

Research on teaching strategies and skills in different educational stages

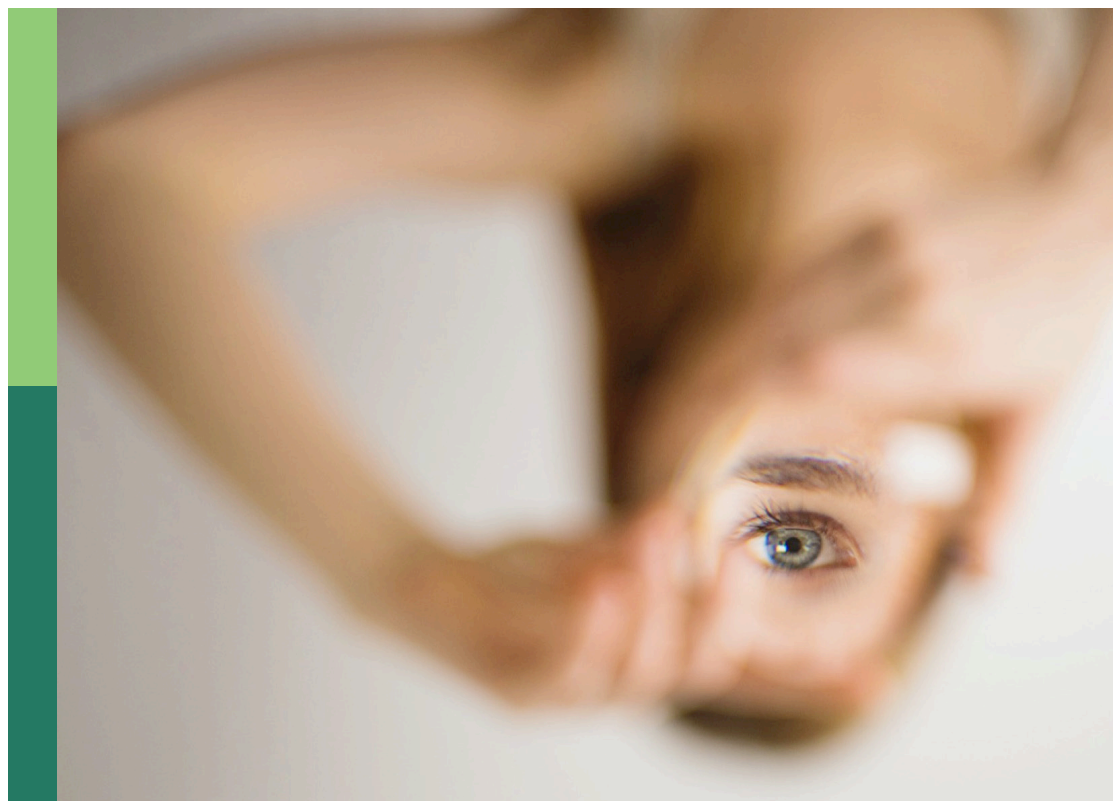
Edited by

Pedro Gil-Madrone, Linda Saraiva and Luisa Losada-Puente

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 978-2-8325-6129-4
DOI 10.3389/978-2-8325-6129-4

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

Research on teaching strategies and skills in different educational stages

Topic editors

Pedro Gil-Madrona — University of Castilla-La Mancha, Spain

Linda Saraiva — Polytechnic Institute of Viana do Castelo, Portugal

Luisa Losada-Puente — University of A Coruña, Spain

Citation

Gil-Madrona, P., Saraiva, L., Losada-Puente, L., eds. (2025). *Research on teaching strategies and skills in different educational stages*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6129-4

Table of contents

06	Online collaborative Padlet-mediated learning in health management studies Lior Naamati-Schneider and Dorit Alt
16	Simulation-based learning in teacher education: Using Maslow's Hierarchy of needs to conceptualize instructors' needs Rivi Frei-Landau and Orna Levin
30	Student-centered teacher responses to student behavior in the classroom: A systematic review Jirina Karasova and Jan Nehyba
44	Preschool teachers display a flexible pattern of pedagogical actions in promoting healthy habits in children Verónica Violant-Holz, Carlota Rodríguez-Silva and Manuel J. Rodríguez
56	Elementary school teachers' perceptions of the characteristics of twice-exceptional students Omar Abdullah Alsamani, Yasir Ayed Alsamiri and Seham Defallah Alfaidi
64	Dyads or quads? Impact of group size and learning context on collaborative learning Mindi Wang, Ling Jiang and Heng Luo
75	Perceived use and value of reproductive, problem-solving, and athlete-initiated teaching by coaches and athletes Koray Kiliç and Mustafa Levent Ince
83	Involvement, leadership and social practice to the development of postgraduate attributes: evidence from extracurricular education with Chinese characteristics Sen Li, Yusheng Hu, Lingling Wang and Enjun Xia
97	Factors associated with implemented teacher-led movement and physical activity in early childhood education and care Ann-Christin Sollerhed
106	Relationship between quality of professional capacity building for kindergarten teachers and children's language development: the mediating role of kindergarten language education activities quality Qiong Wu, Guoxia Wang and Chang Li
115	Exploring native and non-native English speaker teachers' perceptions of English teacher qualities and their students' responses Liwei Deng, Lawrence Jun Zhang and Naashia Mohamed
124	Mapping principles and worked examples for structural learning: effects of content complexity Hsinmei Liao

- 137 **Conceptualising teacher dialogic artistry: a nexus of pedagogical dispositions and language qualities**
Angela Hadjipanteli and Maria Vrikki
- 144 **How much do teachers know about self-regulated learning? To what degree and in what way do they encourage it in students?**
Slavica Šimić Šašić, Anela Nikčević-Milković and Maja Cindrić
- 155 **External assessment of teachers' roles during children's free play and its relation to types of children's play**
Cornelia Rüdisüli, Isabelle Duss, Patricia Lannen and Corina Wustmann Seiler
- 168 **Perceptions and preparedness of secondary teacher trainees to foster inclusive schools for all**
Mercedes Arias-Pastor, Steven Van Vaerenbergh, Jessica Fernández-Solana and Jerónimo J. González-Bernal
- 181 **Focusing on the value of cooperative learning in physical education: a bibliometric analysis**
Tong Zhou, Huayi Wang and Dong Li
- 194 **Research on service quality of professional education based on the PZB service quality model**
Xuejun Deng, Liang Liu and Xiaojun Liu
- 206 **Adaptation of the SAGEFS scale on attitudes toward gender equality in football in the international school context**
Luz Marina Méndez-Hinojosa, Pedro Gil-Madrona, David Zamorano-García, Juan Ángel Simón-Piqueras, Juan Carlos Padierna-Cardona, Elizabeth Flores-Ferro, Martín Francisco González-Villalobos, Ratko Pavlovic, Fernando Maureira-Cid, Natalia Gómez-Santos, Sonia Morales-Calvo and Magaly Cárdenas-Rodríguez
- 215 **Is it possible to identify physical-motor profiles of preschool children on their association with selected biosocial factors?**
Pedro Gil-Madrona, Luisa Losada-Puente, Paula Mendiri, César Sá, Inês P. Silva and Linda Saraiva
- 227 **Analogical thinking modifiability and math processing strategy**
David Tzuriel
- 237 **Retrieval-based concept mapping makes a difference as a retrieval practice activity: a study with high school students**
María Teresa Lechuga, Juana M. Ortega-Tudela and Carlos J. Gómez-Ariza
- 249 **Formative assessment in higher education: an exploratory study within programs for professionals in education**
Davide Parmigiani, Elisabetta Nicchia, Emiliana Murgia and Marcea Ingersoll

- 265 **Word meaning types acquired before vs. after age 5: implications for education**
Andrew Biemiller
- 279 **Biotechnology project-based learning encourages learning and mathematics application**
Yuliana Jiménez-Gaona and Oscar Vivanco-Galván
- 288 **Physical education students' reflections about the learning outcomes of different teaching methods: a mixed methods study**
Arne Sørensen and Pål Lagestad
- 299 **Challenge-obstacle stressors and cyberloafing among higher vocational education students: the moderating role of smartphone addiction and Maladaptive**
Emilio Jesús Lizarte Simón, Meriem Khaled Gijón, María Carmen Galván Malagón and José Gijón Puerta
- 309 **A study of factors and their influence on mathematics-related beliefs and performance in Turkey**
Mehmet Ali Kandemir, Nurullah Eryilmaz and Mack Shelley
- 319 **Critical thinking pedagogical practices in medical education: a systematic review**
Beatriz Araújo, Sandra F. Gomes and Laura Ribeiro
- 334 **Shaping future-ready graduates with mindset shifts: studying the impact of integrating critical and design thinking in design innovation education**
Nadya Shaznay Patel, Shermain Puah and Xiao-Feng Kenan Kok
- 350 **Strategic and self-regulated instruction in synthesis tasks in the university context**
P. Robledo and M. L. Álvarez
- 363 **Flipping the anatomy classroom: a comparative analysis of 16-week and 8-week courses in a community college**
Chaya Gopalan, Elizabeth L. Bruno, Sheyenne E. Daugherty and Kim-Leiloni T. Nguyen



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña,
Spain

REVIEWED BY

David Zamorano,
University of Castilla La Mancha,
Spain
María Martínez López,
University of Castilla La Mancha,
Spain

*CORRESPONDENCE

Lior Naamati-Schneider
✉ liorna@hac.ac.il

SPECIALTY SECTION

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

RECEIVED 02 February 2023

ACCEPTED 13 March 2023

PUBLISHED 30 March 2023

CITATION

Naamati-Schneider L and Alt D (2023) Online
collaborative Padlet-mediated learning in
health management studies.
Front. Psychol. 14:1157621.
doi: 10.3389/fpsyg.2023.1157621

COPYRIGHT

© 2023 Naamati-Schneider and Alt. 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.

Online collaborative Padlet-mediated learning in health management studies

Lior Naamati-Schneider^{1*} and Dorit Alt²

¹Health Systems Management Department, Hadassah Academic College, Jerusalem, Israel, ²Faculty of Education and Instruction, Tel Hai College, Upper Galilee, Israel

Introduction: The Padlet interactive platform constitutes a virtual online board on which users can post various types of multimedia content, such as documents, questions, comments, images, video clips, and audio clips. This platform has gained popularity in higher education yet remains ancillary in the fields of medical education and medical management.

Methods: This case study sought to initially design an intervention program that employs online collaborative learning enabled by Padlet and to assess how Health Management students perceived the activity and its ramifications for their learning. 85 students' reflective journals were content analyzed.

Results: five main categories stood out: (1) user experience of the platform; (2) visibility and visibility of the platform; (3) collaborative learning and OCL; (4) active online learning and student engagement; and (5) cognitive flexibility.

Discussion: This study emphasizes the importance of incorporating digital tools in education, particularly the use of the Padlet platform to facilitate collaborative learning and improve the quality of teaching and learning. It demonstrates that Padlet-mediated online collaborative learning can be an effective digital learning tool because of its ease of use and ability to accommodate both pedagogical and technological challenges.

KEYWORDS

online collaborative learning, Padlet, health management, higher education, collaborative learning

Introduction

The healthcare industry has faced numerous challenges in recent years due to rapid global changes, increased access to information, and evolving labor market demands. As a result, the training of healthcare administrators must be competency-based and focus on developing essential skills such as cognitive and interpersonal abilities (Senkubuge et al., 2014; Naamati-Schneider, 2020). The dynamic nature of healthcare systems, which involves interactions between individuals and their environments, requires teaching methods that can help healthcare providers adapt to the changing demands of the industry (Reis, 2019).

This study proposes the evaluation of the effectiveness of online collaborative learning (OCL) as a teaching method for health management students, utilizing Padlet as a digital platform for interactive collaboration. Previous research in the field of health professions, including health administration, has demonstrated that OCL in digital environments can enhance student learning through collaborative problem-solving and the acquisition of knowledge, as well as the cultivation of higher-order thinking skills (Xiong et al., 2016).

The acquisition of skills such as flexibility, effective decision-making, and interpersonal communication is crucial for health professionals, who often work in multidisciplinary and culturally diverse teams. These skills can be developed through exposure to online collaborative

learning (OCL) environments early on in professional training (Männistö et al., 2020). The adoption and advancement of OCL pedagogies across all levels of education, from undergraduate studies to professional development, is widely believed to benefit both the health system and its patients by enabling health workers to meet the demands of the healthcare market (Alt et al., 2022). This is particularly relevant in light of the practical nature of work in the health industry (Grosser et al., 2020).

The current study hence sought to demonstrate the introduction of an innovative pedagogy of online collaborative learning (OCL) with the aid of Padlet technology (Ramachandiran and Mahmud, 2018), for health management undergraduate students. Our aim was to examine students' perceptions of their experiences of OCL enabled by Padlet from their point of view. Indeed, students' perception of the use of technologies and digital platforms has been found to be an important factor in the extent of the adoption of technology and the achievement of the learning goals (Mehta et al., 2021), therefore scholars have been called upon to extend the knowledge in this field.

In this study, we analyzed students' reflective journals with the aim of understanding how the students perceived the activity and its ramifications. Content analysis helped identify themes and patterns regarding students' perceptions of the newly designed activity. The findings of this study will help in understanding the benefits of Padlet as a collaborative learning tool, and how it should be implemented so as to develop the ability to work in varied teams, to collaborate in decision-making, and to adapt to changes in the learning environment. With these skills and abilities, students will undoubtedly be better able to adapt to the labor market in health professions and in general (Männistö et al., 2020; Alt et al., 2022).

Literature review

In this section we will review the advantages as well as challenges of OCL. We will also address Padlet-mediated learning in technology-enhanced environments by reviewing relevant recent literature. Further, we will present the use of this teaching methodology in the healthcare professions and examine how it impacts the skills required of today's healthcare students.

Online collaborative learning

Zheng et al. (2014) defined online collaborative learning (OCL) as a goal-oriented activity of a group of students committed to achieving a shared goal and creating new knowledge by learning interactively in a digital environment. This methodology has become widespread because of its many advantages. For example, studies show that encouragement of students' engagement, high accessibility to the course material (Sukendro et al., 2020; Balouchi and Samad, 2021), extensive access to colleagues and instructors, and encouragement of interaction and collaboration among the students contribute to better learning outcomes (Ornellas and Muñoz Carril, 2014). OCL was found to add interest and to raise motivation to conduct research and study, and as such to encourage the transition from passive to active learning, and to encourage interaction and connection between the students through the sharing of the products of their study and new ideas, leading to greater social engagement and contributing to the

development of higher-order intellectual skills (Malik et al., 2017; Park et al., 2019; Yadegaridehkordi et al., 2019).

Various studies (Ng et al., 2022) have explored the factors that affect the degree of students' adoption of OCL. For example, Cole et al. (2021) showed that students' perceptions of the online learning environment were an important predictive factor in their engagement in learning by means of online courses. Studies have also found that the nature of the social interaction in the learning environment, support and positive interaction with the facilitator, and the type of technology used are considered key factors in the degree of adoption of OCL and the extent to which it influenced the achievement of the learning goals and improvement of performance (Molinillo et al., 2018; Qureshi et al., 2021). Other studies evaluated the outcomes of OCL and examined their level and quality. Breen (2015) and DiPasquale (2017), for example, presented means of evaluating the outcomes of OCL by analyzing the initial stages of the establishment of the higher-order learning outcomes by means of brainstorming, group discussion, and presenting various ideas, orally or in writing. Blicek et al. (2019) pointed to the development and encouragement of ideas as outcomes of higher-order learning with this methodology and the guidance and support of the instructor. Similarly, Aldholay et al. (2018) demonstrated the development of higher-order thinking by means of processing, reflection, and analysis of collaborative critical thinking regarding specific topics relevant to the course content. This supportive and collaborative learning environment enables the existence of an effective learning perspective for attaining the maximum learning goals.

Online collaborative learning mediated by Padlet

The Padlet interactive platform constitutes a virtual online board on which users can post various types of multimedia content, such as documents, questions, comments, images, video clips, and audio clips. The platform and the board are available to learners throughout their studies and even beyond (Rosnida and Zainor, 2018; Mehta et al., 2021). Recent studies have shown the contribution of Padlet-mediated learning and its usefulness in encouraging learners by means of active teaching and collaborative learning (Rosnida and Zainor, 2018; Rashid et al., 2019).

According to the study of Gasmi and Thomas (2017), this technology has been seen to promote collaboration among learners and to promote students' engagement in seminar courses. The study found that the technology contributed to the efficiency of their studies and constituted an easy and secure learning environment for communication, teamwork, and asking questions. These studies shed a positive light on the use of Padlet-mediated collaborative learning. They note the advantages of the tool in its ease of use, accessibility to materials over time (even after the conclusion of the course), and work with diverse resources (including links, images, and video clips), which make it a convenient tool for student-focused active collaborative learning (Mehta et al., 2021).

Several studies have demonstrated the benefits of using Padlet as a tool for collaborative learning. Garnham and Betts (2018) found that student engagement increased in seminar courses that utilized Padlet, and Beltrán-Martín (2019) reported that students experienced satisfaction and improved academic performance when using the platform. Padlet has been identified as having several advantages for collaborative learning, including ease of use, visual simplicity and

long-term accessibility of content, even after the end of the course. These advantages of the Padlet created a friendly tool for users that led to high student satisfaction and better performance. Moreover, Padlet is a very easy-to-operate technological tool that students can use with various apps and resources such as presentations, links, images, and videos (Beltrán-Martín, 2019).

In the contemporary academic landscape, the availability of various technological tools has become an essential aspect for students to present their knowledge and communicate their ideas. However, while these multiple options provide numerous opportunities, they may also pose a challenge, as learners must devote significant time and effort to acquire the necessary technological skills. As a result, the Padlet can prove to be advantageous in this regard. Furthermore, these features allow for a holistic and student-centered approach to teaching and learning, as students are able to use their preferred medium and become actively involved in creating their own content (Mehta et al., 2021).

The literature review indicates that Padlet-mediated OCL methodologies are not yet widely used in higher education in the fields of medical education and medical management and that there is little research on how the methodologies are used in higher education. Therefore, the current study aimed to implement an intervention program (described below), that combined Padlet-mediated OCL. Next, it sought to evaluate the perception among health management students of its effectiveness in collaborative learning. Mapping the students' perceptions and understanding the effects of using this technology-mediated methodology is essential for achieving learning goals and acquiring the requisite skills, especially in studies that examine the students' perspectives, because of the difficulty in tracing students' engagement in online learning activities (as opposed to face-to-face learning encounters). The data from such studies can be important for evaluating the effectiveness of the learning methodology (Mehta et al., 2021). The current qualitative study is based on a thematic analysis of health management students' reflective journals that were submitted at the end of the learning process. It presents the challenges and opportunities of the proposed learning method as viewed by the students.

Method

The current study uses the case study method. In this method, as defined by Yin (2014), there is an in-depth empirical investigation of the current phenomenon and its relation to the real world. Data are gathered regarding human activities in a specific time and place, enabling the learning of several things about the processes that take place in the case under research. In a qualitative study in general and in case study research in particular, the researcher constructs themes that arise from the participants' words in such a manner that an argument or generalization derived inductively from the participants' words receives empirical support from quotations (Cresswell, 2009).

Participants

Data for the analysis were gathered from 85 Israeli undergraduate students of a Health Management program (covering patient-doctor relations, quality of service in the healthcare system, and ethics and patient rights). The students' mean age was $M = 25.70$ ($SD = 5.85$), and

84% were females. In relation to ethnicity, 58% were Jews, and 42% Arabs (Muslims). Thirty-one percent reported working in health professions (*Profession*).

Data gathering

The students were asked to keep a reflective journal throughout the course and to submit it at the end of the semester. They were instructed to keep a record of the learning process as they saw it and to describe the challenges, difficulties, and positive outcomes in light of their experience (Tripto et al., 2016). For example, they were asked to describe the main challenges, the difficulties with which they had coped, and the positive and negative experiences of the process. Keeping the journals enabled reflection regarding the significant learning processes and adoption of a multidimensional view of this process (Zohar and Barzilai, 2013).

