz et al., 2017](#); [Cui et al., 2021](#)) was based on 10 items to assess the involvement of the participants in activities and events, such as the design competition, visits or internships, and talks by entrepreneurs, and the impact of these activities on them.

Entrepreneurial mindset

We focused on five factors. The need for achievement was measured on a five-item scale adopted by [Zeffane \(2013\)](#); tolerance of ambiguity and alertness to opportunity were measured on four items developed by [Geller et al. \(1993\)](#) and [Tang et al. \(2012\)](#), respectively, and risk propensity and dispositional optimism were

measured on three items adopted by [Hung et al. \(2012\)](#) and [Crane \(2014\)](#), respectively.

Learning motivation and entrepreneurial exposure

Learning motivation was assessed on six items adopted by [Hytti et al. \(2010\)](#), three items for intrinsic and three for extrinsic motivation. Entrepreneurial exposure, adopted from [Peterman and Kennedy \(2003\)](#) and [Botha \(2020\)](#), concerned any experience the respondents had of starting or running a business, or if they had ever worked (including as an intern) in entrepreneurial business.

[Table 2](#) lists the measurement instruments in some detail.

To avoid possible bias and to improve the reliability of the empirical results, we have included personal background information, including gender, age, grade, major, and educational level as control variables.

Statistical analysis

We used SPSS 27.0 software to cleanse the data and Mplus 8.0 for the analysis. The first step was to conduct an exploratory factor analysis and a confirmatory factor analysis of two random half-samples to ensure reliability and validity. Second, we tested the hypotheses by means of structural equation modeling. Finally, we

tested the mediating role and moderating effect as in Preacher et al. (2007).

Results

Reliability and validity

Table 3 gives the reliability, validity, and descriptive statistics for the variables in the research model. Based on Cronbach's alpha (α) and composite reliability (CR), the α values were above 0.8, ranging from 0.944 to 0.989, and the CR value was above 0.6, ranging from 0.943 to 0.989 (Bagozzi and Yi, 1988), indicating high reliability.

The standardized coefficient loading of the items on the corresponding construct was significant (above 0.5 and ranging from 0.828 to 0.987, see Table 2), and the average variance extracted (AVE) values were higher than 0.5 (Bagozzi and Yi, 1988), indicating convergent validity for each variable (Fornell and Larcker, 1981). The square root of the AVE (the diagonal elements in Table 3) was larger than that of the off-diagonal elements at the level of significance (Hulland, 1999), indicating discriminant validity (Fornell and Larcker, 1981).

Common method variance

We used Harman's single-factor method to test common method variance (Podsakoff et al., 2003). The percentage of variance explained by the first factor in the exploratory factor analysis was far below the threshold of 0.50. In the confirmatory factor analysis, the fitness of the single-factor model failed to meet the criteria (CFI = 0.448 < 0.90, TLI = 0.418 < 0.90, RMSEA = 0.221 > 0.08, SRMR = 0.195 > 0.08), the values of the variance inflation factor were <3, and the values of tolerability were more than 0.30. Therefore, common method variance did not affect the outcome of this study.

Hypothesis testing

Correlation

Table 3 also shows the results of the correlation analysis, namely, the positive correlation ($p < 0.01$) between entrepreneurship education (curriculum attendance and extracurricular activity), entrepreneurial mindset, entrepreneurial intention, learning motivation, and entrepreneurial exposure. However, further analysis is necessary to test the validity of the research hypotheses.

Direct effect

Table 4 presents the results of the path analysis of the effect of entrepreneurship education on entrepreneurial intention and an entrepreneurial mindset. The path coefficients indicating the direct effects of (a) curriculum attendance and extracurricular activity on entrepreneurial intentions, (b) curriculum attendance and extracurricular activity on an entrepreneurial mindset, and (c) an entrepreneurial mindset on entrepreneurial intentions were all positive and significant ($p < 0.001$), thereby supporting H1, H2, and H3.

TABLE 4 Path analysis of the direct and indirect predictions.