Data were collected through the reflective diaries, following the intervention plan described in the next section (see the Instructional Method). Prior to its implementation, and for the purpose of maintaining the ethical principles concerning research with humans, the following topics were kept among others: The theme of the research was clarified to the students. It was clarified that the reflective diaries will be used to evaluate the intervention plan. The students' consent was obtained by signing an "informed consent" form. The participants were assured that every name would be replaced by a pseudonym and every identifying item would be omitted when writing the research report. It was made clear to the students that the materials (the reflective diaries) would be processed anonymously only. The research was approved by the college's Ethics Committee.

Procedure of the instructional method (intervention plan)

Underlying this study is the transformation of a required course titled Introducing Quality into the Health System, which had been taught in a traditional frontal manner, to a course using a methodology of Padlet-mediated OCL. This methodology enabled OCL that included work in small groups; building a collaborative data pool; final projects focusing on dilemmas and analyses of them in organizations, using the data pool the students had constructed; presentation of the projects on the collaborative platforms; and peer review (Levenberg and Barak, 2015; Ng et al., 2022). At the beginning of the course, the students were exposed to the new methods of teaching and received information about the Padlet technology and how to use it. The technology was also demonstrated by being used for students' introduction. This introduction included students' presentations of themselves using the Padlet board, thus enabling familiarity with the tool, constituting an informal and equal experience of the tool, and generating interaction between the students to create groups and increase students' engagement in the course (DeWitt et al., 2015; Rashid et al., 2019; Qureshi et al., 2021).

The course consisted of several stages, each one based on the preceding one. After familiarizing themselves with the course outline and the technology, the students were asked to form online study groups with which they would work throughout the semester. In the second stage, the students were asked to choose a topic from among the course

topics presented to them on a Padlet board. In the next stage, the study groups were asked to formulate a question, problem, or dilemma related to the selected topic and to gather information from articles on the topic. The groups were asked to present the articles and analyze them in the context of the selected topic and to post the information on a new Padlet board. This board was organized by topics and subtopics, in accordance with the selected topics. Topics selected included dilemmas in introducing accreditation processes, challenges in maintaining medical confidentiality, clinician-patient relationships through digital tools, and measuring the quality of service in the organization.

By the end of this stage, a categorical board containing information, articles, and links to academic content was conducted. This collaborative information bank was congruent with the course topics and divided in a categorical manner that was easy to navigate and accessible to all the students in all the groups (Mehta et al., 2021). In the next stage, the groups were asked to prepare and present a final project that was posted on yet another Padlet board. In this project, which was based on the preceding work, they were required to broaden the scope of the original project into a comprehensive research question or discussion of a specific dilemma in a specific organization so as to include a number of perspectives, and based on, among other sources, the collaborative knowledge pool of the entire class (and even to extend their project by adding new materials).

As part of the final project, the students were required to present organizational and managerial perspectives as well as the perspectives of patients and health professionals in the system. This was to be accomplished by means of online group meetings and collaborative work, including brainstorming, discussion, and decision-making regarding the construction of the final project and its visual presentation on a Padlet board, in various formats and using various technologies that they chose. These included, for example, e-posters, presentations, and video clips using such technologies as PowerPoint and ThingLink, with which they had become familiar in other courses. This enabled variety and empowerment of the individual student by allowing freedom of choice of the manner of presentation, which is an advantage

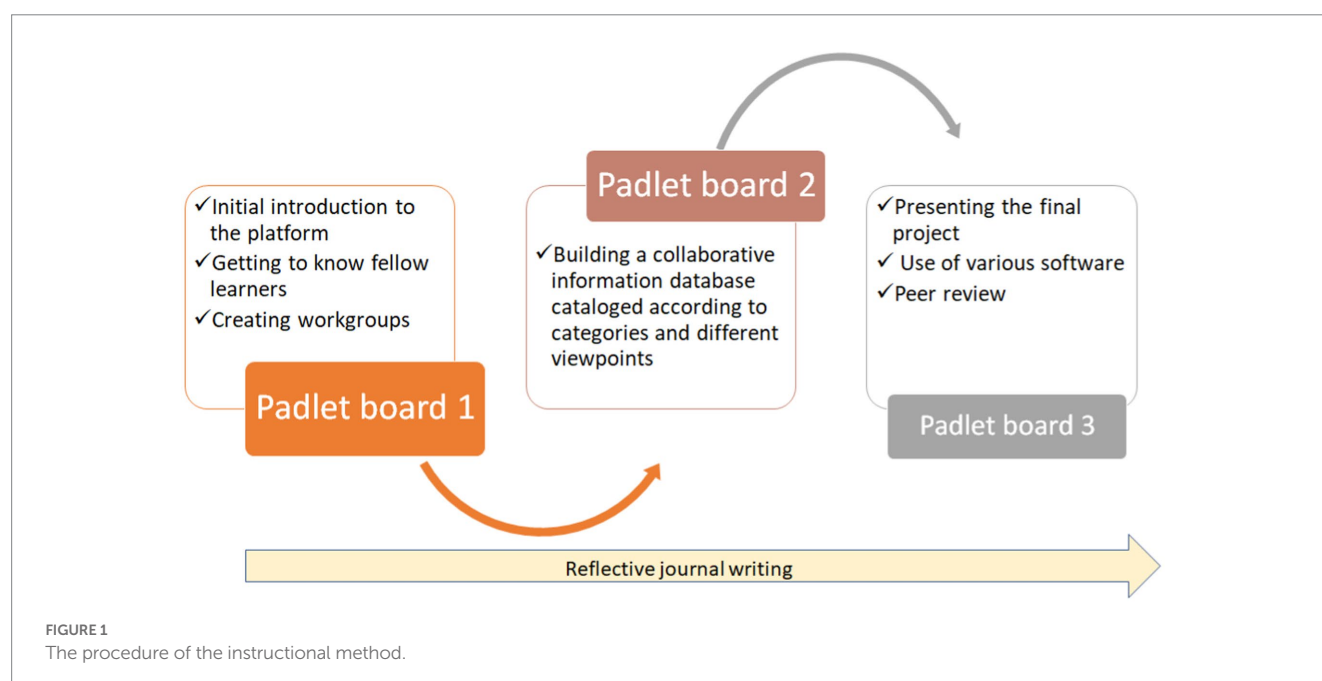
of the Padlet technology (Ramachandiran and Mahmud, 2018). The third board brought together all the projects in the course. The students were required to use the Padlet board to critique their classmates' projects exhibited on that board, ask questions, add information, or highlight topics to be added to the discussion of the dilemma. Figure 1 summarizes the procedure of the instructional method.

Data analysis

This study used a categorical analysis of the data in accordance with the deductive- inductive approach. This combination of approaches throughout the analysis is considered essential because it enables the addressing of categories derived from the literature review and the existing theoretical knowledge in the field combined with new categories that arise from the data analysis (Cresswell, 2009). Thematic analysis was used in this study to identify topics that appeared in the data and to report them. Each theme was summarized to provide an overview of the participants' reports. Then the dominant topics in the reflective journals were identified and were gathered together as main categories (Braun and Clarke, 2006). Disagreements and differences between the researchers were resolved through discussion. The themes regarding which the disagreement was not resolved were presented to a researcher not connected with the study who determined whether the theme was established in the data. Episodes, thoughts, and feelings that the students raised are reported here in order to support the reliability of the themes found. The external validity was derived from the congruence of the themes with the knowledge found in the theoretical literature, as we show in the discussion.

Findings

The data analysis of the reflective journals revealed five main categories: (1) user experience of the platform; (2) visibility and



visibility of the platform; (3) collaborative learning and OCL; (4) active online learning and student engagement; and (5) cognitive flexibility. These categories are summarized in Figure 2. Exemplary quotes from the data are provided to illustrate each category.

User experience of the platform

The analysis of the students' journals revealed that the use of this interface was easy for the students and enabled a positive experience. Most of the students noted the platform's ease of use and the fact that it did not require much learning and experience; after a relatively brief explanation, it was possible to use it freely. In this way, the platform makes the user experience easy and pleasant.

It was very easy in my opinion, and therefore also made it [the Padlet assignments] very enjoyable. They were almost easier and more enjoyable than all the other assignments in other courses. (B22)

Many of the students noted that for them, the platform was almost as experiential as using social media and familiar, favorite applications that they use for online social interaction. This experience contributed to confidence in using the tool that was familiar and informal, reducing resistance and anxiety in carrying out the assignments.

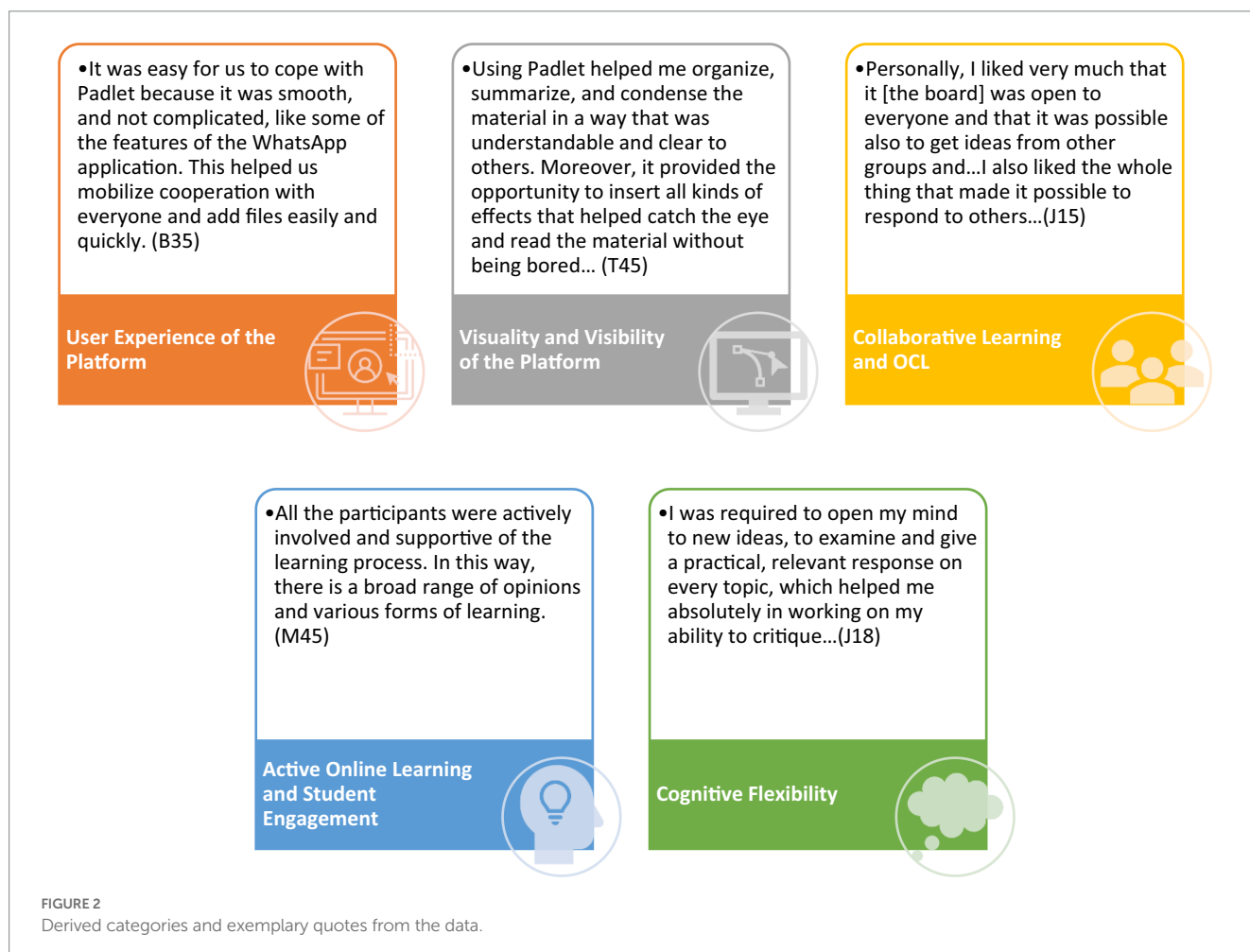
It was easy for us to cope with Padlet because it was smooth, and not complicated, like some of the features of the WhatsApp application. This helped us mobilize cooperation with everyone and add files easily and quickly. (B35)

The students' positive user experience was addressed also in terms of the platform's availability, which was an important element in this category. The fact that the learning and the tools were digital enabled high accessibility to all the material at all times and in every place, including the use of students' phones.

The Padlet was very helpful, I quickly found an article that was relevant to my project and while I was searching, I already understood what I was going to say. The information was accessible every hour of the day. (B11)

Visuality and visibility of the platform

Many students noted positively the viscosity of the platform and the fact that it can be changed according to the students' and instructors' preferences. For example, the background of the Padlet and its organization could be determined by the instructor, but each group may choose the format whereby materials are added.



Using Padlet helped me organize, summarize, and condense the material in a way that was understandable and clear to others. Moreover, it provided the opportunity to insert all kinds of effects that helped catch the eye and read the material without being bored. Colors and the appearance of every post helped me remember and connect [the topic] to an article that needed to be used for the second assignment. In addition, using Padlet helped me very much in choosing the material that was most important for others to read, instead of putting long and boring material. Personally, when I read the posts I was attracted to what drew my eye. (T45)

Nevertheless, a few students noted that the colorfulness and visuality of this tool were perceived as creating visual overload and difficulty in organizing the material.

In my view, the main difficulty was manifested in the visibility of the Padlet. On the one hand it is very friendly and that is colorful and divided into various windows. Here the problem begins, when I felt an overload of information at a certain stage, and it was no longer pleasant to look at and primarily made it difficult to follow the projects of other people. (C12)

As the semester and the students' use of the platform progressed, students reported that they had honed these abilities and that they were able to cope with the ensuing assignments with fewer difficulties in this respect. They noted that these abilities continued to improve with the use of the platform and with comparisons and receipt of help and comments from group members or from members of other groups.

By the final project, we already understood better what is of secondary importance and what is of primary importance and how exactly to condense [it] so that it will catch the eyes of our classmates so that they will want to read our project. (D5)

groups and ... I also liked the whole thing that made it possible to respond to others and to add other things that could enrich the knowledge of the topic of my [project] or to connect to it some kind of creativity, whether a video clip or a survey or a picture. (J15)

In their reflective journals many students noted their initial concern over posting material that included assignments. Some reported fearing criticism. Those who had low self-confidence, were concerned about being criticized because of language difficulties (in a multicultural class room), or were concerned over quality.

There is, for example, concern that a project that is posted on the Padlet is not good enough, and then a feeling of doubt about the work is created and thoughts that "maybe yours is less good than what the others did" begin. (I61)

The students reported that despite their concerns about collaboration, very quickly they were exposed to a variety of projects. The possibility of revising their projects in light of the comparison with other projects or criticism from their classmates had a positive influence on their learning and on the learning product, while relating to other new ideas and other perspectives.

Sharing the projects and assignments enabled the students to critique their own projects and final assignments, compare them, and observe critiques of other projects. This enabled them to revise their own projects and thus they were exposed to a variety of ideas and perspectives as well as the implementation of a variety of abilities, such as self-criticism, and comparison and synthesis of materials.

The Padlet helped me get additional opinions of the project I submitted which would not have happened when I sent the project directly to the lecturer on the website and only the lecturer could see it. This made me understand other additional opinions of other students and to know what things to improve. (Q34)

Collaborative learning and OCL

The experience of collaboration contributed to both the learning level and the social level. Thus, at the learning level, for example, the construction of a collaborative pool of papers and the division of the material and its exposure to all the students made it easier for them in the processes of searching for and locating material and even sorting it.

The pool of articles on Padlet is excellent, broad and truly useful and helpful. Usually when I am looking for an article for other academic projects, just finding the article is a project in itself. Thanks to Padlet, the search is fast and easy. The user experience with Padlet was enjoyable and quick and I even found several articles from which to choose. (F67)

On the social level, the students wrote in their journals that this learning methodology enabled learning in groups and comparison, critiques by peers, and learning from classmates in a positive and less threatening way.

Personally, I liked very much that it [the board] was open to everyone and that it was possible also to get ideas from other

Active online learning and student engagement

Another category that arose from the analysis of the students' reflective journals concerns active learning and active engagement of students throughout the course, using the Padlet digital platform. This categorical analysis revealed that the students found themselves actively involved in the learning processes and even experienced the active engagement of other students.

I feel that I had a very active learning [experience] and that it was not oppressive, which is a very big bonus. (V11)

All the participants were actively involved and supportive of the learning process. In this way, there is a broad range of opinions and various forms of learning. (M45)

In addition, there was brainstorming. That is, all the students participate in the discussion and express opinions and suggestions on a specific topic. (M11)

Students' engagement increased the feeling of commitment to the process and the learning experience, "All the members of the group had a feeling of personal responsibility for the success of their comrades in the group and understand that their personal success depends on the group's success." (C33)

The students also noted that the option of adding comments and communicating about the material enabled brainstorming and discussion in writing, often over time, and led to an open discussion, editing of the material, construction of arguments, and grounding of the arguments in the material studied, individually and even anonymously, "Mostly I would [post] arguments that were based on academic articles. That opened the option of discussing with others, I had such fun using the Padlet" (F31).

Cognitive flexibility

Another category that derives from the proposed theory concerns cognitive flexibility and its development in students throughout the intervention. The students reported being aware of interaction and solution alternatives, considered different opinions while interacting with their classmates, and were open to embrace new perspectives targeted at achieving better results. The students reported that collaborative learning enabled them to experience learning processes that included brainstorming, group discussion, ironing out of disagreements, and exchanges of ideas.

Sharing the information makes it possible to respond to it, to conduct a multi-participant discussion about certain issues in the field of health, to think about new and creative solutions. (K52)

I was required to open my mind to new ideas, to examine and give a practical, relevant response on every topic, which helped me absolutely in working on my ability to critique. Also, this tool exposed me to topics and articles I had not heard about in the past and demonstrated material for me in a broader way by providing a demonstration of the topic as it pertains to medical organizations. All this enabled me to examine information and ways of presentation from perspectives that were different from mine. In addition, during the first assignment we sat as a group and wrote responses together, which enabled us to see together and separately the different ways of thinking. (J18)

Discussion

This study presented the process of introducing and using the methodology of Padlet-mediated OCL with students of health management. It demonstrated how it is possible to guide the OCL process using the Padlet digital internet platform, encourage collaborative learning, and develop cognitive flexibility, through the active engagement of students even in large courses, while achieving the learning goals and outcomes and higher-order thinking among the students.

The analysis of student reflective journals showed that the students found the platform to be user friendly. The term "user experience," which in the past applied mainly to the user interface, that is, the

interface between the technology and the individual, has been broadened and now refers to additional aspects such as design, presentation of information, technology, and even emotional, personal, and experiential aspects (Minichiello et al., 2018). In this context, the students noted that their user experience of the platform was positive, which encouraged them to use it. The platform even increased the interactions among the students, both because of its availability and accessibility and because of the user experience, which included attractive visuality and ease of operation and thus resembled the user experience of familiar platforms that serve in social interactions.

According to the students, the use of the Padlet platform facilitated the establishment of social and learning interactions and motivated them to maintain these relationships. The use of Padlet also promoted a positive and open atmosphere for collaborative learning, allowing for exposure to diverse viewpoints and the cultivation of flexible thinking in collaborative decision-making. This supports the findings of Cole et al. (2021), who showed that students' perceptions of an online learning environment were an important predictive factor for student engagement in learning in online courses. In our study, the participants reported that there was a high level of activity on the part of all the students in the group, including engagement of students from various groups. The assignments were structured in a way that allowed for the accumulation of shared collaborative knowledge, which the students could then draw upon in subsequent assignments. Additionally, the platform provided a positive user experience that contributed to the success of this approach.