Variable	DV = Intention		DV = Mindset	
	Estimate	S.E.	Estimate	S.E.
Intercept	2.081***	0.068	1.463***	0.075
EL	−0.138***	0.007	−0.117***	0.007
Gender	−0.214***	0.007	−0.055***	0.007
Age	−0.090***	0.003	0.145***	0.004
Grade	−0.027***	0.005	−0.207***	0.005
Major 2	0.146***	0.009	−0.200***	0.009
Major 3	0.120***	0.008	−0.094***	0.008
Curriculum	0.170***	0.009	0.128***	0.008
Extra	0.182***	0.002	0.190***	0.002
Mindset	0.718***	0.004		
R^2	0.539	0.003	0.268	0.004
F	196.219***		75.464***	
IND_C	0.092***	0.006		
	[0.080, 0.104]			
IND_E	0.137***	0.001		
	[0.134, 0.139]			

Control variables are as follows: EL (educational level and junior college as reference), gender (female as reference), age, grade, major (humanities and social science as a reference, major 2 = economics and management, major 3 = science and engineering). IND_C/IND_E the mediating effect of entrepreneurial mindset from curriculum attendance/extracurricular activity to entrepreneurial intention.

N = 91,510. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, two-tailed test. Numbers in [] are confidence intervals at the 95% level and bootstrapping $n = 5,000$.

Mediating effect

Path analysis and bootstrapping ($n = 5,000$) were used in the process of testing mediation (Preacher et al., 2007), as shown in Table 4. The indirect path to an entrepreneurial mindset from curriculum attendance and extracurricular activity was positive and significant ($p < 0.001$), and the bootstrapping confidence intervals at 95% were [0.080, 0.104] and [0.134, 0.139], which skip 0. Hence, the indirect path of an entrepreneurial mindset from entrepreneurship education to entrepreneurial intention was significant. The direct coefficients of entrepreneurship education were significant, too, confirming the role of an entrepreneurial mindset as a partial mediator. Therefore, H4 was supported.

Moderating effect

Table 5 presents the results of the tests for a moderating effect between learning motivation among college students and prior entrepreneurial exposure. The coefficients of the interaction connecting curriculum attendance and intrinsic learning motivation with entrepreneurial intention and an entrepreneurial mindset are positively significant, whereas those concerning extracurricular activity are negatively significant. Therefore, H5a and H6a are partially supported: in other words, intrinsic learning motivation among college students positively moderates the effect of curriculum attendance on entrepreneurial intention and mindset.

TABLE 5 Path analysis of the moderating effects.

Variable	DV = Intention				DV = Mindset			
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	2.226***	0.063	2.065***	0.067	1.565***	0.075	1.466***	0.075
EL	−0.138***	0.007	−0.132***	0.007	−0.107***	0.007	−0.113***	0.007
Gender	−0.196***	0.007	−0.210***	0.007	−0.038***	0.006	−0.051***	0.007
Age	−0.081***	0.003	−0.090***	0.003	0.140***	0.004	0.145***	0.004
Grade	−0.024***	0.005	−0.035***	0.005	−0.164***	0.005	−0.212***	0.005
Major 2	0.128***	0.009	0.145***	0.009	−0.196***	0.009	−0.201***	0.009
Major 3	0.102***	0.008	0.120***	0.008	−0.108***	0.008	−0.094***	0.008
Mindset	0.629***	0.005	0.717***	0.004				
Curriculum	−0.672***	0.025	0.197***	0.010	−2.033***	0.023	0.146***	0.009
Extra	0.194***	0.003	0.177***	0.002	0.154***	0.003	0.181***	0.002
Movi	0.232***	0.006			0.152***	0.005		
Move	−0.115***	0.006			0.106***	0.005		
Exposure			0.073***	0.022			−0.038	0.020
Curriculum* movi	0.232***	0.006			0.152***	0.005		
	[0.220, 0.243]				[0.141, 0.162]			
Curriculum* move	−0.115***	0.006			0.106***	0.005		
	[−0.127, −0.103]				[0.095, 0.116]			
Extra* movi	−0.034***	0.003			−0.007***	0.003		
	[−0.040, −0.028]				[−0.012, −0.002]			
Extra* move	0.024***	0.003			−0.003	0.003		
	[0.018, 0.030]				[−0.008, 0.003]			
Curriculum* exposure			−0.122***	0.021			−0.075***	0.020
			[−0.161, −0.081]				[−0.113, −0.035]	
Extra* exposure			0.020***	0.004			0.035***	0.004
			[0.013, 0.028]				[0.028, 0.042]	
R ²	0.562	0.003	0.539	0.003	0.410	0.004	0.269	0.004
F	212.625***		196.566***		114.655***		75.957***	

N = 91,510. *p < 0.05, **p < 0.01, ***p < 0.001, two-tailed test. Numbers in [] are confidence intervals at the 95% level and bootstrapping n = 5,000.