Together the students created a sense of collectivity and a shared goal, a desire to share knowledge, and construction of a basis for teamwork and collaboration with the other members of the small group or the entire class. The results of the present study align with previous research suggesting that the nature of interactions with the instructor and peers, as well as the user experience of the technology, are crucial factors influencing the adoption and effectiveness of OCL in achieving learning goals and improving performance (Molinillo et al., 2018; Qureshi et al., 2021). Studies of the Padlet technology specifically, such as Gasmi and Thomas's (2017), found that the platform was perceived as promoting collaboration among the students and student engagement, asking of questions, discussions, and teamwork.

Students, who were accustomed to individual learning, initially struggled with the transition to collaborative approaches. They reported being anxious about sharing their work due to fear of criticism and lack of confidence in their abilities. However, they eventually recognized that this collaboration could be an opportunity for growth and development. With this realization, the students were able to freely provide feedback on their peers' projects and receive comments on their own work, resulting in a positive user experience akin to that of social media rather than a threatening learning environment. These findings align with previous research indicating that OCL can promote higher-order outcomes such as brainstorming, group discussion, idea generation, organization of ideas, and critical thinking through active student participation (Breen, 2015; DiPasquale, 2017; Aldholay et al., 2018).

One particularly notable benefit of using Padlet in large classes is the ability to divide students into groups and facilitate high levels of engagement. This is in contrast to traditional compulsory courses with similar class sizes, which may struggle to allocate sufficient time for group presentations and individualized instruction for the instructor

due to the size of the class (Young and Nichols, 2017). According to student journals and other research on the platform, Padlet enables highly engaged participation (Mehta et al., 2021). Therefore, the combination of this teaching method with the Padlet platform offers a solution to the challenge of effectively instructing large classes in the 21st century. This advantage is particularly relevant in the current context of social distancing measures and the shift towards hybrid learning. The use of Padlet maintains these benefits and also allows for greater accessibility in terms of both time and location, as the boards can be accessed from any internet-enabled device, including smartphones, at any time. This feature is particularly useful in the new era of education.

According to the analysis of student journals, the online methodology and the features of the Padlet platform facilitated the development of cognitive flexibility. The students reported being exposed to a variety of opinions and perspectives from their group members and the entire class, which required them to engage in written discussions, consider alternative ideas, and potentially adjust their project based on feedback from their peers. The processes described by the students align with the concept of cognitive flexibility, which involves the ability to respond adaptively and dynamically to changing circumstances, interactions, and personal experiences in order to achieve a desirable outcome (Hayes et al., 2006). It also involves the consideration of multiple approaches to a problem, negotiation, and the ability to arrive at implementable solutions in complex, varied, and multicultural environments (Boot et al., 2017). This development of cognitive flexibility is also linked to open-mindedness, defined as the willingness to consider alternative possibilities, opinions, and solutions (Barak and Levenberg, 2016).

Indeed, the students reported initial difficulty in this context, which they expressed, *inter alia*, as fear of negative criticism and low valuation, but as the semester progressed and they experienced the advantages of the methodology, this difficulty was overcome and the students began to examine new opinions and ideas, critically and objectively, and even examined their own work by comparing projects, editing, and improving while relating to the critique of their projects by their comrades. In this way they often improved the learning product and achieved higher-order learning outcomes while developing cognitive flexibility.

Limitations of the study and future directions of research

This study utilized online collaborative learning (OCL) and the Padlet platform as tools to support a specific teaching method. While the researchers focused on the pedagogical aspect, future research should compare this digitally supported activity to a non-digital equivalent and measure the impact on students' cognitive skills. This comparison would allow for an assessment of the specific contribution of the digital instrument to skill development and the identification of any challenges encountered. It would also be beneficial to examine the use of this instrument in different teaching methodologies to determine if it is more effective for certain types of assignments or requirements, such as the creation of concept maps or presentations of projects.

This study focused on the field of health management, but the development of these skills at the beginning of professional training in the medical field could benefit health professionals in meeting the demands of contemporary healthcare systems. As we enter a new era

that requires adaptability and flexibility, it is suggested that this pedagogy be implemented in other academic fields to evaluate its impact on the acquisition of relevant skills. The results of this study highlight the need for further research on the use of digital learning environments based on learning outcomes-oriented pedagogy. This may contribute to the development of teaching strategies that foster lifelong learning skills.

Conclusions and recommendations

The dynamic worldwide changes in the digital era and the new needs they are creating have challenged many traditional teaching methods that were widespread in academia in general and in the area of medical education in particular. Digital processes in teaching that were accelerated by the limitations of social distancing during the COVID-19 pandemic started becoming established as essential in these processes, especially because online learning processes mediated by digital platforms will leave their mark in creating the new normal.

These changes have become an opportunity for introducing and establishing innovative and active teaching methods that encourage great engagement of the learner in the learning process, instead of the traditional, passive methods that are widespread in academia in general and the fields of medicine and health management in particular. Consequently, there is a real need today for research and implementation of innovative teaching methods that are suited to the learning goals of the 21st century. Developing these abilities is crucial, especially in the health professions, which are undergoing a host of changes that require the development of these skills at an early stage of training and throughout the more advanced stages, including professional training. Researchers, academics, and educators in higher education are needed to research and actively implement online learning technologies (Mehta et al., 2021).

The findings of the current study support previous studies of the use of Padlet-mediated online learning that were found to be effective in developing engagement and higher-order learning outcomes (Breen, 2015; DiPasquale, 2017; Aldholay et al., 2018) and they demonstrate specifically the effect of implementing Padlet-mediated OCL on the development of learning abilities among students of health management. The findings of this study support assumptions that combining OCL and teaching methods with digital platforms that encourage collaborative learning enable the development and acquisition of cognitive flexibility, teamwork, and higher-order learning outcomes of the kind that enhance crucial abilities of employees in health systems in the 21st century (Männistö et al., 2020). Developing these abilities enables the transformation of individuals into lifelong learners who are better adapted to the changing market conditions and demands of the 21st century. These abilities are crucial for learners and employees in many fields, including the fields of health management and medicine.

This study adds to the existing literature by emphasizing the importance of incorporating digital tools in education, particularly the use of the Padlet platform to facilitate collaborative learning and improve the quality of teaching and learning. It demonstrates that Padlet-mediated online collaborative learning (OCL) can be an effective digital learning tool because of its ease of use and ability to accommodate both pedagogical and technological challenges. This method can be conducted openly or anonymously, depending on the

instructor's preference, and is suitable for students who may be hesitant to participate or fear criticism. Additionally, the accessibility and availability of Padlet make it suitable for use in large classes. It is worth noting that despite the benefits of Padlet-mediated OCL, this methodology may also pose challenges that require careful management and support. This may include gradual integration into the curriculum and ongoing communication between students and the instructor.

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 Hadassah Academic College, Jerusalem, Israel. The patients/participants provided their written informed consent to participate in this study.

References

- Aldholay, A. H., Isaac, O., Abdullah, Z., and Ramayah, T. (2018). The role of transformational leadership as a mediating variable in DeLone and McLean information system success model: the context of online learning usage in Yemen. *Telematics Inform.* 35, 1421–1437. doi: 10.1016/j.tele.2018.03.012
- Alt, D., Naamati-Schneider, L., and Meirovitch, A. (2022). Future Problem-Solving Practiced During COVID-19: Implications for Health Management Students' E-Health Literacy Identity. *Front Psychol.* 293.
- Balouchi, S., and Samad, A. A. (2021). No more excuses, learn English for free: factors affecting L2 learners intention to use online technology for informal English learning. *Educ. Inf. Technol.* 26, 1111–1132. doi: 10.1007/s10639-020-10307-z
- Barak, M., and Levenberg, A. (2016). Flexible thinking in learning: an individual differences measure for learning in technology-enhanced environments. *Comput. Educ.* 99, 39–52. doi: 10.1016/j.compedu.2016.04.003
- Beltrán-Martín, I. (2019) "Using Padlet for collaborative learning." in *HEAD'19. 5th international conference on higher education advances*. pp. 201–211. Editorial Universitat Politècnica de València.
- Blieck, Y., Ooghe, I., Zhu, C., Depryck, K., Struyven, K., Pynoo, B., et al. (2019). Consensus among stakeholders about success factors and indicators for quality of online and blended learning in adult education: a Delphi study. *Stud. Contin. Educ.* 41, 36–60. doi: 10.1080/0158037X.2018.1457023
- Boot, N., Baas, M., van Gaal, S., Cools, R., and De Dreu, C. K. (2017). Creative cognition and dopaminergic modulation of fronto-striatal networks: integrative review and research agenda. *Neurosci. Biobehav. Rev.* 78, 13–23. doi: 10.1016/j.neubiorev.2017.04.007
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Breen, H. (2015). Assessing online collaborative discourse. *Nurs. Forum* 50, 218–227. doi: 10.1111/nuf.12091
- Cole, A. W., Lennon, L., and Weber, N. L. (2021). Student perceptions of online active learning practices and online learning climate predict online course engagement. *Interact. Learn. Environ.* 29, 866–880. doi: 10.1080/10494820.2019.1619593
- Cresswell, J. W. (2009). *Research designs: Qualitative, quantitative, and mixed methods approaches*. California: Sage.
- DeWitt, D., Alias, N., and Siraj, S. (2015). Collaborative learning: Interactive debates using Padlet in a higher education institution, presented at the international educational technology conference (IETC 2015), Istanbul, Turkey. Available at: <https://core.ac.uk/reader/162014460> (Accessed February 1, 2023).
- DiPasquale, J. (2017). Wiki'd transgressions: scaffolding still necessary to support online collaborative learning. *Canadian J. Act. Res.*, 18, 47–61. doi:10.33524/cjar.v18i3.355
- Garnham, W. A., and Betts, T. (2018). The Padlet project: transforming student engagement in foundation year seminars. *J. Learn. Teach.* 11:2. doi: 10.21100/compass.v11i2.714
- Gasmi, A. A., and Thomas, M. (2017). "Academic writing in the flipped EFL classroom: a case study on student engagement in Oman," in *Flipped instruction methods and digital Technologies in the Language Learning Classroom*. eds. J. P. Loucky and J. L. Ware (Pennsylvania: IGI Global), 232–251.
- Grosser, J., Bientzle, M., and Kimmerle, J. (2020). A literature review on the foundations and potentials of digital teaching scenarios for interprofessional health care education. *Int. J. Environ. Res. Public Health* 17:3410. doi: 10.3390/ijerph17103410
- Hayes, S. C., Luoma, J., Bond, F., Masuda, A., and Lillis, J. (2006). Acceptance and commitment therapy: model, processes, and outcomes. *Behav. Res. Ther.* 44, 1–25. doi: 10.1016/j.brat.2005.06.006
- Levenberg, A., and Barak, M. (2015). A table for four: Collaborative writing in shared cloud documents. In *9th Chais Conference. The Open University of Israel, Raanana, February*. Available at: https://scholar.google.co.il/citations?view_op=view_citation&hl=en&user=GfelyEgAAAAJ&citation_for_view=GfelyEgAAAAJ:UeHWp8X0CEIC
- Malik, M., Fatima, G., Abid Hussain, C., and Sarwar, A. (2017). E-learning: students' perspectives about asynchronous and synchronous resources at higher education level. *Bull. Educ. Res.* 39, 183–195.
- Männistö, M., Mikkonen, K., Kuivila, H. M., Virtanen, M., Kyngäs, H., and Kääriäinen, M. (2020). Digital collaborative learning in nursing education: a systematic review. *Scand. J. Caring Sci.* 34, 280–292. doi: 10.1111/scs.12743
- Mehta, K. J., Miletich, I., and Detyna, M. (2021). Content-specific differences in Padlet perception for collaborative learning amongst undergraduate students. *Res. Learn. Technol.* 29:2551. doi: 10.25304/rlt.v29.2551
- Minichiello, A., Hood, J. R., and Harkness, D. S. (2018). Bringing user experience design to bear on STEM education: a narrative literature review. *J. STEM Educ. Res.* 1, 7–33. doi: 10.1007/s41979-018-0005-3
- Molinillo, S., Aguilar-Illescas, R., Anaya-Sánchez, R., and Vallespín-Arán, M. (2018). Exploring the impacts of interactions, social presence and emotional engagement on active collaborative learning in a social web-based environment. *Comput. Educ.* 123, 41–52. doi: 10.1016/j.compedu.2018.04.012
- Naamati Schneider, L. (2020). Strategic management as adaptation to changes in the ecosystems of public hospitals in Israel. *Isr. J. Health. Policy Res.* 9, 1–11. doi:10.1186/s13584-020-00424-y
- Ng, P. M., Chan, J. K., and Lit, K. K. (2022). Student learning performance in online collaborative learning. *Educ. Inf. Technol.* 27, 8129–8145. doi: 10.1007/s10639-022-10923-x
- Ornellas, A., and Muñoz Carril, P. C. (2014). A methodological approach to support collaborative media creation in an e-learning higher education context. *Open Learn. J. Open Dist. e-Learn.* 29, 59–71. doi: 10.1080/02680513.2014.906916
- Park, C. W., Kim, D. G., Cho, S., and Han, H. J. (2019). Adoption of multimedia technology for learning and gender difference. *Comput. Hum. Behav.* 92, 288–296. doi: 10.1016/j.chb.2018.11.029

Author contributions

LN-S and DA: conceptualization, data curation, methodology, writing – original draft preparation, and writing – reviewing and editing. All 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.

- Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., and Yousufi, S. Q. (2021). Factors affecting students' learning performance through collaborative learning and engagement. *Interact. Learn. Environ.* 1–21, 1–21. doi: 10.1080/10494820.2021.1884886
- Ramachandiran, C. R., and Mahmud, M. M. (2018). "Padlet: a technology tool for the 21st century students skills assessment" in *International Conference on Educational Awareness and Policy (ICEAP) Proceeding*, Vol. 1.
- Rashid, A. A., Yunus, M. M., and Wahi, W. (2019). Using Padlet for collaborative writing among ESL learners. *Creat. Educ.* 10, 610–620. doi: 10.4236/ce.2019.103044
- Reis, S. (2019). The doctor in the digital age: competencies needed and a road map for their achievement [review]. *Harefuah* 58, 648–653.
- Rosnida, A. D., and Zainor, Z. (2018). "Padlet as an educational tool: pedagogical considerations and lessons learnt" in *Proceedings of the 10th international conference on education technology and computers (ICETC '18)*, pp. 156–162. Association for Computing Machinery.
- Senkubuge, F., Modisenyane, M., and Bishaw, T. (2014). Strengthening health systems by health sector reforms. *Glob. Health Action* 7:23568. doi: 10.3402/gha.v7.23568
- Sukendro, S., Habibi, A., Khaeruddin, K., Indrayana, B., Syahrudin, S., Makadada, F. A., et al. (2020). Using an extended technology acceptance model to understand students' use of e-learning during Covid-19: Indonesian sport science education context. *Heliyon* 6:e05410. doi: 10.1016/j.heliyon.2020.e05410
- Tripto, J., Ben-Zvi Assaraf, O., Snapir, Z., and Amit, M. (2016). The 'what is a system' reflection interview as a knowledge integration activity for high school students' understanding of complex systems in human biology. *Int. J. Sci. Educ.* 38, 564–595. doi: 10.1080/09500693.2016.1150620
- Xiong, P., Zhang, J., Wang, X., Wu, T., and Hall, B. (2016). Effects of a mixed media education intervention program on increasing knowledge, attitude and compliance with standard precautions among nursing students: a randomized controlled trial. *Am. J. Infect. Control* 45, 389–395. doi: 10.1016/j.ajic.2016.11.006
- Yadegaridehkordi, E., Shuib, L., Nilashi, M., and Asadi, S. (2019). Decision to adopt online collaborative learning tools in higher education: a case of top Malaysian universities. *Educ. Inf. Technol.* 24, 79–102. doi: 10.1007/s10639-018-9761-z
- Yin, R. K. (2014). *Case study research design and methods*. 5th Edn. California: Sage.
- Young, S., and Nichols, H. (2017). A reflexive evaluation of technology-enhanced learning. *Res. Learn. Technol.* 25:1998. doi: 10.25304/rlt.v25.1998
- Zheng, L., Chen, N.-S., Huang, R., and Yang, K. (2014). A novel approach to assess collaborative learning processes and group performance through the knowledge convergence. *J. Comp. Educ.* 1, 167–185. doi: 10.1007/s40692-014-0009-7
- Zohar, A., and Barzilai, S. (2013). A review of research on metacognition in science education: current and future directions. *Stud. Sci. Educ.* 49, 121–169. doi: 10.1080/03057267.2013.847261



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña,
Spain

REVIEWED BY

Olga Chernikova,
Ludwig Maximilian University of Munich,
Germany
Cristina Honrubia,
Universidad Isabel I de Castilla,
Spain

*CORRESPONDENCE

Rivi Frei-Landau
✉ rivipsy@gmail.com

SPECIALTY SECTION

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

RECEIVED 22 January 2023

ACCEPTED 09 March 2023

PUBLISHED 06 April 2023

CITATION

Frei-Landau R and Levin O (2023) Simulation-
based learning in teacher education: Using
Maslow's Hierarchy of needs to conceptualize
instructors' needs.
Front. Psychol. 14:1149576.
doi: 10.3389/fpsyg.2023.1149576

COPYRIGHT

© 2023 Frei-Landau and Levin. 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.

Simulation-based learning in teacher education: Using Maslow's Hierarchy of needs to conceptualize instructors' needs

Rivi Frei-Landau ^{1,2*} and Orna Levin ¹

¹The School of Education, Achva Academic College, Shikmim, Israel, ²The Department of Education and Psychology, the Open University, Ra'anana, Israel

Introduction: Simulation-based learning (SBL) has become an effective tool in the education field, and instructors play a significant role in leading it. Although much is known about participants' needs, SBL instructors' needs have yet to be addressed. The study's goal was to explore SBL instructors' needs while guiding an SBL workshop using the psychological framework of Maslow's Hierarchy of Needs.

Methods: We employed a qualitative case-study design, consisting of 68 four-hour-long workshops, held at a teacher-education simulation center by the center's professional instructors. Data collection comprised 211 statements derived from instructors' open-ended reflections, the transcripts of two focus groups held with the instructors, and 98 interpersonal communication documents.

Results: Data were analyzed using both deductive and inductive thematic analysis, which rendered 11 themes spanning Maslow's five levels, and revealing two possible simulation-based learning paths: a complete process, in which all needs are met and an incomplete process, in which needs remain unmet.

Discussion: Theoretical insights and practical implications are provided for attending to instructors' needs (i.e., basic needs, security, belonging, self-esteem, self-actualization) to ensure optimal learning in teacher education when using SBL.

KEYWORDS

simulation-based learning, teacher education, instruction, Maslow's Hierarchy of needs, qualitative case study

1. Introduction

The use of simulation-based learning (SBL) in teacher education has increased rapidly over the last decade (Ferguson, 2017; Thisgaard and Makransky, 2017; Ledger et al., 2022), and SBL instructors play a significant role in leading it. Although much is known about workshop participants' needs, instructors' needs have yet to be addressed, particularly in the field of teacher education. This lack of knowledge may result in a situation in which instructor's needs are overlooked, thus, impeding their ability to perform optimally. Hence, to address this literature gap, the goal of the current study was to explore instructors' needs when leading an SBL workshop and to identify and conceptualize them using Maslow's Hierarchy of Needs, a well-established framework regarding human needs. Enhancing our theoretical understanding of the understudied phenomenon of instructors' needs is imperative, given their significant role in facilitating the workshop goals; thus, addressing their needs is a key factor for the successful implementation of SBL in teacher education. This view coincides with previous claims (e.g., Snoek et al., 2010) that identify the teacher educator as the figure with the greatest impact on the quality of teacher education.

1.1. Theoretical background

1.1.1. Simulation-based learning

Simulation-based learning is a novel methodology that evolved from health disciplines (Zhong and Zhang, 2021), and authentically simulates the conditions of the professional arena at hand for learning purposes such as teaching skills (Chernikova et al., 2020; Ross and Atkinson, 2020; Malone and Brünken, 2021; Frei-Landau et al., 2022; Muchnik-Rozanov et al., 2022). It is accompanied by a reflective debriefing (Garden et al., 2015) and guided by an instructor, during which time constructive feedback is given, with the aim of improving future performance (Girod and Girod, 2008).