The significance of the coefficients ($p < 0.001$) indicates that the interaction between curriculum attendance and extrinsic learning motivation is negatively related to entrepreneurial intention but positively related to an entrepreneurial mindset. Moreover, the interaction between extracurricular activity and extrinsic learning motivation is positively and significantly related to entrepreneurial intention but not significantly related to an entrepreneurial mindset. Therefore, H6b is partially supported, while H5b is not supported. In other words, extrinsic learning motivation among college students negatively moderates the relationship between curriculum attendance and entrepreneurial intention.

The coefficients of the interaction linking curriculum attendance and prior entrepreneurial exposure to entrepreneurial intention and an entrepreneurial mindset are negatively significant, whereas those of the interaction linking extracurricular activity and prior entrepreneurial exposure are positive and significant. Therefore, H5c and H6c are partially supported. In other words, previous

entrepreneurial exposure positively moderates the predicted effect of extracurricular activity on entrepreneurial intention and an entrepreneurial mindset.

Robustness

Five dimensions of an entrepreneurial mindset were introduced as five mediators in the robustness tests, and the results are shown in Table 6. Without the change, the mediating effect was relatively robust, not only for an entrepreneurial mindset but also for the five dimensions.

Discussion

With regard to H1, the results showed correlations between entrepreneurship education and entrepreneurial intention, which is in line with existing empirical evidence both at home and abroad (Cui et al., 2021; Wang et al., 2021), namely, entrepreneurship education

TABLE 6 Robustness testing: The results of a mediating role.

Variable	DV = Intention				
CVs	Yes	Yes	Yes	Yes	Yes
Curriculum	0.193***	0.166***	0.198***	0.200***	0.183***
Extra	0.224***	0.217***	0.237***	0.225***	0.198***
Mind_na	0.580***				
Mind_rp		0.471***			
Mind_at			0.505***		
Mind_do				0.551***	
Mind_ao					0.613***
R ²	0.484	0.476	0.447	0.477	0.506
F	173.656***	157.637***	158.621***	170.566***	178.956***
Variable	DV = Mind_na	DV = Mind_rp	DV = Mind_at	DV = Mind_do	DV = Mind_ao
CVs	Yes	Yes	Yes	Yes	Yes
Curriculum	0.119***	0.204***	0.127***	0.113***	0.129***
Extra	0.164***	0.216***	0.163***	0.171***	0.197***
R ²	0.177	0.532	0.153	0.153	0.244
F	65.642***	58.418***	63.191***	68.467***	73.815***
IND_C	0.069***	0.096***	0.064***	0.062***	0.079***
	[0.059, 0.079]	[0.086, 0.106]	[0.055, 0.073]	[0.052, 0.072]	[0.068, 0.090]
IND_E	0.095***	0.102***	0.082***	0.094***	0.121***
	[0.093, 0.098]	[0.099, 0.104]	[0.080, 0.084]	[0.092, 0.097]	[0.118, 0.123]

N = 91,510. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, two-tailed test. Numbers in [] are confidence intervals at the 95% level and bootstrapping $n = 5,000$.

helps to equip students with the relevant knowledge and skills to take advantage of opportunities in the entrepreneurial field (Nadelson et al., 2018). Furthermore, related to H5 and H6 together, the above-mentioned experiences of learning and entrepreneurial exposure may enhance business success later on.