In its early stages, the term ‘simulation’ was used to describe role playing, whereas over the last two decades it has been expanded to include advanced designs that involve professional actors and the latest technology (Frei-Landau and Levin, 2022). For instance, *computer-based* simulations, such as augmented-reality environments use video devices that imitate the physical world (Enyedy et al., 2015; Tang et al., 2021) and *human* simulations - often termed ‘clinical simulations,’ involve professional actors (Dotger, 2010). Despite the diversity of SBL designs, they share several common features: learning in a safe environment but under conditions that resemble those of the professional field in question (Rayner and Fluck, 2014), learning from mistakes under conditions that involve no risk or deleterious impact (Kaufman and Ireland, 2016), and learning through active and reflective experience (Butvilofsky et al., 2012).

1.1.2. Simulations in the field of teacher education

Until recently, SBL has been used mostly in medical and health education (Fischetti et al., 2021). Although its use in the field of teacher education is still in development (Shapira-Lishchinsky, 2013; Levin and Flavian, 2020; Frei-Landau et al., 2022), it has been purported to help teacher-education graduates to be better prepared for facing future challenges (Dieker et al., 2014; Kaufman and Ireland, 2016), by providing an opportunity to apply theoretical knowledge to behavior (Theelen et al., 2019). As such, simulations are recommended and implemented in initial teacher education (Ledger and Fischetti, 2020; Ledger, 2021; Ade-Ojo et al., 2022), and recent reviews have demonstrated SBLs pedagogical benefits, specifically, promoting cognitive, behavioral, and affective learning processes (Vlachopoulos and Makri, 2017).

Theoretically speaking, learning through simulations is grounded in sociocultural frameworks for embodied cognition (Danish et al., 2020). Specifically, according to Vygotsky’s Zone of Proximal Development (ZPD in Vygotsky, 1978), learners acquire skills *via* guidance and encouragement of a more skilled partner—in the case of SBL—the instructor. In fact, the SBL instructor guides a reflective debriefing during the learning process. This is in line with Dewey’s suggestion that educators undertake reflective practices for self-development, as well as improved self-confidence and self-efficacy. To this end, current simulation studies rely on these theoretical frameworks (Fischetti et al., 2021).

Studies focusing on SBL outcomes have demonstrated the affordances of SBL in the field of teacher education, which include the acquisition of skills specific to teacher education (Theelen et al., 2019) and special education (Larson et al., 2020); the opportunity to formulate cognitive knowledge (Ferguson, 2017; Dalinger et al., 2020);

and the promotion of learners’ reflective abilities (Levin and Flavian, 2020) and critical thinking (Codreanu et al., 2020) in routine practice, as well as in the context of ethical dilemmas in education (Yablon et al., 2021). However, it is important to note that studies have also indicated that there are also drawbacks to SBL. For instance, in the course of the simulation, learners can experience negative feelings, such as embarrassment and demotivation (Bautista and Boone, 2015; Dalgarno et al., 2016), as well as learning overload (Makransky et al., 2019), each of which might hinder the learning process. Given these findings, the efficacy of SBL in the field of teacher education requires further study. Interestingly, SBL advantages and disadvantages have been examined from participants’ perspectives but rarely from those of instructors (Judge et al., 2013).

1.2. The significant role of the instructor in simulation workshops

As the reflective debriefing phase is the most significant component of SBL (Tutticci et al., 2018) and given that the instructors are the ones to lead it, their needs should be examined. Furthermore, as the quality of instruction affects the ultimate efficacy of SBL (Garden et al., 2015), providing instructors with optimal conditions is imperative. In this sense, the instructor’s role coincides with the teacher educator’s role (i.e., guiding the learner through the learning process). Hence, in accordance with prior arguments that the teacher educator is the figure with the greatest impact on the quality of teacher education (Snoek et al., 2010), SBL instructors’ needs should be examined. In other words, acknowledging that it was important to study the needs of teacher educators in the field of teacher education (Van Velzen et al., 2010; Ghouseini, 2017) given the critical role they play (Orland-Barak, 2014), we must conclude that it is equally important to address the needs of simulation instructors in the context of SBL in teacher education.

The literature in the field of medical simulations indicates that the instructors deal with complex challenges, such as creating a supportive group climate (Tutticci et al., 2018) and managing the technical operation of equipment (Sellberg, 2018), especially the video-system equipment. By contrast, the challenges and needs of simulation instructors in the field of teacher education have yet to be precisely identified and mapped. This gap in the literature is surprising, considering the vital role of instructors during SBL. Therefore, expanding our understanding of instructors’ needs by mapping them is expected to provide insight into ways to optimize the instructors’ functioning (Der-Sahakian et al., 2015; Campbell and Daley, 2017). In the current study, we addressed this overlooked aspect, by examining SBL instructors’ needs through the lens of Maslow’s Hierarchy of Needs (Maslow, 1943).

1.2.1. The theoretical framework: Maslow’s Hierarchy of needs

Maslow’s Hierarchy of Needs (MHN) deals with the essence of human needs and motivations (Fives and Mills, 2016) and provides a hierarchical scale onto which such needs can be mapped. According to Maslow, this hierarchy ranges from basic to complex needs, as follows: (1) *Basic needs* include physiological needs such as hunger, thirst, and sleep, the fulfilment of which are necessary for survival. From the perspective of scholars in the field of education

(Riley and Mort, 1981; Fisher and Royster, 2016), the equivalent of this level of basic need in the context of the education field also includes tangible equipment and supplies as well as technology-related needs (Bailey and Pownell, 1998); (2) *Safety needs* involve a sense of security and protection from physical as well as emotional harm (freedom from threat and the need for self-preservation); (3) *Belongingness needs* involve the desire for acceptance, affection, and friendship, as well as the ability to exist in harmony with others; (4) *Esteem needs* consist of two elements: the need to feel competent, strong, and successful, and the need for recognition, appreciation, and a positive reputation (Maslow, 1943 in Adams et al., 2015); and (5) *Self-actualization needs* include the realization of one's goals/ambitions (Shaughnessy et al., 2018), and are gratified when individuals reach their potential.

Maslow's main claim was that 'The appearance of one need usually rests on the prior satisfaction of another' (Maslow, 1943, p. 370). Hence, the satisfaction of basic existential needs is a preliminary condition for engaging in processes of more complex growth (such as learning). It may be for this reason that Maslow's theory is used so extensively in the field of education (Fives and Mills, 2016).

Maslow's theory reached the height of its popularity in the 1960s and 70s. From the end of the 1960s to the mid-1980s, numerous studies applied the theory to gain an understanding of candidates' motivations for pursuing a career in the field of education (Coon, 2006), the needs of preservice teachers when learning, and the desired components of teacher-education programs (Riley and Mort, 1981). Between the mid-1980s and the end of the 1990s, the theory was rarely referenced in studies.

In the last two decades, there has been a renewed interest in Maslow's theory, with a focus on contemporary issues. Maslow's theory has been found to be a suitable framework for understanding schoolteachers' needs (Fisher and Royster, 2016), as well as scholars' needs in higher education (Schulte, 2018). Furthermore, in the field of teacher education, it was used for conceptualizing instructors' process of self-actualization through mentoring (Fletcher, 2006). Thus, in line with these developments, as well as the findings of a preliminary pilot study that highlighted MHN as a relevant theoretical framework for mapping instructors' needs (see the data analysis section), we chose to use MHN as a theoretical lens for exploring the overlooked issue of SBL instructors' needs. Surprisingly, no studies about simulation instructors' needs exist despite the centrality of instructors in the SBL process as the debriefing facilitator. Providing insight into SBL instructors' needs is crucial, as meeting people's needs is linked with their success (Freitas and Leonard, 2011). In conclusion, considering the lack of empirical information regarding the needs of SBL instructors' in teacher education, and the importance of this information given the instructors' central role in facilitating the workshop goals, the current study aimed to explore SBL instructors' needs in the teacher education arena.

1.3. Research questions

As mentioned, not only is it crucial to identify SBL instructors' needs, but once identified, we must be prepared to interpret and evaluate their impact, to understand their relative significance to the overall SBL process. To this end, we formulated the following research questions:

- (1) Instructors' perceived needs and their mapping according to MHN: what are the needs of simulation instructors when leading a simulation workshop, and in what ways do these needs correspond to Maslow's Hierarchy of Needs theory?
- (2) Instructors' perceived SBL process in light of their needs' gratification: What is the perceived impact of gratification of the needs in the hierarchy vis-à-vis the SBL process overall? What is the perceived impact of unmet needs on the SBL process?

2. Methodology

2.1. The study context

The research was conducted at a simulation center at a higher-education institution in Israel that operates in accordance with the clinical simulation model. This model involves professional actors (as opposed to virtual simulations) and is frequently used by Dotger (2015) to enhance teacher preparation. Generally, workshops are directed toward promoting communication skills, teamwork, and multiculturalism (Yablon et al., 2021; Frei-Landau and Avidov-Ungar, 2022; Levin et al., 2022).

The center is located in the school of teacher education and its content is coordinated with the teacher education program. It provides services to both preservice and inservice teachers in separate groups. Inservice teachers participate in SBL as part of their professional development courses, whereas preservice teachers typically participate in SBL as part of a course in pedagogy. The students attending the center represent a multicultural cross-section of Israeli society, which includes Israeli Jewish students (secular and religious alike), Israeli Arab students (mostly from local Bedouin settlements), and new immigrants.

The current study was carried out over 2 years. The staff of the simulation center consists of 29 instructors and 14 professional actors. Workshops are guided by a single instructor, who has received special training for this purpose. Most of the instructors are the college lecturers, whereas paucity hold a background and contemporary employment that is highly related to teacher education. Specifically, most of them are the college's pedagogical teacher educators, whereas some are senior schoolteachers or hold a managerial position in their school. In addition to the specialized training they receive at the simulation center, they usually hold prior additional certifications as group mentors or coaches; thus, they have education-oriented backgrounds as well as field experience leading groups. The simulation center's staff, including the simulation instructors, the center's administrators and the research unit meet regularly to enable knowledge sharing and brainstorming on how to best design and improve the quality of the teacher-education simulations. Each simulation workshop lasts 4 h and comprises approximately 12–15 participants. Pre-workshop preparation includes an in-depth conversation held between the group's coordinator and a center staff member, during which the main needs of the group are identified. Next, the instructor receives additional information about the incoming group's characteristics; then, tailored scenarios are written by the center's professional content developers (scenario writers). Each workshop consists of three different scenarios typically involving conflictual situations that are common in the field of education. Table 1

TABLE 1 A sample of simulation scenarios: Contexts, descriptions, goals, and key issues for instructor's debriefing.

Simulated context	Scenario's figure characters	Scenario description	Summary of key issues for instructor's debriefing
Teacher*-student	Dan, 13 years old, diagnosed with ADHD	Teacher-initiated conference to address student's classroom disruptive behaviors (chattering, inattention, etc.) and learning difficulties	Facilitating cooperation Finding empathy Evoking motivation for learning Assertiveness
	Sara, 14 years old, has recently entered junior high school	Teacher-initiated meeting to discuss student's social behaviors (social-media shaming)	Active listening Inquiring through open questions Expressing empathy Assertiveness
Teacher*-parent	Lea, 45 years old, a lawyer, mother of a 10-year-old student diagnosed with ASD	Parent-initiated conference to discuss with the teacher the problems of the inclusion process of her fifth-grade daughter with autism	Facilitating trust Legitimizing emotions Dealing with antagonism Collaborative problem-solving
	Jacob, 55 years old, a delivery man, father of a nine-year-old student	Teacher-initiated conference to discuss student's off-task behaviors and low grades, aiming to suggest professional assessment	Facilitating trust Managing goal-oriented conversation Managing resistance Legitimizing emotions
School principal – teacher*	Kim, an experienced school principal, a strong believer in parent-school cooperation	School-principal-initiated conference to acquaint the teacher with the school's educational philosophy, procedures, and policies, following a parent's complaint	Accepting criticism Discussing expectations Facing an authority figure Managing impressions
	Alex, an experienced school principal, highly ambitious and eager to prove himself	Teacher-initiated conference to discuss the teacher's overload with extra school assignments	Assertiveness alongside empathy Exhibiting problem-based attitude (rather than emotion-based attitude) Managing a collaborative dialogue

*The volunteer participant.

illustrates a sample of simulation scenarios, presenting the contexts, descriptions, goals, and key issues for instructors' debriefing.

The SBL workshop process involves three experiential cycles that follow the same procedure but are diverse in terms of content. Each of these cycles includes a scenario-based simulative experience – that is, a five-minute enactment of the scenario by professional actors interacting with a group member in real-time, during which the actor responds according to predefined schemes written into the scenario – followed by a debriefing phase. The interactions are video-recorded (using the SimBoost platform), with the assistance of a technical support team, and the debriefing phase is conducted by the instructor, during which a variety of theoretical coping-strategy models are presented. Finally, as part of the debriefing process, the actors present their feedback to the workshop participants, providing the participants with a unique opportunity to receive direct and authentic feedback on how their behavior was experienced by others. Figure 1 presents a step-by-step description of the simulation workshop process, emphasizing the role of the instructor as well as the stages of the debriefing phase, as these are major components of SBL.

2.2. The study design

Gaining insight into the subjective needs of instructors called for a qualitative methodology (Creswell, 2007). In this domain, a qualitative case-study design was selected (Stake, 2013). A case-study involves the exploration of 'a real-life, contemporary bounded system

(a case), through detailed, in-depth data collection (Creswell, 2007, p. 97). In the current study, all 68 SBL workshops were conducted at a single site (the simulation center) and all dealt with promoting communication skills within conflictual situations that are typically inherent to the teaching experience. Hence, the SBL workshop can be considered a case through which instructors' needs could be examined. The goal of this approach was to perform an in-depth case analysis of instructors' needs while they were managing an SBL workshop. Thus, the identical structure of the workshop, which was used consistently to train teachers (albeit of different disciplines and varying levels of experience) served as an optimal setting for thoroughly exploring SBL instructors' needs. As per Creswell's (2007) guidelines for case-studies, we gathered a variety of data using several information sources, to provide an in-depth view of the case, which was limited in terms of time and or space.

2.3. Participants and data collection

Employing a purposive sampling method, 29 simulation center instructors were approached, yet it was emphasized that participation was voluntary. Eventually, 22 instructors agreed to participate. Data from 68 workshops were collected over a two-year period (28 from the first year; 40 from the second year), using multiple data sources, to provide a comprehensive understanding of the issue, allow for the triangulation of findings (Bogdan and Biklen, 2007), and the performance of cross-validity checks (Creswell, 2007). The following data sources were used.

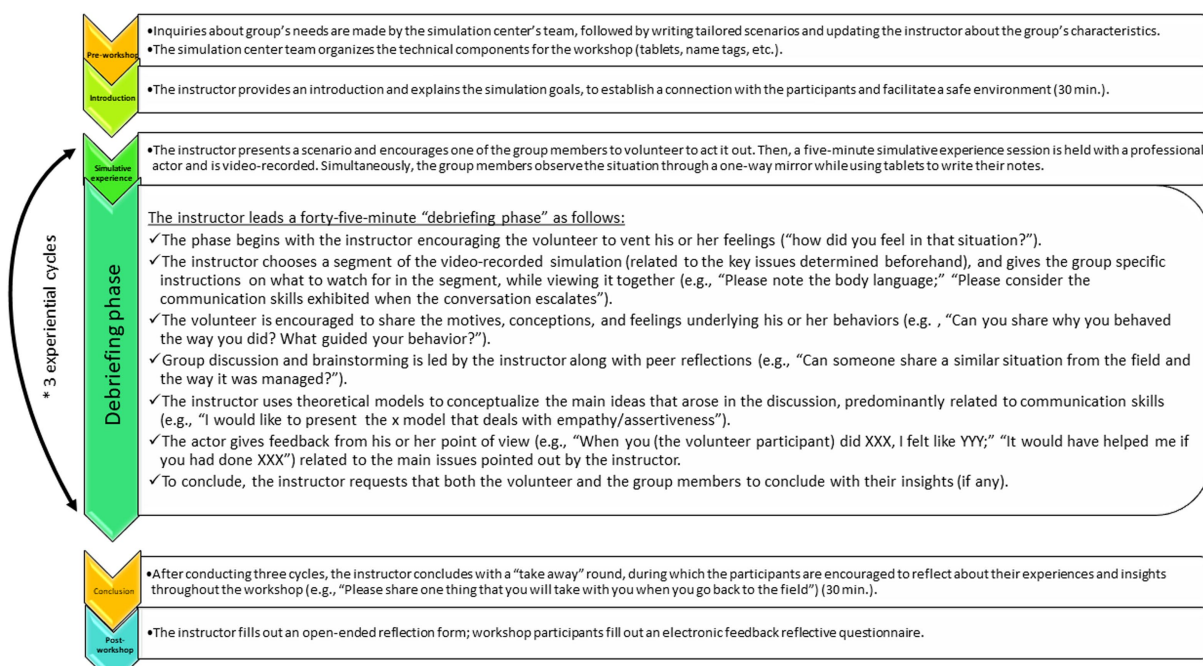


FIGURE 1

Instructor's role throughout the SBL workshop. A step-by-step description of the simulation workshop process, emphasizing the role of the instructor throughout the process. [Note that although all workshops have the same structure, they differ in terms of their contents (different scenarios)].

- (1) *Instructors' post-workshop reflections.* Following each workshop, instructors were asked to freely elaborate about their experiences during the simulation workshop (including perceptions, challenges, etc.). Overall, 68 reflections were collected, yielding 211 statements.
- (2) *Focus groups.* Two focus groups were conducted at the end of each year (10 participated in the first-year focus group and 12 participated in the second-year focus group). During the focus groups, the participants were requested to discuss their needs when leading a workshop as well as to respond to others' comments (sample questions: 'We wish to learn more about your needs as instructors when leading a simulation workshop. Could you please share your experience in this regard?'; 'Please describe what promotes or impedes your needs being met in the context of leading an SBL workshop?'; 'How does having your needs met affect the learning process and workshop overall?'; etc.). Focus group sessions lasted one and a half hours and were audio-taped and then transcribed.
- (3) *Instructors' email correspondence documents.* Documents containing email correspondence between the center's academic manager and the 22 instructors were collected. Overall, 98 documents were reviewed.

The rationale for collecting data from multiple sources lies in the assumption that richer information will be derived this way. The inclusion of personal reflections are recorded privately and convey thoughts and emotions that might not be shared in a group. By contrast, focus groups are conducted collaboratively and publicly, where interpersonal interactions can trigger new insights and perspectives. The rationale for examining email correspondences between instructors and academic managers was that this medium is often used by

instructors to discuss challenges and experiences following an SBL workshop, thus, providing important data about the instructor's needs.

Prior to data analysis, all data were coded as follows: instructors' reflections were coded sequentially (REF-W1—REF-W68); focus group data were coded in accordance with the year in which they were conducted (FOGR1, FOGR2); and instructors' correspondence documents were coded chronologically (ECO1—ECO98).

2.4. Data analysis

Data from a preliminary pilot study of simulation instructors (comprising a focus group, $N=8$; and 26 email correspondences) revealed that the instructors' needs could be mapped to the MHN framework. Thus, the MHN framework was chosen as the theoretical lens through which the entire data were analyzed and through which the issue of needs was examined. This analysis was conducted in a two-step thematic analysis procedure, as follows. In the first round, we performed a deductive analysis, using MHN as a theoretical framework through which to investigate the perceived needs of the instructors when leading simulation workshops. At this stage, data from all sources were coded based on Maslow's five levels of human needs, while we remained open to additional levels or needs not mentioned in MHN. We calculated the frequency with which each theme appeared in each of Maslow's levels, across all data sources, to gain a sense of the prevalence of the various themes. During the counting procedure, we took note of instances where the same instructor mentioned the same workshop-related theme in different data sources. In such cases, the theme in question was counted only once.