The results related to H2, H3, and H4 verified the role of a mediator in developing an entrepreneurial mindset, correlating with entrepreneurship education, and entrepreneurial intention. Growth orientation could develop through training and acquiring knowledge of entrepreneurship (Schmidt and Ford, 2003), boosting achievement needs, ambiguity tolerance, alertness, risk propensity, and dispositional optimism, for example. These psychological factors help college students to identify and exploit opportunities associated with business ownership (Davis et al., 2016). Again, according to the results related to H5 and H6, learning motivation moderates the correlations between entrepreneurship education and entrepreneurial intention: similar to mindset, it represents the psychological approach to behavioral patterns (Woolfolk, 2021). Thus, the educational environment should better embody growth-mindset principles and internal-motivation practices (Zhang, 2022).

In sum, creating an entrepreneurial learning climate facilitates the acquisition of entrepreneurship-related knowledge and skills on the one hand and encourages students to be self-efficient on the other, thereby supporting them in the business-startup process. We assume that the key issue of transiting the entrepreneurial environment relies on the shaping of implicit theories of entrepreneurial capability and the building of internal motivation.

Implications

On the theoretical level, this study contributes to the literature on entrepreneurship education, especially from the perspective of educational psychology. First, the empirical study conducted among Chinese students verified the correlation between entrepreneurship education and entrepreneurial intention, thereby enriching the existing literature from Eastern cultures. Second, the findings acknowledge the supporting role of psychological aspects such as mindset and motivation in entrepreneurship education, verifying the importance of students' subjective initiatives in validating learning experiences. Third, earlier entrepreneurial exposure was introduced as a contextual factor, thereby reflecting the importance of the entrepreneurial climate in validating entrepreneurship education.

On the practical level, the findings offer insights that could help administrators in government and policymakers in institutions to develop entrepreneurship education further. First, they attest to the value of education in fostering innovativeness among students in higher education institutions. Thus, more attention in government and these institutions should be placed on supporting entrepreneurship education, such as in the funding and reinforcement of the curriculum. Second, mindset, motivation, and prior experience are advantageous to students in developing behavioral patterns related to entrepreneurship. Teachers should therefore be aware of what students enjoy in the learning process so that they can motivate them further and tailor relevant courses or activities to maximize the benefits.

Limitations and future directions

The current study has several limitations that future research might draw on. First, given the major revision of the entrepreneurial mindset adopted in the current study from Dweck's original, there is a need to gather more empirical evidence to test the validity and generality of the five-item scale of the entrepreneurial mindset adopted from existing literature. Second, according to our findings, the moderating roles of learning motivation and entrepreneurial exposure varied with aspects of entrepreneurship education (curriculum attendance and extracurricular activity), but we failed to identify the underlying details. Future research should, therefore, focus on specific issues of education that may affect its impacts, such as curriculum planning, pedagogical strategies, and teaching materials. Third, given the tremendous sample size of our cross-sectional survey, although it thoroughly identified the correlations between entrepreneurship education and corresponding intentions *via* the mindset, it was not a scientifically causal analysis. A quasi-experiment comprising a randomized controlled trial or dynamic tracking would shed deeper light on the causal mechanism. Fourth, the empirical study was merely conducted among Chinese students; future research on patterns of entrepreneurial thinking that may be culture dependent should focus on the cross-cultural perspective, thereby tailoring education to improving the entrepreneurial climate.

Conclusion

Using a cross-sectional survey of almost 100,000 college students, we examined the correlations between entrepreneurship education and the corresponding mindset and intention, including the mediating role of the entrepreneurial mindset and the moderating role of learning motivation and entrepreneurial exposure. The findings indicate that, first, entrepreneurship education, comprising curriculum attendance or extracurricular activities, directly predicts an entrepreneurial mindset and entrepreneurial intention. Second, entrepreneurship education predicts entrepreneurial intention indirectly by stimulating the mindsets of college students, verifying the mediating role of the entrepreneurial mindset. Third, intrinsic motivation positively moderates the relationships between curriculum attendance and entrepreneurial intention/mindset, whereas extrinsic motivation moderates them negatively. Fourth, previous entrepreneurial exposure among college students positively moderates the prediction of entrepreneurship education (particularly extracurricular activity) in terms of entrepreneurial intention. In addition to shedding light on the ordinary routine of entrepreneurship education, the results verify the importance of individuals' psychological initiatives and the learning climate in establishing adaptive behavioral patterns. The

vital aspect of transiting the entrepreneurial environment may rely on the shaping of subjective entrepreneurial initiatives.

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 Ethical Review Board of the School of Education at Nanjing University, as well as the colleges or universities involved in the survey. All participants gave their informed consent. Written informed consent was not required.

Author contributions

JSu designed the study, conducted the analysis, drafted the manuscript, and reviewed and revised the manuscript. JSh designed the study and reviewed the manuscript. JZ drafted, reviewed, and revised the manuscript. All authors approved the final submission and publication.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 71974093).

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

- Arranz, N., Ubierna, F., Arroyabe, M. F., Perez, C., and Arroyabe, F. D. (2017). The effect of curricular and extracurricular activities on university students' entrepreneurial intention and competence. *Stud. High. Educ.* 42, 1979–2008. doi: 10.1080/03075079.2015.1130030
- Bae, T. J., Qian, S., Miao, C., and Fiet, J. O. (2014). The relationship between entrepreneurship education and entrepreneurial intentions: a meta-analytic review. *Entrep. Theory Pract.* 38, 217–254. doi: 10.1111/etap.12095

- Bagozzi, R. P., and Yi, Y. (1988). On the evaluation of structural equation models. *J. Acad. Mark. Sci.* 16, 74–94. doi: 10.1007/BF02723327
- Biggs, J. B. (1987). *Study Process Questionnaire Manual*. Melbourne, VIC: Australian Council for Educational Research.
- Botha, M. (2020). Prior entrepreneurial exposure and action of women entrepreneurs: exploring the moderation effects of entrepreneurial competencies in a developing country context. *Front. Psychol.* 11, 922. doi: 10.3389/fpsyg.2020.00922
- Chen, C. C., Greene, P. G., and Crick, A. (1998). Does entrepreneurial self-efficacy distinguish entrepreneurs from managers? *J. Bus. Ventur.* 13, 295–316. doi: 10.1016/S0883-9026(97)00029-3
- Chlosta, S., Patzelt, H., Klein, S. B., and Dormann, C. (2012). Parental role models and the decision to become self-employed: the moderating effect of personality. *Small Bus. Econ.* 38, 121–138. doi: 10.1007/s11187-010-9270-y
- Crane, F. G. (2014). Measuring and enhancing dispositional optimism and entrepreneurial intent in the entrepreneurial classroom: an Bahamian study. *J. Acad. Bus. Educ.* 15, 94–104.
- Cui, J., Sun, J., and Bell, R. (2021). The impact of entrepreneurship education on the entrepreneurial mindset of college students in China: the mediating role of inspiration and the role of educational attributes. *Int. J. Manag. Stud.* 19, 100296. doi: 10.1016/j.ijme.2019.04.001
- Daspit, J. J., Fox, C. J., and Findley, S. K. (2021). Entrepreneurial mindset: an integrated definition, a review of current insights, and directions for future research. *J. Small Bus. Manag.* doi: 10.1080/00472778.2021.1907583
- Davis, M. H., Hall, J. A., and Mayer, P. S. (2016). Developing a new measure of entrepreneurial mindset: reliability, validity, and implications for practitioners. *Consult. Psychol. J. Pract. Res.* 68, 21–48. doi: 10.1037/cpb0000045
- Egan, A., Maguire, R., Christophers, L., and Rooney, B. (2017). Developing creativity in higher education for 21st century learners: a protocol for a scoping review. *Int. J. Educ. Res.* 82, 21–27. doi: 10.1016/j.ijer.2016.12.004
- Fayolle, A., and Gailly, B. (2015). The impact of entrepreneurship education on entrepreneurial attitudes and intention: hysteresis and persistence. *J. Small Bus. Manag.* 53, 75–93. doi: 10.1111/jsbm.12065
- Fayolle, A., and Liñán, F. (2014). The future of research on entrepreneurial intentions. *J. Bus. Res.* 67, 663–666. doi: 10.1016/j.jbusres.2013.11.024
- Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.1177/002224378101800104
- Geller, G., Tambor, E. S., Chase, G. A., and Holtzman, N. A. (1993). Measuring physicians' tolerance for ambiguity and its relationship to their reported practices regarding genetic testing. *Med. Care.* 31, 989–1001. doi: 10.1097/00005650-199311000-00002
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strateg. Manage. J.* 20, 195–204. doi: 10.1002/(SICI)1097-0266(199902)20:2<195::AID-SMJ13>3.0.CO;2-7
- Hung, K., Tangpong, C., Li, J., and Li, Y. (2012). Robustness of general risk propensity scale in cross-cultural settings. *J. Manage. ISS.* 24, 78–96. Available online at: <https://www.jstor.org/stable/43488124>
- Hytti, U., Stenholm, P., Heinonen, J., and Seikkula-Leino, J. (2010). Perceived learning outcomes in entrepreneurship education: the impact of student motivation and team behavior. *Educ. Train.* 50, 587–606. doi: 10.1108/00400911011088935
- Jack, S. L., and Anderson, A. R. (1999). Entrepreneurship education within the enterprise culture: producing reflective practitioners. *Int. J. Entrep. Behav. Res.* 5, 110–125. doi: 10.1108/13552559910284074
- Krueger, N. (2015). *Entrepreneurial Education in Practice: Part 1—The Entrepreneurial Mindset*. Paris: OECD.
- Krueger, N. F., Reilly, M. D., and Carsrud, A. L. (2000). Competing models of entrepreneurial intentions. *J. Bus. Ventur.* 15, 411–432. doi: 10.1016/S0883-9026(98)00033-0
- Kuh, G. D. (2001). Assessing what really matters to student learning inside the national survey of student engagement. *Change* 33, 10–17. doi: 10.1080/00091380109601795
- Li, L., and Wu, D. (2019). Entrepreneurial education and students' entrepreneurial intention: does team cooperation matter? *J. Glob. Entrep. Res.* 9, 1–13. doi: 10.1186/s40497-019-0157-3
- Lindberg, E., Bohman, H., Hultén, P., and Wilson, T. (2017). Enhancing students' entrepreneurial mindset: a Swedish experience. *Educ. Train.* 59, 68–779. doi: 10.1108/ET-09-2016-0140
- Nabi, G., Walmsley, A., Liñán, F., Akhtar, I., and Neame, C. (2018). Does entrepreneurship education in the first year of higher education develop entrepreneurial intentions? The role of learning and inspiration. *Stud. High. Educ.* 43, 452–467. doi: 10.1080/03075079.2016.1177716
- Nadelson, L. S., Palmer, A. D., Benton, T., Basnet, R., Bissonnette, M., Cantwell, L., et al. (2018). Developing next generation of innovators: teaching entrepreneurial mindset elements across disciplines. *Inter. J. High. Educ.* 7, 114–126. doi: 10.5430/ijhe.v7n5p114
- Ni, H., and Ye, Y. (2018). Entrepreneurship education matters: exploring secondary vocational school students' entrepreneurial intention in China. *Asia-Pac. Educ. Res.* 27, 409–418. doi: 10.1007/s40299-018-0399-9
- Peterman, N. E., and Kennedy, J. (2003). Enterprise education: influencing students' perceptions of entrepreneurship. *Entrep. Theory Pract.* 28, 129–144. doi: 10.1046/j.1540-6520.2003.00035.x
- Pfeifer, S., Šarlija, N., and Zekić Sušac, M. (2016). Shaping the entrepreneurial mindset: entrepreneurial intentions of business students in Croatia. *J. Small Bus. Manag.* 54, 102–117. doi: 10.1111/jsbm.12133
- Podsakoff, P. M., MacKenzie, S. B., Lee, J., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88, 879–903. doi: 10.1037/0021-9010.88.5.879
- Preacher, K. J., Rucker, D. D., and Hayes, A. F. (2007). Addressing moderated mediation hypotheses: theory, methods, and prescriptions. *Multivariate Behav. Res.* 42, 185–227. doi: 10.1080/00273170701341316
- Rauch, A., and Hulsink, W. (2015). Putting entrepreneurship education where the intention to act lies: an investigation into the impact of entrepreneurship education on entrepreneurial behavior. *Acad. Manag. Learn. Educ.* 14, 187–204. doi: 10.5465/amle.2012.0293