In the second round, we employed an open thematic analysis, which is inductive in nature. During this inductive analytic process,

data segments that were identified in the previous round as reflecting instructors' needs were analyzed, as we searched for salient themes emerging from the text. Similar or related themes were grouped together. This microanalysis was used to ensure that no important ideas, themes, or constructs were overlooked. This process yielded 11 main themes along Maslow's five levels. Eventually, we were able to organize the emergent themes that represented instructors' needs according to MHN. Then, we calculated the frequency with which each theme appeared in each of Maslow's levels across all data sources, to gain a sense of the prevalence of the various themes. During the counting procedure, we took note of instances where the same instructor found a particular workshop-related theme in more than one data source (e.g., in the reflections and in the email correspondence). In such cases, the theme in question was counted only once. Table 2 presents an example of the qualitative content analysis coding process.

Working separately, we (the two researchers) initially read and analyzed the material, marking and coding the relevant themes and subthemes, and then ascribed each of these themes to one of the five levels of MHN. Subsequently, we compared our findings and discussed them until we reached a joint decision regarding the relevant themes at each level. Finally, member checking, which is considered the gold standard for establishing trustworthiness (Kornbluh, 2015), was employed, by sharing the findings with the instructors and incorporating their suggestions into our interpretations.

2.5. Procedure and ethics

This study's protocol was approved by the higher-education institution's Ethics Committee. The participants gave their informed consent. The request that they write their reflections was sent to instructors *via* a link to a Google Form, enabling them to refuse without being inconvenienced. Confidentiality was maintained, as answers were anonymous. Before conducting the focus group, the instructors gave their consent to have the session recorded, and they were explicitly told that they could choose what they wished to share, and were free to leave at will, without consequence. Focus groups were recorded and transcribed; all personal information was concealed and, hence, pseudonyms are used herein.

3. Findings

First, we will describe the themes that emerged as related to instructors' needs during SBL, while considering the ways in which these themes correspond to MHN. Then, we will focus on two possible SBL pathways that demonstrate SBL outcomes when needs are met as opposed to when they are unmet.

3.1. Instructors' needs as mapped onto MHN

The findings unveiled the various themes that represented instructors' needs and the ways in which these needs resembled those of MHN. Figure 2 presents the integration of these 11 themes into the five hierarchical levels.

3.2. Basic needs

Based on the analysis, 39% of the instructors' reports (82 phrases) described needs that corresponded to basic needs on the MHN framework. That is, the instructors indicated that these items were essential for them to be able to perform their duties during workshops. These comprised both generic needs and specific simulation-workshop-related needs, as elaborated below.

3.2.1. Generic workshop equipment

The *equipment* needed for operating the workshop was mentioned as a primary and significant concern: 'The binders with the forms were not well-organized. I had to search for the correct form ... We had to do last-minute photocopying ... how am I meant to manage the workshop efficiently like this?' (REF-W31). The presence or absence of technical equipment was experienced as enabling or impeding the instructor's operation of the workshop. Judging from the tone of these comments, it appears that having equipment-related difficulties was a frustrating experience.

3.2.2. Specific SBL workshop components

Other basic components needed were essential to the simulation, specifically the video recording and the scenario. First, *being able to use the video* system properly during the debriefing process was experienced as essential to the instructors' functioning: 'Because of technical problems in the video system, I was unable to use it throughout the debriefing stage. It complicated everything for me' (ECO2). In such cases, instructors noted that the assistance of a technical support team was beneficial.

Second, an *adequate and adapted scenario* was experienced as fundamental to instructors' ability to successfully lead the workshop.

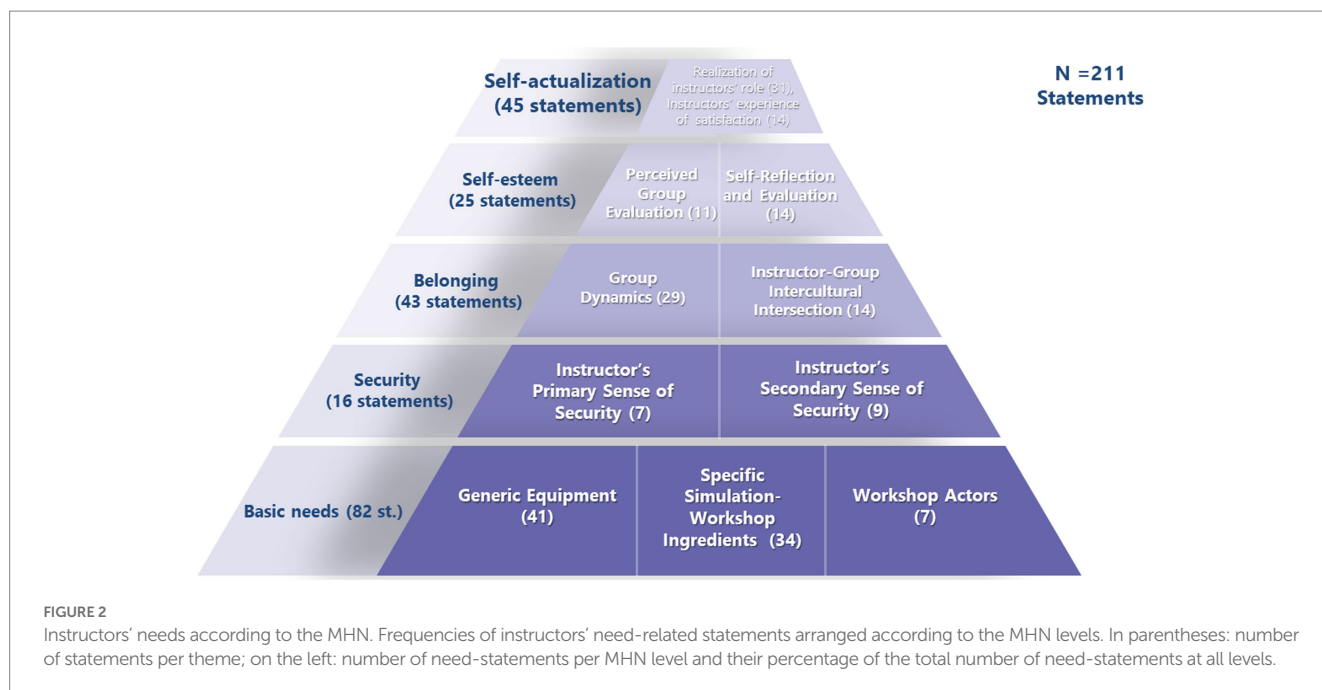
I once had a team of teachers who worked at a boarding school for high-risk students and had no interactions with the students' parents. Yet the scenario focused on teacher-parent relationships. The debriefing was dull. I had to try much harder to make something out of it ... in order to best function I need you to carefully select the scenario (FOCGR1).

In other words, when the scenario fails to correspond precisely to the profile of the workshop participants, the instructor's ability to focus the debriefing phase on the predefined goals is hindered, and thus the workshop fails to address the participants' needs.

Interestingly, the abovementioned needs were typically mentioned at the beginning of the instructors' descriptions, and if these needs were not met, their style became laconic, and they offered no further elaborations about other, higher-level needs. This observation further suggests that these needs were fundamental to instructors' functioning. Instructors also mentioned a human-based element as basic to their functioning: the professional actors.

3.2.3. The workshop actors

According to the instructors, the participation of professional actors provided them with the optimal conditions under which to conduct the debriefing phase. Two elements were of particular importance: the *quality of the actor's performance* and the *actor's feedback*. Specifically, the better the actor's performance, the more authentic the experience was for the participants, thus facilitating the



instructors' work: 'The actor performed his role so well! It significantly helped me deliver my main message, as the situation was experienced as authentic and thus initiated behaviors that align with the predefined key issues' (ECO88). Furthermore, the *actor's feedback* during the debriefing helped instructors to lead the group to the predefined goals.

The actor is able to deliver a complex message since it is based on his or her own experience during the simulation. This is all the more evident when the actor listens to the discussion held and internalizes the key points I am striving for, and then accommodates the feedback to these key points (FOGR2).

In fact, some of the instructors noted that they deliberately coordinated with the actors to obtain their cooperation in the debriefing stage. Thus, it appears that the actors' professional performance, feedback, and collaboration were perceived as essential basic ingredients for instructors' optimal functioning.

3.3. Security needs

Data analysis indicated that 7.5% (16 phrases) of the instructors' reports referred to issues of security. According to Maslow (1943), this stage refers to a sense of protection from emotional harm, as part of the need for self-preservation. Instructors' security-related needs involved both a primary and secondary sense of security.

3.4. Instructors' primary sense of security

Instructors related to their internal sense of emotional security, which stemmed predominantly from how well they were able to lead the workshop and seemed to depend on their prior professional training: 'Unfortunately, I did not come sufficiently prepared to lead the workshop ... I would have felt less threatened if, during the initial training, we would have gained more first-hand practice in leading a

workshop' (ECO12). The better instructors' preparation was, the more they seemed to experience a primary sense of internal security in leading the workshop. Yet this sense of security was also affected by externally-based influences, as elaborated below.

3.5. Instructors' secondary sense of security

Frequently, an authority figure (school principal, supervisor, etc.) attends the workshop along with the participants, a situation which can undermine the instructors' status and emotional security: 'The only challenge I had to cope with was the presence of the supervisor, who openly expressed her dissatisfaction with the way I was leading the workshop' (FOGR2). Receiving criticism from authority figures during workshops may directly impact instructors' sense of security, but their presence may also have an indirect effect on instructors' sense of security:

It would be very helpful if the group supervisor was not present ... This had a negative effect on me, as it seemed that some of the teachers felt completely blocked and the conversation was stilted ... I felt their tension ... They couldn't speak openly. It made my job much harder (REF-W61).

As shown, the authority figure may indirectly affect the instructor's sense of security: when the participants feel threatened, they are less willing to openly engage in the debriefing process, and the instructor is thus forced to lead the workshop under challenging conditions. Figure 3 illustrates the assumed paths through which the authority figure impacts the instructor's sense of security: directly (represented by a bold line) and indirectly (represented by a dotted line).

3.6. Belongingness

Forty-three phrases (20%) dealt with issues of belongingness. Within this level, two themes emerged: the first relating to 'regular'

ongoing situations and the second being relevant in cases of group diversity.

3.6.1. Instructor-group dynamics during the workshop

Instructors' collaboration with the group during the workshop was described as crucial to their ability to create an atmosphere that promotes learning. When this collaboration was suboptimal, instructors' functioning was hindered:

My interaction with the group was problematic. No one volunteered to participate in the simulation; they were quiet and not responsive. Some spent much of their time on their mobile phones. It was really difficult because I need to properly communicate with them to be able to function at my best (REF-W17).

Generally, instructors mentioned the need to work harmoniously with the group when guiding the workshop, for the sake of optimal operation of the workshop (recruiting a volunteer, achieving cooperation, facilitating the debriefing). This need for belongingness was potentially challenged when facing specific cases of cultural diversity.

3.6.2. Instructor-group intercultural gap

Some instructors perceived the group's cultural background as potentially challenging to their sense of belongingness. To face this challenge, they searched for information about the group beforehand, to become familiar with its practices.

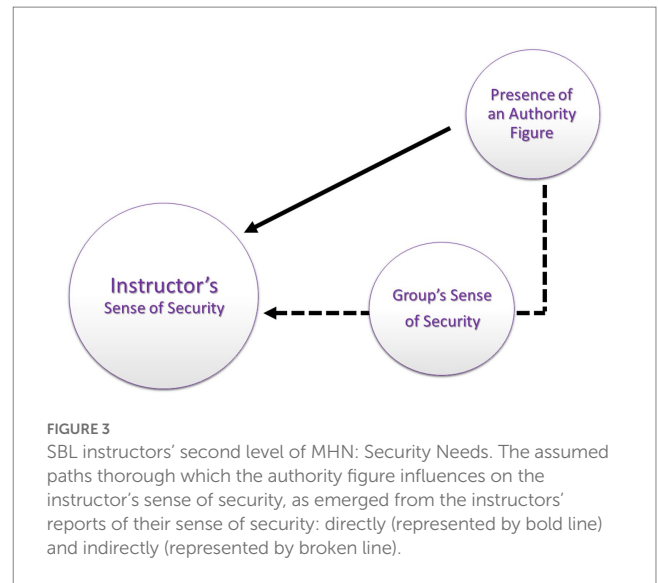
I had a workshop with a religious school, and I'm glad that a few days beforehand I entered their school's website and learned who they are and which projects they undertake. I did that because I am not religious, and their worldview is so different from my own. Later on, during the workshop, when they spoke about the project, I knew what they were talking about, so I wasn't an outsider (FOGR1).

As shown, familiarizing themselves with the group's cultural characteristics may help instructors to 'work in harmony with the group,' particularly when the group's background is unfamiliar to the instructor.

Not sharing the same native language presents another intercultural challenge. As described by a Jewish Hebrew-speaking instructor who managed an SBL workshop for a Bedouin group:

During the group discussion, they repeatedly commented to each other in Arabic, and I don't know Arabic. If I knew Arabic, these comments could have been further developed during the discussion ... and to be honest, at a certain point they were laughing and I wasn't sure whether they were laughing about me or about the topic discussed (REF-W25).

It is worth noting that the instructor positioned herself outside to the group (*they know ... I do not*), which precluded any sense of belonging. In such cases, the instructor's ability to communicate and 'join' the group is impeded. Consequently, instructors' ability to lead the collaborative discussions is impaired, and they may experience a sense of disconnection from the group.



3.7. Self-esteem needs

Findings indicated, that 12% (25 phrases) of the instructors' reports dealt with issues related to self-esteem needs, which is in line with Maslow's (1943) self-esteem needs.

3.7.1. Instructors' self-reflection and evaluation of functioning

Instructors' self-esteem needs manifested first and foremost in their self-reflection about their competence and functioning: 'I wonder how I can do it more effectively ... I'm not sure whether we went deep enough' (REF-W46). This engagement in a self-reflective process was present in many instructors' reports and appeared to focus on instructors' sense of competence in achieving their highest level of professional functioning. This concern aligns closely with one of the aspects (i.e., the need for competence and achievement) that comprise self-esteem needs according to Maslow (1943).

3.7.2. Instructors' perception of group evaluations

Another aspect of self-esteem, according to Maslow, is the need for recognition and appreciation. Indeed, positive participant responses seemed to be very important to the instructors: 'The participants' feedback was very warm; they said it had been a meaningful experience. They even asked to have an additional future workshop with me' (FOGR1). The phrase 'with me' implies that the instructor viewed participants' feedback as reflecting the appreciation of his or her professional performance, thus nurturing the instructor's self-esteem. In fact, most of the instructors explicitly asked to review their group's formal feedback forms, an indication of their perception that the group's final feedback was an evaluation of their performance. Figure 4 illustrates the two sources that nurture instructors' self-esteem.

3.8. Self-actualization

The analysis revealed that 21.5% (45 phrases) of the instructors' reports dealt with issues related to self-actualization, which in MHN involves reaching one's potential and desired goals. Two interrelated

themes emerged at this level: *realization of the instructor's role* and the *instructor's experience of satisfaction*.

3.8.1. Realization of instructor's role

Interestingly, many instructors began their responses by referring to what they perceived as the learning outcomes.

There was meaningful learning ... students said that they felt better prepared to face challenges in their practical work ... I can clearly see that they leave the workshop better equipped for their jobs and I find this very fulfilling ... at the end of the day, this is why I'm here (REF-W38).

Self-actualization occurs, according to Maslow (1943), when individuals realize their full potential, or when they attain their desired goals (Shaughnessy et al., 2018). Hence, given that the instructors' role is to mediate the learning process, their need for self-actualization is gratified when such a process has successfully occurred. By the same token, when such a process is not achieved, instructors' sense of self-realization is absent. Unsurprisingly, the achievement of their goals was significantly related to instructors' overall experience of satisfaction.

3.8.2. Instructors' experience of satisfaction

When they were able to successfully realize their role as mediators of the SBL learning process, the emotions experienced by the instructors included enjoyment, reward, and a sense of satisfaction: 'I felt that I had contributed to the participants' professional development and that was so rewarding. I enjoyed it' (ECO-47).

In contrast, when self-realization was not achieved, the emotions experienced included frustration, despair, disappointment, and sometimes bitterness. 'Frankly, I am quite disappointed We did not gain much learning. So frustrating!!' (ECO-56).

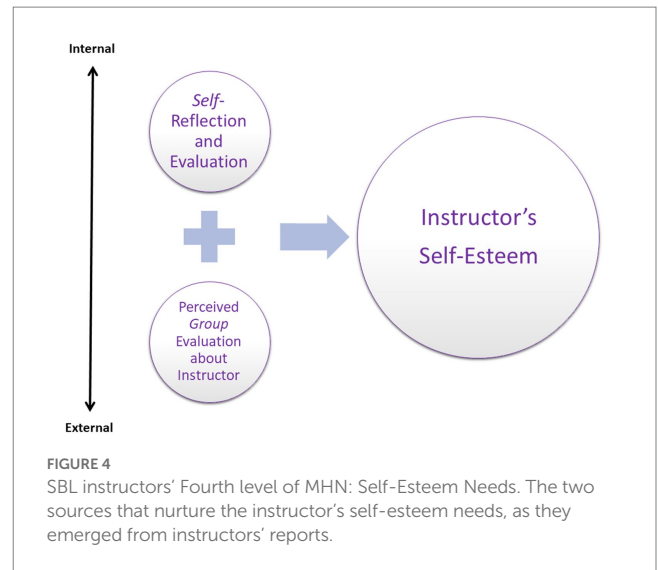
In conclusion, self-actualization consisted of two aspects: The first involved a *cognitive* process, during which instructors examined whether their ultimate goal of mediating the learning was achieved, whereas the second focused on their *emotional* experience as a result of this achievement or lack thereof.

3.9. The SBL process in light of instructors' needs, as related to MHN – Two possible paths

The following section describes the process of a successfully completed SBL process vs. an incomplete SBL process, depending on whether the hierarchical needs were fully or partially met, from the instructors' perspectives.

3.9.1. The complete process, in which all needs are met

The analysis revealed that 62 of the 68 workshops reviewed in this study corresponded to the complete process. In this path, all five of the hierarchical levels of needs were addressed in the simulation workshop: a condition that enabled the instructor to achieve self-actualization. As we analyze Debra's (all participant names are pseudonyms) description of her experience as an SBL workshop instructor, we find that her sense of self-actualization (level 5 need) includes the fulfilment of her needs from all of the preceding levels.



The simulations contributed a great deal [level 5–self-actualization]. The students explicitly expressed their enthusiasm [level 4–self-esteem]. I have been working with this group since the beginning of the academic year and there was a pleasant atmosphere [level 3–belonging].

As can be seen, no mention was made of first or second-level needs (neither basic nor security needs). Disregarding these first two levels was typical of cases in which a complete process had taken place, most likely because when fundamental needs are met, there is no need to mention them. However, when these two basic-level needs were not met, their absence was very conspicuous, as shown in the following section.

3.9.2. The incomplete process, in which needs remain unmet

Six of the 68 workshops conducted corresponded to the path of ungratified needs. In these cases, the instructor reported that the goals of the SBL were not achieved and, to explain why, referred to one or more needs that remained unfulfilled throughout the workshop. Workshops that do not achieve their predefined goals constitute an important test case, through which lessons can be learned and improvements can be made. It is thus worthwhile to present the following two examples of workshops that were conducted by different instructors and involved different target populations. The analysis and MHN-related conceptualizations of the excerpted quotes regarding ungratified needs are presented in brackets.

Lea's report, after conducting an SBL workshop for preservice teachers

The pages and forms were disorganized; the tablets didn't work (at the last moment I had to photocopy pages) [level 1 – basic needs; equipment]. In general, I'm not sure whether we were able to delve deep enough or whether I allowed the group enough room for participants to express themselves [level 4 –self-esteem; self-reflection]. ... When the students arrived, they were very tired and resistant, so they did not really cooperate. I didn't feel comfortable in the group, I actually had to disarm their resistance [level 3 – belonging; dynamics during the workshop]. The group's

TABLE 2 A demonstration of the coding process using qualitative content analysis: An example of instructors’ needs that correspond to MHN-level 1, Basic Needs.

MHN level	Themes	Description of need	Sub-themes	Unit of Analysis (Examples of coded data)
Basic needs (level 1)	Equipment	Initial needs that enable the instructor to operate the workshop appropriately	Satisfactory technical equipment	- “The binders with the forms were not well-organized. I had to search for the correct form...” (REF-W31) - “The name badges should have been prepared beforehand. I had to waste a lot of time searching for these stickers instead of focusing on what really matters” (REF-W4)
	Workshop Ingredients	Initial needs that relate to the simulation’s unique auxiliary components	Video system facilities	- “Because of technical problems in the video system, I was unable to use it throughout the debriefing stage. It complicated everything for me” (ECO2) - “Observing the recorded video segment is one of the most important things that help me to guide the group towards the points I wish to teach during the workshop” (REF-W18)
			Adjusted scenario	- “The scenarios were so precise and adjusted to the group – in a way that it seemed to imitate what they were dealing with in real life; the debriefing went quite smoothly. I think this was one of my best simulation workshops!” (REF-W44)
	The Workshop Actors	The need for effective acting as well as collaboration between the actor and the instructor	Quality of actor’s performance	- “The actor did not do a professional acting job and it made the debriefing very challenging. In contrast, in last week’s workshop, the actor did a good job acting, and it really made a difference when I guided the debriefing” (REF-W9)
			Actor’s feedback	- “Generally, the actor to is able to deliver a complex message since it is based on his or her own experience during the simulation. This is all the more evident when the actor listens to the discussion held and internalizes the key points I am striving for, and then accommodates the feedback to these key points” (FOGR2) - “The actor and I took a few minutes to consult and coordinate: ‘I will focus on this and you’ll focus on that.’ This way, the students do not get bored and I know what I want to focus on” (REF-W33)

pedagogical mentor was present at the workshop and she felt it was important that she have her say. I asked her to enable the group members to express themselves [level 2 – security; authority figure]. But all in all, I’m not sure to what extent we were able to delve into the essential matters. Quite disappointing, I must admit [level 5 – self-actualization].

Naif’s report after conducting a workshop for a group of inservice teachers

From the beginning I felt that the relationships among the school staff were loaded and problematic, I found myself wondering how I should fit in [level 3 – belonging]. In the first simulated round, the participant could not be heard, due to a bug in the video system [level 1 – basic needs; workshop ingredients]. Throughout the workshop, the passive resistance was palpable and had a significant effect on the entire process [level 3 – belonging; group dynamics during the workshop]. In the second round, there were no volunteers. The vice-principal of the school was present and intervened, which in my view, prohibited an open discussion during the workshop [level 2 – security; authority figure]. My numerous attempts to break down participants’ resistance were only partially successful [level 4 –self-esteem; self-reflection]. What a waste of valuable resources!! [level 5 – self-actualization]

The SBL workshops that were characterized as ‘incomplete’ were described in greater detail (manifested in the length of the transcript text) than those characterized as ‘complete’ (shorter segments in the transcripts). Additionally, these longer descriptions tended to describe the process in a way that aligns with Maslow’s hierarchy. By providing these lengthy descriptions, the instructors inadvertently attributed the lack of completion of the SBL process to their needs that remained unmet.

4. Discussion

Using Maslow’s Hierarchy of Needs as a theoretical framework, the current study provides theoretical insight into the needs of SBL instructors. As these instructors are the ones mediating the learning process (Sellberg, 2018) by leading the debriefing – the most essential ingredient in SBL – it is imperative to examine their needs. Knowledge about SBL (particularly about instructors’ needs when leading SBL) in the field of teacher education is limited. As such, this study contributes to the existing literature by enhancing the theoretical understanding of instructors’ needs and highlighting the importance of meeting them. In turn, this information can be used to improve the quality of SBL-based teacher education. That is, the goal of the SBL in teacher education is to provide preservice and inservice teachers with the opportunity to analyze challenges that resemble those they encounter in the professional field; the likelihood of realizing this goal increases as more of the SBL

instructors' needs are met. Consequently, gaining an understanding and thus being able to fully address the SBL instructors' needs is ultimately more beneficial for the teachers participating in the SBL workshops. To this end, practical recommendations for optimizing SBL instructors' training and promoting SBL best practices can be formulated.

Overall, we were able to link instructors' needs with the principles of Maslow's theory (Maslow, 1943), which posits that self-actualization depends on meeting the needs of the hierarchy's preceding levels. Accordingly, when instructors' needs were gratified, they felt that they successfully fulfilled their role as instructor (i.e., mediating the learning), and they experienced satisfaction. Conversely, when instructors' needs were not met, they felt frustrated and disappointed. In such cases, instructors' reports included an elaborate explanation of why this 'failure' occurred, revealing that when lower-level needs were not met, then higher-level needs could not be met either, bearing out the principles underlying MHN (i.e., that self-actualization is achieved only when the needs of preceding levels are met). The theoretical insights and implications drawn from this study are presented and discussed according to the five levels of Maslow's hierarchy.

4.1. Basic needs

The fact that certain needs (equipment, scenario, and actor) were mentioned early on in the instructors' descriptions suggests that these were fundamental to the SBL instructors. Moreover, whenever these needs were not met, the instructors were concise in their comments and did not elaborate further about higher-level needs. These two features together indicate the fundamental nature of these needs, in line with MHN. Findings suggest that to meet instructors' basic needs, the following three issues should be addressed: First, in accordance with scholars' recent claims (Fisher and Royster, 2016), who argue that equipment and supplies are considered basic needs, *providing adequate equipment* is basic to instructors' ability to manage the simulation. This finding aligns with Ferguson's (2017) prior recommendations.

Second, designing *suitable scenarios* is imperative to instructors, as it creates a sense of authenticity that facilitates instructors' role. Although authenticity was previously documented as crucial to *workshop participants'* experience in SBL (Campbell and Daley, 2017), the current study highlights its importance for *instructors'* ability to lead SBL. In the same vein, as technology-related needs in this context coincide with Maslow's basic needs (Bailey and Pownell, 1998), instructor training should include technology orientation, as previously argued (Fanning and Gaba, 2007).

Third, as the *workshop's actors* bring the scenarios to life, employing professional actors is an essential basic requirement. The necessity of professional actors was previously advocated for the purposes of promoting authenticity (Pascucci et al., 2014), yet the current study highlights its necessity for instructors' best functioning.

4.2. Security needs

Maslow's theory was previously applied to explain the need for a safe learning environment in general (Shaughnessy et al., 2018), and particularly in the context of SBL *participants* (Tutticci et al., 2018). However, the current study underscores that security needs should be met not only for learners but also for *instructors*. Indeed, it was argued that instructors must be prepared to face unexpected challenging situations (Der-Sahakian et al., 2015). The current study exemplifies one

of these challenges: the presence of an authority figure who may undermine the instructor. When the instructor feels threatened, the ability to function optimally is impaired, and it is likely that the workshop participants' learning experience will likewise be suboptimal. This echoes previous claims that learners' perceptions of teachers as authority figures may affect the teaching-learning process (Gil-Madrona et al., 2020). Hence, instructors should be provided with preliminary training focusing on how to handle such perceived threats in the context of leading a workshop to ensure that teachers receive the best SBL conditions.

4.3. Belongingness needs

Corresponding to Maslow's sense of belonging, which includes the desire for collaboration and harmonious existence with others, establishing a sense of belongingness with the group was experienced by instructors as a vital need, allowing them to perform optimally. This finding echoes a previous study that found that an important factor that influenced teacher mentors' self-actualization was the relationship that developed during the mentoring process (Fletcher, 2006). Our study further expands these findings by emphasizing the importance of relationships in the context of SBL instruction as well. In fact, the sense of belonging is particularly relevant as SBL involves group peer learning (Levin and Flavian, 2020). Hence, it may be challenged in cases of instructor-group diversity. Instructor training should therefore aim to instill cultural sensitivity and social justice norms (Grogan, 2014; Martinek et al., 2021; Losada Puente et al., 2022), to bridge potential gaps. Noteworthy, intercultural competence was previously explored as an outcome variable of SBL (Harder, 2018), yet this study highlights the need to view cultural differences as a factor that affects instructors' ability to function effectively. Training instructors to speak the group's cultural 'lingo' is vital to instructors' belongingness needs and overall performance.

4.4. Self-esteem needs

In accordance with MHN's fourth level, which involves the need to be competent and to achieve, as well as to be recognized and appreciated (Maslow, 1943 in Adams et al., 2015), instructors reported engaging in a dual process of reflection, based on two sources: instructors' *internal* self-reflection (about their competence and achievements) and the perceived feedback from the group (i.e., *external* recognition and appreciation).

Whereas prior SBL research has demonstrated that reflection promotes effective learning among *participants* (Butvilofsky et al., 2012), the current study underscores the importance of reflection for addressing *instructors'* needs during SBL. Thus, allocating time to instructors for receiving participants' reflections, granting them access to feedback questionnaires, encouraging them to engage in self-reflection, and using peers' reflective discourse (Der-Sahakian et al., 2015) may all be channels through which instructors' need for self-evaluation (as part of their self-esteem needs) is addressed.

4.5. Self-actualization

The peak of MHN involves reaching self-actualization, which essentially involves the realization of one's desired goals (Shaughnessy et al., 2018) and occurs when individuals realize their full potential

(Maslow, 1943). Instructors' reports indicated two components of self-actualization in the context of SBL: (a) Instructors' actualization of their job-related potential as mediators of learning, manifested in the achievement of learning outcomes, and (b) The experience of satisfaction once instructors' goals are achieved.

Findings further showed two possible paths that the SBL process can take. As expected, and in accordance with Maslow's principles: (1) the self-actualization of instructors requires the meeting of all prior needs in the hierarchy, and (2) conversely, instructors' unmet needs hinder their self-actualization and result in experiences of frustration and disappointment. Hence, it is essential to design workshops in a manner that ensures that all instructors' needs are met, to optimize the likelihood that they will realize the workshop's goals.

5. Implications and limitations

The study's major implication for teacher education lies in the focus on the instructors' needs, given that they are the ones who facilitate teachers' learning through simulation. To date, the literature on educational simulations has focused solely on the workshop participants' needs (Tutticci et al., 2018). The current study, however, sheds light on the instructors' needs and the components that may promote effective SBL from the instructors' perspectives. In fact, instructors' needs may be viewed as an indirect factor that ultimately affects the quality of teachers' SBL experience. As such, the study contributes to the ongoing conversation regarding the ways in which learning through simulation can be understood and improved to promote teachers' learning. The study has a few limitations. First, it is possible that instructors' reflections were influenced by response biases. To minimize this possibility, a triangulation of research sources was established. Second, as the sample comprised mainly females, transferability may be limited. However, given the gender imbalance in the field of teacher education (Robinson et al., 2017), we believe that the sample is adequate. Third, the research was conducted in one specific teacher-education simulation center, which is located at a teacher-education college. Thus, transferability to teachers attending other colleges may be limited. As such, a multisite investigation of this issue is needed to further examine the research findings. Furthermore, given that this is a case study, transferring and comparing these findings to findings of other studies should be conducted with extra caution (Stake, 2013). For example, this study involved clinical simulations and, therefore, applying its findings about the SBL instructors' needs to virtual simulations requires careful consideration. Consequently, additional studies are needed, to expand the paradigm to a multiple-case study. In addition, we call for a longitudinal examination of instructors' needs, to determine whether these needs change over time. Finally, to better understand the impact of an authority figure's presence during SBL, further research is recommended.

6. Conclusion

As instructors play a significant role in SBL, the current study contributes to the existing literature of teacher education by providing an in-depth theoretical conceptualization of SBL instructors' needs and highlighting the importance of addressing them, to provide teachers with the best possible learning conditions. Addressing the study's goal

by demonstrating the SBL instructors' needs vis-à-vis MHN enhances the theoretical understanding of this phenomenon and provides applicable insights that can be assimilated in SBL instructors' training. Specifically, efforts should be made to address instructors' varied needs, as follows: *basic needs* may be attended to by providing adequate equipment, designing suitable and authentic scenarios, and employing professional actors; *security needs* may be met by preparing instructors to face unexpected challenging situations and by providing preliminary training focused on handling perceived sources of threats when leading a workshop; to create a *sense of belonging*, instructors should receive cultural-competence training, to learn of ways to bridge potential gaps; *self-esteem* needs may be addressed by allocating time for receiving participants' reflections and encouraging instructors to engage in self-reflection as part of their need for self-evaluation. Ultimately, all of these are ways to help instructors experience *self-actualization* and satisfaction. Given that SBL is becoming an integral part of teacher education and considering the essential role of the instructors in facilitating teachers' learning through simulation, understanding how to address instructors' needs and how to design optimal training for SBL instructors in teacher education is imperative. Future studies may opt to focus on instructors' needs in the context of virtual simulations to further broaden our understanding of whether and/or how instructors' needs are manifested differently in this context. To this end, studies should be conducted in a range of contexts, for example, in a multicultural framework, and with different participant populations. Furthermore, future studies should aim to explore whether these needs are shared across different learning environments or whether these are unique to SBL. To conclude, the current study underscores the importance of attending to instructors' needs, to ensure that they can use the simulation tool in teacher education to its full potential.

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 Achva academic college ethics committee. 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.

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

References

- Adams, J., Harris, C., and Bohley Martin, K. A. (2015). Explaining small-business development: a small-business development model combining the Maslow and the Hayes and wheelwright models. *J. India. Acad. Soc. Sci.* 18, 26–36.
- Ade-Ojo, G. O., Markowski, M., Essex, R., Stiell, M., and Jameson, J. (2022). A systematic scoping review and textual narrative synthesis of physical and mixed-reality simulation in pre-service teacher training. *J. Comput. Assist. Learn.* 38, 861–874. doi: 10.1111/jcal.12653
- Bailey, G. D., and Pownell, D. (1998). Technology staff-development and support programs: applying Abraham Maslow's hierarchy of needs. *Learn. Lead. Technol.* 26, 47–51.
- Bautista, N. U., and Boone, W. J. (2015). Exploring the impact of TeachMETM lab virtual classroom teaching simulation on early childhood education Majors' self-efficacy beliefs. *J. Sci. Teach. Educ.* 26, 237–262. doi: 10.1007/s10972-014-9418-8
- Bogdan, R. C., and Biklen, S. K. (2007). *Qualitative research in education: An introduction to theory and methods*. Boston, MA: Allyn & Bacon.
- Butvilofsky, S. A., Escamilla, K., Soltero-González, L., and Aragon, L. (2012). Promoting reflective teaching through simulation in a study in Mexico program. *J. Hisp. High. Educ.* 11, 197–212. doi: 10.1177/1538192711435561
- Campbell, S. H., and Daley, K. (2017). *Simulation scenarios for nursing educators: making it real*. Springer Publishing Company.
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., and Fischer, F. (2020). Simulation-based learning in higher education: a meta-analysis. *Rev. Educ. Res.* 90, 499–541. doi: 10.3102/0034654320933544
- Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., and Seidel, T. (2020). Between authenticity and cognitive demand: finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teach. Teach. Educ.* 95:103146. doi: 10.1016/j.tate.2020.103146
- Coon, D. J. (2006). Abraham H. Maslow: reconnaissance for Eupsychia. *Port. Pioneers Psychol.* 6, 255–271.
- Creswell, J. W. (2007). *Qualitative inquiry and research method: choosing among five approaches*. Sage.
- Dalgarno, B., Gregory, S., Knox, V., and Reiners, T. (2016). Practicing teaching using virtual classroom role plays. *Aust. J. Teach. Educ.* 41, 126–154. doi: 10.14221/ajte.2016v41n1.8
- Dalinger, T., Thomas, K. B., Stansberry, S., and Xiu, Y. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. *Comput. Educ.* 144:103696. doi: 10.1016/j.compedu.2019.103696
- Danish, J. A., Enyedy, N., Saleh, A., and Humburg, M. (2020). Learning in embodied activity framework: a sociocultural framework for embodied cognition. *Int. J. Comput.-Support. Collab. Learn.* 15, 49–87. doi: 10.1007/s11412-020-09317-3
- Der-Sahakian, G., Alinier, G., Savoldelli, G., Oriot, D., Jaffrelot, M., and Lecomte, F. (2015). Setting conditions for productive debriefing. *Simul. Gaming* 46, 197–208. doi: 10.1177/1046878115576105
- Dieker, L. A., Rodriguez, J. A., Lignugaris Kraft, B., Hynes, M. C., and Hughes, C. E. (2014). The potential of simulated environments in teacher education: current and future possibilities. *Teach. Educ. Spec. Educ.* 37, 21–33. doi: 10.1177/0888406413512683
- Dotger, B. H. (2010). "I had no idea": developing dispositional awareness and sensitivity through a cross-professional pedagogy. *Teach. Teach. Educ.* 26, 805–812. doi: 10.1016/j.tate.2009.10.017
- Dotger, B. H. (2015). Core pedagogy: individual uncertainty, shared practice, formative ethos. *J. Teach. Educ.* 66, 215–226. doi: 10.1177/0022487115570093
- Enyedy, N., Danish, J. A., and DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *Int. J. Comput.-Support. Collab. Learn.* 10, 7–34. doi: 10.1007/s11412-015-9207-1
- Fanning, R. M., and Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simul. Healthc.* 2, 115–125. doi: 10.1097/SIH.0b013e3180315539
- Ferguson, K. (2017). Using a simulation to teach Reading assessment to preservice teachers. *Read. Teach.* 70, 561–569. doi: 10.1002/trtr.1561
- Fischetti, J., Ledger, S., Lynch, D., and Donnelly, D. (2021). Practice before practicum: simulation in initial teacher education. *Teach. Educ.* 57, 155–174.
- Fisher, M. H., and Royster, D. (2016). Mathematics teachers' support and retention: using Maslow's hierarchy to understand teachers' needs. *Int. J. Math. Educ. Sci. Technol.* 47, 993–1008. doi: 10.1080/0020739X.2016.1162333
- Fives, H., and Mills, T. M. (2016). Making motivation meaningful by mastering Maslow. In M. C. Smith and Frates DenschN. De (Eds.), *Challenges and innovations in Educational Psychology teaching and learning* (137–150). Information Age Publishing.
- Fletcher, S. (2006). Attaining self-actualisation through mentoring. *Eur. J. Teach. Educ.* 21, 109–118. doi: 10.1080/0261976980210110
- Frei-Landau, R., and Avidov-Ungar, O. (2022). Educational equity amidst COVID-19: exploring the online learning challenges of Bedouin and Jewish female preservice teachers in Israel. *Teach. Teach. Educ.* 111:103623. doi: 10.1016/j.tate.2021.103623
- Frei-Landau, R., and Levin, O. (2022). The virtual Sim (HU) lation model: conceptualization and implementation in the context of distant learning in teacher education. *Teach. Teach. Educ.* 117:103798. doi: 10.1016/j.tate.2022.103798
- Frei-Landau, R., Orland-Barak, L., and Muchnick-Rozonov, Y. (2022). What's in it for the observer? Mimetic aspects of learning through observation in simulation-based learning in teacher education. *Teach. Teach. Educ.* 113:103682. doi: 10.1016/j.tate.2022.103682
- Freitas, F. A., and Leonard, L. J. (2011). Maslow's hierarchy of needs and student academic success. *Teach. Learn. Nurs.* 6, 9–13. doi: 10.1016/j.teln.2010.07.004
- Garden, A. L., Fevre, D. M., Waddington, H., and Weller, J. M. (2015). Debriefing after simulation-based non-technical skill training in healthcare: a systematic review of effective practice. *Anaesth. Intensive Care* 43, 300–308. doi: 10.1177/0310057X1504300303
- Ghouseini, H. (2017). Rehearsals of teaching and opportunities to learn mathematical knowledge for teaching. *Cogn. Instr.* 35, 188–211. doi: 10.1080/07370008.2017.1323903
- Gil-Madrona, P., Hinojosa, L. M. M., Pérez-Segura, J. J., Sáez-Sánchez, M. B., and Poblete, G. Z. (2020). Scale of pedagogical authority meanings in the classroom (ESAPA) for Ibero-America built on the opinions of teaching students. *Teach. Teach. Educ.* 93:103079. doi: 10.1016/j.tate.2020.103079
- Girod, M., and Girod, G. (2008). Simulation and the need for quality practice in teacher preparation. *J. Technol. Teach. Educ.* 16, 307–337.
- Grogan, M. (2014). Educational leadership and social justice in the United States. *Bildung und Erziehung* 67, 299–312. doi: 10.7788/bue-2014-0305
- Harder, N. (2018). Determining the effects of simulation on intercultural competency in undergraduate nursing students. *Nurse Educ.* 43, 4–6. doi: 10.1097/NNE.0000000000000403
- Judge, S., Bobzien, J., Maydosz, A., Gear, S., and Katsioloudis, P. (2013). The use of visual-based simulated environments in teacher preparation. *J. Educ. Train. Stud.* 1, 88–97. doi: 10.11114/jets.v1i1.41
- Kaufman, D., and Ireland, A. (2016). Enhancing teacher education with simulations. *Tech Trends* 60, 260–267. doi: 10.1007/s11528-016-0049-0
- Kornbluh, M. (2015). Combatting challenges to establishing trustworthiness in qualitative research. *Qual. Res. Psychol.* 12, 397–414. doi: 10.1080/14780887.2015.1021941
- Larson, K. E., Hirsch, S. E., McGraw, J. P., and Bradshaw, C. P. (2020). Preparing preservice teachers to manage behavior problems in the classroom: the feasibility and acceptability of using a mixed-reality simulator. *J. Spec. Educ. Technol.* 35, 63–75. doi: 10.1177/0162643419836415
- Ledger, S. (2021). *Resilience building for pre-service teachers: BRiTE, micro-teaching and augmented reality/simulation (BRiTE-AR)*, vol. 245 Cultivating Teacher Resilience.
- Ledger, S., Burgess, M., Rappa, N., Power, B., Wong, K. W., Teo, T., et al. (2022). Simulation platforms in initial teacher education: past practice informing future potentiality. *Comput. Educ.* 178:104385. doi: 10.1016/j.compedu.2021.104385
- Ledger, S., and Fischetti, J. (2020). Micro-teaching 2.0: technology as the classroom. *Australas. J. Educ. Technol.* 36, 37–54.
- Levin, O., and Flavian, H. (2020). Simulation-based learning in the context of peer learning from the perspective of preservice teachers: a case study. *Eur. J. Teach. Educ.* 45, 373–394. doi: 10.1080/02619768.2020.1827391
- Levin, O., Frei-Landau, R., Flavian, H., and Miller, E. C. (2022). Creating authenticity in simulation-based learning scenarios in teacher education. *Eur. J. Teach. Educ.*, 1–22.
- Losada Puente, L., Fiuza Asorey, M., and Bana Castro, M. (2022). What defines inclusion in higher education institutions? Validation of an instrument based on the 'index for inclusion'. *Int. J. Disabil. Dev. Educ.* 69, 91–105. doi: 10.1080/1034912X.2021.1992752
- Makransky, G., Terkildsen, T. S., and Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learn. Instr.* 60, 225–236. doi: 10.1016/j.learninstruc.2017.12.007