- Santos, S. C., Neumeyer, X., and Morris, M. H. (2019). Entrepreneurship education in a poverty context: an empowerment perspective. *J. Small Bus. Manag.* 57 (Suppl. 1), 6–32. doi: 10.1111/jsbm.12485
- Schmidt, A. M., and Ford, J. K. (2003). Learning within a learner control training environment: the interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Pers. Psychol.* 56, 405–429. doi: 10.1111/j.1744-6570.2003.tb00156.x
- Sieger, P., Fueglistaller, U., and Zellweger, T. (2014). *Student Entrepreneurship Across the Globe: A Look at Intentions and Activities*. St. Gallen: Swiss Research Institute of Small Business and Entrepreneurship at the University of St.Gallen (KMU-HSG).
- Solesvik, M. Z., Westhead, P., Matlay, H., and Parsyck, V. N. (2013). Entrepreneurial assets and mindsets: benefit from university entrepreneurship education investment. *Educ. Train.* 55, 748–762. doi: 10.1108/ET-06-2013-0075
- Sommarström, K., Ruskovaara, E., and Pihkala, T. (2017). Company visits as an opportunity for entrepreneurial learning. *J. Inter. Bus. Entrep. Dev.* 10, 298–315. doi: 10.1504/JIBED.2017.085505
- Soria-Barreto, K., Honores-Marin, G., Gutiérrez-Zepeda, P., and Gutiérrez-Rodríguez, J. (2017). Prior exposure and educational environment towards entrepreneurial intention. *J. Tech. Manag. Innov.* 12, 45–58. doi: 10.4067/S0718-27242017000200006
- Sutter, C., Bruton, G. D., and Chen, J. (2019). Entrepreneurship as a solution to extreme poverty: a review and future research directions. *J. Bus. Ventur.* 34, 197–214. doi: 10.1016/j.jbusvent.2018.06.003
- Tang, J., Kacmar, K. M., and Busenitz, L. (2012). Entrepreneurial alertness in the pursuit of new opportunities. *J. Bus. Ventur.* 27, 77–94. doi: 10.1016/j.jbusvent.2010.07.001
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C., and Vallières, E. F. (1992). The academic motivation scale: a measure of intrinsic, extrinsic, and amotivation in education. *Educ. Psychol. Meas.* 52, 1003–1017. doi: 10.1177/0013164492052004025
- Wang, Q., Sun, Z., and Wu, C. (2021). The impact of university innovation and entrepreneurship education on entrepreneurial intention from the perspective of educational psychology. *Front. Psychol.* 12, 745976. doi: 10.3389/fpsyg.2021.745976
- Woolfolk, A. (2021). 教育心理学：主动学习版[*Educational Psychology: Active Learning Edition*] (13th Edn). Beijing: China Machine Press.
- Wu, L., Jiang, S., Wang, X., Yu, L., Wang, Y., and Pan, H. (2022). Entrepreneurship education and entrepreneurial intentions of college students: the mediating role of entrepreneurial self-efficacy and the moderating role of entrepreneurial competition experience. *Front. Psychol.* 12, 727826. doi: 10.3389/fpsyg.2021.727826
- Zapkau, F. B., Schwens, C., Steinmetz, H., and Kabst, R. (2015). Disentangling the effect of prior entrepreneurial exposure on entrepreneurial intention. *J. Bus. Res.* 68, 639–653. doi: 10.1016/j.jbusres.2014.08.007
- Zeffane, R. (2013). Need for achievement, personality and entrepreneurial potential: a study of young adults in the United Arab Emirates. *J. Enterp. Cult.* 21, 75–105. doi: 10.1142/S0218495813500040
- Zhang, J. (2022). What characterises an effective mindset intervention in enhancing students' learning? A systematic literature review. *Sustainability* 14, 3811. doi: 10.3390/su14073811
- Zhang, J., and Huang, J. (2021). Entrepreneurial self-efficacy mediates the impact of the post-pandemic entrepreneurship environment on college students' entrepreneurial intention. *Front. Psychol.* 12, 643184. doi: 10.3389/fpsyg.2021.643184

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