- Malone, S., and Brünken, R. (2021). Hazard perception, presence, and simulation sickness—a comparison of desktop and head-mounted display for driving simulation. *Front. Psychol.* 12:647723. doi: 10.3389/fpsyg.2021.647723
- Martinek, T. J., Santos, F., Sá, C., and Saraiva, L. (2021). Teaching personal and social responsibility (TPSR) in teacher education: reflections and future direction. *J. Educ. Cult. Stud.* 5, p32–p13. doi: 10.22158/jecs.v5n2p32
- Maslow, A. H. (1943). A theory of human motivation. *Psychol. Rev.* 50, 370–396. doi: 10.1037/h0054346
- Muchnik-Rozanov, Y., Frei-Landau, R., and Avidov-Ungar, O. (2022). Mobile-learning adoption in teacher education amidst COVID-19: identifying two critical stages by exploring teachers' emotions. *Front. Educ.* 7:1077989. doi: 10.3389/educ.2022.1077989
- Orland-Barak, L. (2014). Mediation in mentoring: a synthesis of studies in teaching and teacher education. *Teach. Teach. Educ.* 44, 180–188.
- Pascucci, R. C., Weinstock, P. H., O'Connor, B. E., Fancy, K. M., and Meyer, E. C. (2014). Integrating actors into a simulation program: a primer. *Simul. Healthc.* 9, 120–126. doi: 10.1097/SIH.0b013e3182a3ded7
- Rayner, C., and Fluck, A. (2014). Pre-service Teachers' perceptions of Sim School as preparation for inclusive education: a pilot study. *Asia Pac. J. Teach. Educ.* 42, 212–227. doi: 10.1080/1359866X.2014.927825
- Riley, R. D., and Mort, K. (1981). Teacher center responses to teacher needs. *Clearing House* 54, 227–230. doi: 10.1080/00098655.1981.9957164
- Robinson, K., Shakeshaft, C., Grogan, M., and Newcomb, W. S. (2017). Necessary but not sufficient: the continuing inequality between men and women in educational leadership, findings from the American Association of School Administrators mid-decade survey. *Front. Educ.* 2:12. doi: 10.3389/educ.2017.00012
- Ross, P., and Atkinson, A. P. (2020). Expanding simulation models of emotional understanding: the case for different modalities, body-state simulation prominence, and developmental trajectories. *Front. Psychol.* 11:309. doi: 10.3389/fpsyg.2020.00309
- Schulte, M. (2018). Adult learning degree and career pathways: allusions to Maslow's hierarchy of needs. *J. Contin. High. Educ.* 66, 62–64. doi: 10.1080/07377363.2017.1368767
- Sellberg, C. (2018). From briefing, through scenario, to debriefing: the maritime Instructor's work during simulator-based training. *Cogn. Tech. Work* 20, 49–62. doi: 10.1007/s10111-017-0446-y
- Shapira-Lishchinsky, O. (2013). Team-based simulations: learning ethical conduct in teacher trainee programs. *Teach. Teach. Educ.* 33, 1–12. doi: 10.1016/j.tate.2013.02.001
- Shaughnessy, M. F., Moffitt, B., and Cordova, M. (2018). Maslow, basic needs and contemporary teacher training issues. *Arch. Curr. Res. Int.* 14, 1–7. doi: 10.9734/ACRI/2018/42858
- Snoek, M., Swennen, A., and van der Klink, M. (2010). "The teacher educator: a neglected factor in the contemporary debate on teacher education" in *Advancing quality cultures for teachers' education in Europe: Tensions and opportunities*. eds. B. Hudson, P. Zgaga and B. Åstrand (Sweden: Umeå University), 33–48.
- Stake, R. E. (2013). *Multiple case study analysis*. New York, NY: Guilford Press.
- Tang, F. M. K., Lee, R. M. F., Szeto, R. H. L., Cheng, J. K. K., Choi, F. W. T., Cheung, J. C. T., et al. (2021). A simulation Design of Immersive Virtual Reality for animal handling training to biomedical sciences undergraduates. *Front. Educ.* 6:239.
- Theelen, H., van den Beemt, A., and Brok, P. D. (2019). Classroom simulations in teacher education to support preservice teachers' interpersonal competence: a systematic literature review. *Comput. Educ.* 129, 14–26. doi: 10.1016/j.compedu.2018.10.015
- Thisgaard, M., and Makransky, G. (2017). Virtual learning simulations in high school: effects on cognitive and non-cognitive outcomes and implications on the development of STEM academic and career choice. *Front. Psychol.* 8:805. doi: 10.3389/fpsyg.2017.00805
- Tutticci, N., Ryan, M., Coyer, F., and Lewis, P. A. (2018). Collaborative facilitation of debrief after high-fidelity simulation and its implications for reflective thinking: student experiences. *Stud. High. Educ.* 43, 1654–1667. doi: 10.1080/03075079.2017.1281238
- Van Velzen, C., Van der Klink, M., Swennen, A., and Yaffe, E. (2010). The induction and needs of beginning teacher educators. *Prof. Dev. Educ.* 36, 61–75. doi: 10.1080/19415250903454817
- Vlachopoulos, D., and Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. *Int. J. Educ. Technol. High. Educ.* 14, 1–33. doi: 10.1186/s41239-017-0062-1
- Vygotsky, L. S. (1978). Socio-cultural theory. *Mind Society* 6, 52–58.
- Yablon, Y. B., Iluz, S., and Hershkowitz, S. (2021). Simulation-based learning for rabbi-teachers dealing with conflicts in education. *Br. J. Relig. Educ.* 44, 315–324.
- Zhong, W., and Zhang, G. (2021). Mental simulation to promote exercise intentions and behaviors. *Front. Psychol.* 12:589622. doi: 10.3389/fpsyg.2021.589622



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo,
Portugal

REVIEWED BY

Elizabeth Fraser Selkirk Hannah,
University of Dundee, United Kingdom
Milan Kubiakto,
J. E. Purkyne University, Czechia

*CORRESPONDENCE

Jirina Karasova
✉ j.j.karasova@gmail.com

SPECIALTY SECTION

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

RECEIVED 01 February 2023

ACCEPTED 17 March 2023

PUBLISHED 12 April 2023

CITATION

Karasova J and Nehyba J (2023)
Student-centered teacher responses
to student behavior in the classroom:
A systematic review.
Front. Educ. 8:1156530.
doi: 10.3389/feduc.2023.1156530

COPYRIGHT

© 2023 Karasova and Nehyba. 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.

Student-centered teacher responses to student behavior in the classroom: A systematic review

Jirina Karasova* and Jan Nehyba

Department of Education, Masaryk University, Brno, Czechia

Introduction: Effective communication skills are essential for successful behavior management in the classroom. Teachers who can respond proactively and in a student-centered manner can create a positive and productive learning environment. However, the empirical support for student-centered communication practices in behavior management is limited.

Methods: To address this gap, a systematic literature review was conducted to identify the characteristics of student-centered behavior management strategies that lead to lower behavior problems and increased student engagement. The review utilized a PRISMA protocol to ensure the rigor of the study selection process.

Results: Five main categories were identified that characterize student-centered behavior management responses. A table of 24 communication strategies was presented based on the findings of the review. The study also discussed the further impact of these strategies on student motivation, learning outcomes, responsibility, and interpersonal classroom climate.

Discussion: The findings of this study highlight the importance of effective communication skills in behavior management and provide valuable insights for teachers to improve their practice. By implementing these student-centered communication strategies, teachers can manage the classroom effectively, creating a more positive and productive learning environment and supporting students in achieving better learning outcomes.

KEYWORDS

communication skills, classroom behavior management, student-centered, systematic review, communication strategies

1. Introduction

Teachers worldwide spend a significant amount of their teaching time managing student behavior and report it as their main challenge in the profession (Eisenman et al., 2015; Kwok, 2020). In the United States, more than one-third of teachers stated that student behavior interfered with their teaching [as reported in Steinberg and Lacoe (2017)]. Similarly, in the recent Talis (OECD, 2019) study, it was reported that 29 percent of teachers spend a significant amount of their instruction time on behavior management. In a survey in Australia (Auditor General Western Australia, 2014), 39 percent of teachers used more than 20 percent of their teaching day on student behavior.

There is not one single widely accepted definition of classroom behavior management. Brophy (2006) describes classroom management as any teachers' actions, which create and facilitate a learning environment for successful instruction. Emmer and Evertson (2012) conclude that for effective behavior management, which is an essential part of classroom management, teachers need to develop a caring relationship with and among students, encourage students' engagement, optimize student access to learning, promote students' social, emotional, and self-regulation skills, and use appropriate interventions to help students with behavior problems. The current study focuses on a key component of classroom behavior management, specifically on teachers' communication responses to student behavior.

Effective teaching and learning cannot happen in poorly managed classrooms (Hattie, 2009, 2012; Korpershoek et al., 2016). If not managed effectively, classroom disturbances increase teachers' stress, drop-out, and job dissatisfaction (Clunies-Ross et al., 2008; Aldrup et al., 2018; Paramita et al., 2020). Further, ineffective classroom management strategies harm student wellbeing, responsibility, self-concept, and result in a greater amount of disruptive behavior in the classroom (Larrivee, 2005; Omoteso and Semudara, 2011). On the other hand, as Emmer and Evertson (2012) suggest, when teachers possess effective behavior management skills, it reduces their stress and burnout (Oliver et al., 2011). It also enables them to deal with problem behavior more efficiently, establish a safe learning climate, and positive teacher-student relationships (Gordon and Burch, 2003; Larrivee, 2005; Ming-Tak and Wai-Shing, 2008; Porter, 2014; Schonert-Reichl, 2017), which in return increase students' motivation (Kunter et al., 2007), learning achievements (Oliver et al., 2011; Omoteso and Semudara, 2011), prosocial behavior, and both teachers' and students' wellbeing (Oliver et al., 2011; Schonert-Reichl, 2017).

2. Literature review

2.1. Student-centered and teacher-directed responses to student behavior

The role of teachers' communication in classroom behavior management has been repeatedly emphasized (Porter, 2014; Burden, 2016). Communication skills to respond to student behavior are considered an integral and crucial part of classroom management skills (Gordon and Burch, 2003; Ming-Tak and Wai-Shing, 2008; Roache and Lewis, 2011). Fogelgarn et al. (2020) emphasize that the way teachers talk to students directly impacts student behavior, teacher-student relationships, student autonomy, and classroom climate. The current study addresses the crucial communication skills teachers need for effective behavior management. Specifically, it focuses on behavior management communication strategies or practices, which are teachers' specific responses to student behavior. These responses include both verbal also non-verbal behaviors such as gestures, voice tone, eye contact, or body posture for they are inevitably tied to teachers' communication practices in the classroom (LaBelle et al., 2013).

As suggested by Porter (2014), teachers' use of communication practices or strategies depend on their approach to classroom

management. These approaches vary from the traditional, teacher-directed behavioral models to the more recent humanistic student-centered models (Alcruz and Blair, 2022). Hart (2010) suggests that the former includes the use of communication strategies such as praise or punishment to reinforce or reduce student behavior. The latter include communication strategies such as listening to views of students, or non-directive non-judgmental language (Larrivee, 2005) to provide students with space to regulate their own behavior.

A recent meta-review shows a growing trend away from teacher-directed approaches promoting compliance and obedience to a more holistic and student-centered approach to classroom management, promoting students' self-regulation and autonomy (Freiberg et al., 2020). The student-centered approach is rooted in humanistic psychology (Rogers, 1969; Gordon and Burch, 2003) and emphasizes the importance of both teachers' and students' needs; it promotes mutual trust, respect, connectedness, shared responsibility, student self-regulation, positive relationships, and safe and positive school climate (Freiberg and Lamb, 2009).

Student-centered communication practices are associated with a positive impact on student behavior, relationships, and learning. Cornelius-White (2007) posited that empathy and non-directivity helps teachers eliminate power struggles, which supports students in their learning and behavior. His meta-analysis reported an association of student-centered education with significant increases in participation, satisfaction, and learning motivation. It also reports effects on self-esteem and social skills, reduction in drop-out, disruptive behavior, and absences (Cornelius-White, 2007). Establishing student-centered classrooms also promotes positive teacher-student relationships, a safe classroom climate, student academic achievements, and social and emotional development (Weinstein and Romano, 2018; Gokalp and Can, 2021). Moreover, it also allows teachers to focus more on student academic and social-emotional outcomes and spend less time managing student behavior (Talvio et al., 2014).

2.2. Proactive and reactive communication practices

Teachers' classroom management communication practices include a wide range of proactive and reactive strategies, both having a different impact on student behavior (Hepburn and Beamish, 2019; Paramita et al., 2021). Proactive strategies emphasize the prevention of classroom disturbances through building relationships, actively engaging students, creating classroom rules, or establishing a safe learning environment. These strategies have wide support in research studies suggesting that proactive teachers have classrooms with fewer disturbances, engaged and motivated students, and higher student achievements (Clunies-Ross et al., 2008; Hepburn and Beamish, 2019). Reactive strategies are those that teachers use directly in response to student behavior. Most studies focusing on reactive strategies include among these disciplinary interventions such as punishments, giving warnings, threatening, administering consequences, making sarcastic comments, yelling angrily at students, reprimands, or directives (Lewis et al., 2005; Paramita et al., 2021). Clunies-Ross et al. (2008) suggest that reactive strategies include mostly negative teachers' responses and increase problem behavior.

Miller et al. (2000) support these findings suggesting that students express problem behavior in response to teachers' unfair, harmful, or aggressive behavior. Also, these classroom management strategies are not effective in the long term and have a negative impact on teacher-student relationships, student autonomy, achievements, and teachers' wellbeing (Lewis et al., 2005; Hepburn and Beamish, 2019; Paramita et al., 2021).

Authors repeatedly emphasize that for effective classroom management, teachers need to implement proactive classroom management strategies (Hepburn and Beamish, 2019; Paramita et al., 2021), however, there are situations when teachers need communication strategies to respond to student behavior directly when it occurs. Although reactive strategies are frequently referred to as teachers' negative responses to student behavior, numerous authors suggest that there are communication responses teachers can use directly in response to student behavior, which promote teacher-student relationship, support student autonomy, social and emotional skills, engagement, self-concept, and decreases student misbehavior (Gordon and Burch, 2003; Larrivee, 2005; Porter, 2014; Talvio et al., 2014). Madden and Senior (2017) label these strategies as responsive rather than reactive and emphasize that these strategies do not always have a negative impact on students; they can be used in accordance with the student-centered approach and can be effective in helping students with their behavior issues and providing the students with space to self-regulate. Gordon and Burch (2003), for instance, include among these I-messages or Active listening. As opposed to the traditional view of reactive strategies, these strategies are reported as effective in the long term (Porter, 2014). Talvio et al. (2014) build on Gordon and Burch's ideas and suggests that these strategies such as I-messages or active listening support student social and emotional development for teachers' effective communication skills are integrated into the social and emotional competence model (Talvio et al., 2014).

2.3. A responsive student-centered approach to behavior management

Effective behavior management practices are widely researched and reviewed. Research reviews and meta-analyses focus mostly on general classroom management practices (Korpershoek et al., 2016; Hepburn and Beamish, 2019), or programs (Oliver et al., 2011; Korpershoek et al., 2016; Freiberg et al., 2020). Further, research reviews focus on specific classroom management strategies in accordance with the behavioral approach, such as Positive Behavioral Interventions and Support (PBIS) practices (Hepburn and Beamish, 2019; Estrapala et al., 2020). However, little attention is given to the student-centered or humanistic approach to behavior management and to student-centered communication strategies that are used directly in response to student behavior. Although these responsive student-centered strategies have broad theoretical support (Larrivee, 2005; Hattie, 2009, 2012; Porter, 2014; Burden, 2016), the empirical support is somewhat limited. According to our search in Web of Science, Scopus, and Education Resources Information Center (ERIC) databases, there were not found any literature reviews or meta-analyses that would address the humanistic communication strategies teachers employ to

manage student behavior. Our search covered the period from 1995 to 2021, and used several key words, including "student-centered," "literature review," "systematic review," "meta-analysis," "behavior management," and "communication strategies." For such an approach yields positive results (Porter, 2014), a literature review gathering these strategies warrants to be interesting for educators and researchers who want to employ these strategies in teacher preparation programs or address these teaching strategies in research studies. The review presented here offers a position from the student-centered approach, aiming to identify a broader range of effective student-centered communication strategies teachers can use in response to student behavior to reduce the problem behavior or increase student engagement in the classroom. Additionally, this review found the impact that these student-centered behavior management communication strategies have on students besides the reported impact on their behavior and engagement. It further describes the student-centered communication strategies, which lead to a decrease in student problem behavior or an increase in student behavior engagement. It also presents the further impact of these communication practices on students and characterizes the reviewed studies.

1. What characterizes student-centered behavior management communication strategies that lead to lower behavior problems or increased behavior engagement?
2. What further impacts do the student-centered behavior management communication strategies have on students?

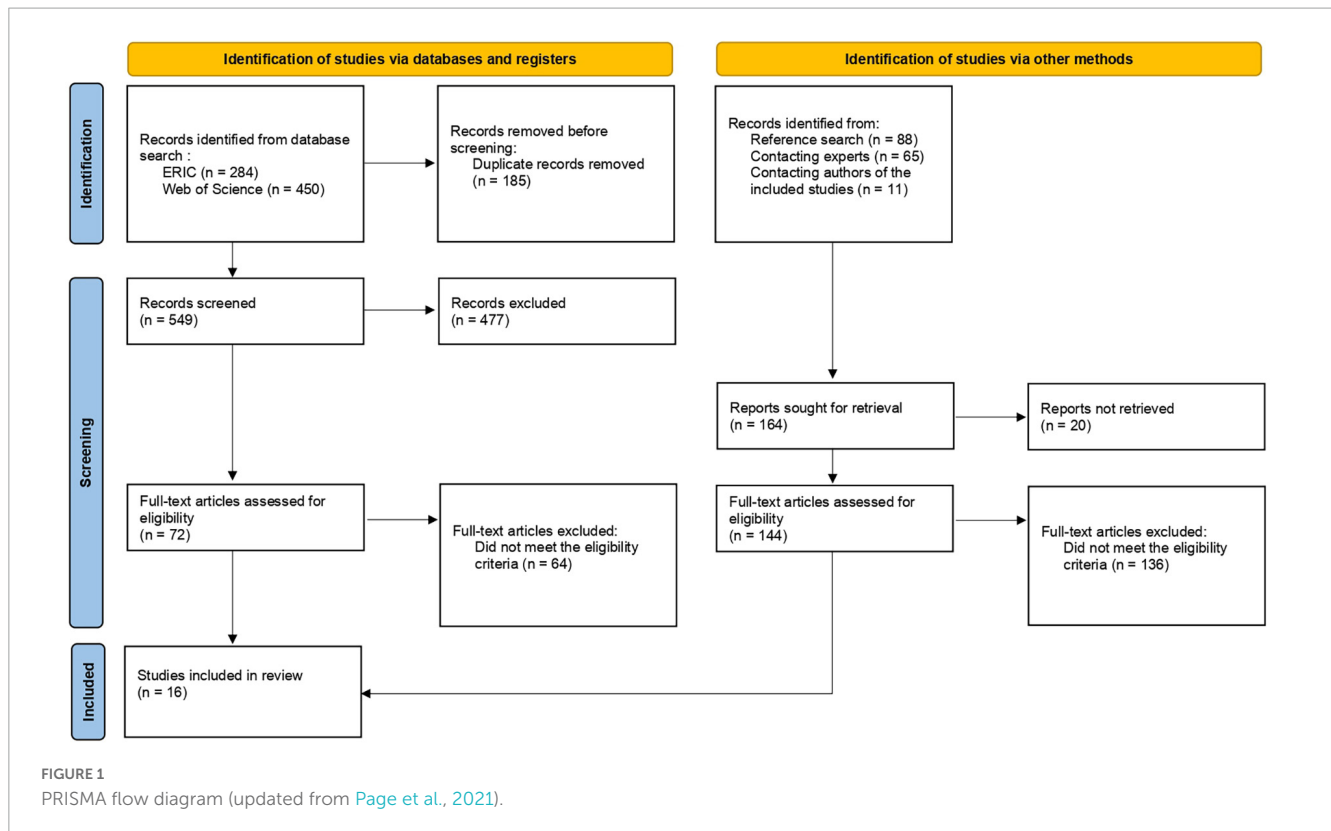
3. Methods

3.1. Data collection and evaluation

This study is a literature review using a narrative approach (Snijlsteit et al., 2012), based on the principles of the systematic review (Booth et al., 2016) to add new insights and recommendations while providing transparency and clarity and formal guidance on the method. Using this approach, this study provides a descriptive overview of the studies on teachers' student-centered communication practices used to effectively deal with student behavior.

A literature search following the PRISMA protocol (Booth et al., 2016) was conducted, using three individual searching methods based on Randolph (2009) literature review methodology: search in databases, reference search, and contacting experts. Before the search, a set of inclusion criteria was designed and piloted by two sample searching processes; subsequently, the criteria were slightly changed to yield reliable results.

The whole process of data collection and evaluation can be seen in Figure 1. The first stage of data collection was an electronic search of academic databases; Web of Science and ERIC databases were used. ERIC and Web of Science (WOS) databases were chosen for our literature review search due to their comprehensive coverage of academic literature in the field of education, as well as their sophisticated search tools which allowed us to refine our search terms and tailor our search to our specific topic. In addition, WOS is known for its high-quality indexing of articles and its citation tracking, which ensures the access to the most relevant and



up-to-date literature on our educational topic. For the initial search, the keywords “classroom management,” “behavior management,” and “strategies” were used and combined using the Boolean operator “OR.” This means that the search results would include articles that contain any one of these keywords. Subsequently, the search was refined five times by adding keywords: “humanism,” “effective,” “autonomy,” “misbehavior,” and “techniques.” These keywords were also combined using the “OR” operator, which means that the search results would include articles that contain any one of these additional keywords in addition to the initial keywords. As noted by [Vongalis-Macrow \(2009\)](#), in the 90’s, globalization emerged as a great impetus for educational reforms. Therefore, a systematic review search was limited to peer-reviewed journals published between 1995 and 2021. This time frame was chosen as it captured a period of significant change and innovation in the field of education, including a growing interest in evidence-based practices ([Hargreaves, 2007](#)). Additionally, this approach allowed for the retrieval of studies that continue to be influential in current education research. The Web of Science search was limited to four subject areas: Education, Psychology, Communication, and Behavioral Sciences.

The initial search identified 734 articles. Of these, 185 duplicates were removed, and 477 were excluded after reading the abstract based on the following inclusion and exclusion criteria: Inclusion criteria: (1) the study reported the impact of communication strategies on student engagement, or behavior, (2) the communication strategy was in accordance with the principles of student-centered humanistic approach; (3) the study was conducted in an educational context, including early childhood, elementary, secondary, and higher education contexts, (4) the study was a research article or a review. Exclusion criteria: (1)

books, or other sources that were not a research article or a review, (2) studies carried out in a non-education context, (3) the communication strategies used were in accordance with or grounded in the behaviorist or teacher-directed approach, (4) the study did not report the impact of communication strategies on student behavior or engagement. Seventy-two articles remained for further examination. The process of inclusion of studies was conducted with another researcher to ensure the quality of the search and inclusion process. A full-text review resulted in five studies meeting the criteria.

Second, as [Randolph \(2009\)](#) recommends, the references of the most relevant full-text reviewed articles were retrieved. Those references, which seemed relevant were found, and their references were read. This process was repeated until no new articles came to light. This search resulted in 88 articles examined; three of them were included based on the inclusion criteria.

After database and reference searching, 28 most-cited experts in the field of classroom management, classroom interaction, behavior management, or autonomy support were contacted *via* email. They were shared with a brief explanation of the problem formulation, literature search objectives, and preliminary findings. The experts were asked about any further references that might be included in the literature review and other further ideas or recommendations regarding the researched topic. A total of 65 articles were gathered from the nineteen responses that were obtained. All of them were full text examined. After excluding already included articles and applying inclusion and exclusion criteria, five studies were included in the review.

Further, the authors of the studies intended for inclusion in the review were contacted. The findings were shared with them, and a brief description of the study was provided, followed by a

request for a critical view on the study's inclusion. They were asked about any further recommendations for studies to be included in the review. In this email, I also asked for additional pieces of information, which were missing in the studies. For the main aim of the study was to provide a picture of student-centered communication practices and provide their individual descriptions, each of the authors was asked to provide a brief quote of the communication strategy listed in the included research articles. Fifteen responses were gathered throughout this process, which helped us to revise the studies and create a table with each communication strategy supported by a quote and a description. This process resulted in ten further studies, of which one was included in the review.

3.2. Quality appraisal

The quality of the studies included in this review was assessed using an adapted version of the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018). This tool was chosen as it is specifically designed for reviews that include a combination of qualitative, quantitative, and mixed-method studies. Two researchers independently evaluated the quality of the selected studies using the MMAT. The tool includes two screening questions for all research designs and five specific criteria for each research design category, such as qualitative research, quantitative randomized controlled trials, quantitative non-randomized studies, quantitative descriptive studies, and mixed-method studies. Each criterion is rated as “Yes,” “No,” or “Can’t tell.” In this review, the reviewers agreed to use only “Yes” or “No” ratings. To report the quality of the studies was used a combination of stars (*). Each star stands for 20% (e.g., three stars stand for 60%, four stars stand for 80%, five stars stand for 100%). Each criterion was assigned a value of one star or 20%. For example, a study that fulfilled three criteria would receive three stars, which equals 60%. Ratings for each study and criterion and the overall score can be found in Figure 2.

3.3. Data analysis and interpretation

For the data analysis and interpretation, the review used a content analysis approach (Silverman, 2020). It enabled us to include and analyze both qualitative and quantitative studies and code the studies by drawing on a theoretical framework underlying this study. The codes were created both deductively and inductively. The studies were coded to capture the general and detailed description of the communication practices used in the studies. The codes were further inductively classified and categorized.

To analyze and report the description of student-centered communication practices affecting student behavior were created categories to reflect their key patterns, similarities, and differences. These categories were based on a qualitative content analysis of the description of each communication strategy; the descriptions were provided in the research studies or obtained directly from the studies' authors. Creating the individual descriptions of the strategies and their categorization was done by two individual researchers through constant interaction between the research studies' content, the theoretical framework supporting the studies, and communication with the authors. The process of the analysis can be seen in [Supplementary Table 1](#).

4. Results

4.1. Characteristics of the reviewed studies

Altogether, 16 studies were included in the in-depth review. [Table 1](#) provides key information for each research paper. Studies were conducted in USA (9), Belgium (2), Israel (2), South Korea (1), Australia (2), China (1), and Ireland (1). One study was conducted in three countries. The studies were

No.	Author (year)	Study design	S1	S2	1	2	3	4	5	Quality (*****)	%
1	Aelterman et al., 2018	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	No	Yes	***	60%
2	Assor et al., 2002	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	No	Yes	***	60%
3	Bucnker & Frisby, 2015	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	No	Yes	***	60%
4	Cheon et al., 2019	Quantitative randomized controlled trials	Yes	Yes	No	Yes	Yes	Yes	Yes	****	80%
5	Chesebro & Lyon, 2020	Qualitative	Yes	Yes	Yes	Yes	Yes	Yes	Yes	*****	100%
6	Goodboy & Myers, 2008	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	No	Yes	***	60%
7	Jang et al., 2010	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	Yes	Yes	****	80%
8	Madden & Senior, 2017	Mixed methods	Yes	Yes	Yes	No	Yes	Yes	No	***	60%
9	Miller et al., 2013	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	Yes	Yes	****	80%
10	Peterson et al., 1979	Quantitative non-randomized studies	No	Yes	No	Yes	Yes	No	Yes	***	60%
11	Roache & Lewis, 2011	Quantitative non-randomized studies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	*****	100%
12	Romi et al., 2009	Quantitative non-randomized studies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	****	80%
13	Vansteenkiste et al., 2012	Quantitative non-randomized studies	Yes	Yes	No	Yes	Yes	Yes	Yes	****	80%
14	Wallace et al., 2014	Quantitative descriptive studies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	*****	100%
15	Weger, 2017	Quantitative non-randomized studies	Yes	Yes	Yes	Yes	No	Yes	Yes	****	80,00%
16	Worley et al., 2007	Qualitative	Yes	Yes	Yes	Yes	Yes	Yes	Yes	*****	100%
Total score											77,50%

FIGURE 2

Overview of the quality appraisal ratings using the MMAT.

TABLE 1 Overview of the reviewed studies.

No.	References	Country	Communication strategies	Variable	Findings	Context	Measurement approach	Quality appraisal
1	Aelterman et al., 2018	Belgium	Discussing rules with students	Misbehavior	Decreased level of misbehavior.	Secondary school	Quantitative	60%
2	Assor et al., 2002	Israel	Fostering relevance Allowing criticism Providing choice	Behavioral engagement	Fostering relevance: increase in positive feelings, and behavioral and cognitive engagement Providing choice: increase in positive feelings, no increase in behavioral and cognitive engagement Allowing criticism: no increase in positive feelings, no increase in engagement	Elementary school (8–14 years), grades 3–8	Quantitative	60%
3	Buckner and Frisby, 2015	USA	Confirmation	Instructional dissent	Lower expressive and vengeful dissent	University	Quantitative	60%
4	Cheon et al., 2019	South Korea	Taking students' perspectives Empathic Listening Acknowledging and accepting student negative affect Provide explanatory rationale Invitational language	Prosocial behavior Academic success Peer task climate Need satisfaction	Increased prosocial behavior, academic outcomes, peer task climate and need satisfaction.	Secondary school	Quantitative	80%
5	Chesebro and Lyon, 2020	USA	Facilitating discussion Providing explanations Staying composed and respectful Managing out of class	Facilitating discussion Providing explanations Staying composed and respectful Managing out of class	Perceived by students as effectively managed incidents	College	Qualitative	100%
6	Goodboy and Myers, 2008	USA	Confirmation	Misbehavior Participation in class Learning outcomes Motivation	Decreased level of misbehavior, increased participation in class, learning outcomes and motivation.	College	Quantitative	60%
7	Jang et al., 2010	USA	Non-controlling informational language Acknowledging students' perspectives and feelings	Student engagement	Increased student engagement.	Secondary school (High school)	Quantitative	80%
8	Madden and Senior, 2017	Ireland	Acknowledging students' feelings	Misbehavior	Decreased level of misbehavior.	Elementary school	Quantitative	60%
9	Miller et al., 2014	USA	Self-disclosure Non-verbal immediacy	Classroom incivility	Decreased level of classroom incivility	College	Quantitative	80%
10	Peterson et al., 1979	USA	I-messages	Student study behavior Disruptions	Increased student study behavior, decreased disruptions.	Elementary school	Quantitative	60%

(Continued)

TABLE 1 (Continued)

No.	References	Country	Communication strategies	Variable	Findings	Context	Measurement approach	Quality appraisal
11	Roache and Lewis, 2011	Australia	Hinting Recognize appropriate behavior Punishment Discussing the impact of student behavior on others Aggressive techniques	Student responsibility Level of misbehavior	Hinting, recognizing appropriate behavior and discussing with students: increased student responsibility and decreased level of misbehavior. Aggressive behavior and punishments increased level of misbehavior and decreased student responsibility.	Secondary + primary schools	Quantitative	100%
12	Romi et al., 2009	Australia China Israel	Hinting Recognize appropriate behavior Punishment Discussing the impact of student behavior on others Aggressive techniques	Student responsibility Level of misbehavior	Hinting, recognizing appropriate behavior, discussing with students: increased student responsibility and decreased level of misbehavior. Aggressive behavior and punishments increased level of misbehavior and decreased student responsibility.	Secondary schools	Quantitative	80%
13	Vansteenkiste et al., 2012	Belgium	Clear expectations	Problem behavior Self-regulated learning outcomes	Decreased problem behavior, increased self-regulated learning.	Secondary school	Quantitative	80%
14	Wallace et al., 2014	USA	Active Listening Taking students' perspectives Accepting students' negative affect Provide rationale Open communication Questioning	Interpersonal climate Student engagement Motivation Students' affective responses	Increased interpersonal climate, student engagement, motivation, and affective responses.	Middle school	Mixed	100%
15	Weger, 2017	USA	Active empathic listening	Misbehavior	Decreased level of misbehavior	Undergraduate students	Quantitative	80,00%
16	Worley et al., 2007	USA	Active listening enthusiasm	Classroom climate	Safe classroom climate	Graduate, undergraduate students	Qualitative	100%

conducted in the context of elementary schools (4), secondary schools (7), colleges (3), and universities (3). One study was conducted in two educational contexts. The prevailing study design in the reviewed studies was quantitative (13), three studies used a qualitative methodology design, and one combined qualitative and quantitative instruments to assess the impact of student-centered communication strategies. All studies reported the relation of student-centered communication strategies to student behavior (16). The classrooms where teachers used student-centered communication practices in response to student behavior reported a decreased level of misbehavior (12), increased prosocial behavior (1), or increased behavioral engagement (3). Further effects of the student-centered communication practices include increased students' participation in class (1), learning outcomes (2), motivation (2), study behavior (1), need satisfaction (1), responsibility (2), self-regulation (1), and interpersonal climate (3).

4.2. Responsive student-centered communication strategies

The student-centered communication strategies are further characterized into five groups (sub-sections "4.2.1. Perspectives and feelings, 4.2.2. Choice and autonomy-support, 4.2.3. Non-directive non-judgmental language, 4.2.4. Explanations and expectations, and 4.2.5. Teacher immediacy and form of the message"), as identified throughout the content analysis. The specific communication strategies gathered throughout the review can be found in [Figure 3](#). The extended version of the [Supplementary Table 2](#) also contains references and definitions for each communication strategy. [Supplementary Table 1](#) describes which communication strategy belongs to each group.

4.2.1. Perspectives and feelings

Taking students' perspectives, recognizing their feelings, and acknowledging them were among the most frequent features of communication strategies listed throughout the reviewed studies. Teachers take and acknowledge students' perspectives when they experience the classroom events as if they were the students. Teachers also communicate to students that they value and follow their perspectives even when they disagree. Further, any students' expression of negative affect is accepted as a valid reaction to classroom demands, rules, or activities ([Jang et al., 2010](#)). Authors emphasize the importance of communicating to students that teachers value and follow their perspectives and feelings even when they disagree ([Wallace et al., 2014](#); [Weger, 2017](#)).

In this review, Taking students' perspectives and Acknowledging students' feelings are reported in one category as they frequently overlap and are used as one communication strategy ([Jang et al., 2010](#); [Vansteenkiste et al., 2012](#); [Cheon et al., 2019](#)). For instance, [Wallace et al. \(2014\)](#) term these strategies together as Open communication. Other authors emphasize the importance of recognizing and acknowledging students' perspectives and feelings when using other communication strategies. [Aelterman et al. \(2018\)](#) conclude that it is key to ask about students' feelings and acknowledge them when discussing rules with students. [Goodboy and Myers \(2008\)](#) and [Buckner](#)

and [Frisby \(2015\)](#) also include in their study communication strategies that communicate to students that they are recognized and acknowledged as valuable individuals with their needs and affects; they use the term Confirmation. Similarly, teachers can recognize and acknowledge students' perspectives and feelings when providing clear expectations for behavior, asking questions, or actively listening ([Vansteenkiste et al., 2012](#); [Wallace et al., 2014](#); [Weger, 2017](#)).

Besides acknowledging students' feelings and perspectives, the authors also point to the importance of supporting students to see the perspectives of others. [Roache and Lewis \(2011\)](#), [Miller et al. \(2014\)](#), and [Wallace et al. \(2014\)](#) emphasize the importance of talking about teachers' perspectives. Teachers use I-messages, Open communication, or Self-disclosure to describe their feelings, explain why specific behavior is unacceptable, illustrate a rule, or provide a particular rationale for expected behavior ([Peterson et al., 1979](#); [Roache and Lewis, 2011](#); [Miller et al., 2014](#); [Wallace et al., 2014](#)).

4.2.2. Choice and autonomy-support

When talking about providing students with choice and autonomy, authors repeatedly emphasize the importance of teachers' autonomy-supportive language or behavior ([Assor et al., 2002](#); [Jang et al., 2010](#); [Vansteenkiste et al., 2012](#); [Wallace et al., 2014](#); [Cheon et al., 2019](#)). They stress that it is crucial to communicate with the students in an autonomy-supportive way as frustrating the need for student autonomy may lead to negative behavioral, motivational, and affective outcomes. [Vansteenkiste et al. \(2012\)](#) conclude that autonomy is not about unlimited freedom where students lack sufficient guidance. Instead, teachers adopt the autonomy-supportive language and provide students with structure (such as communicating requests and clear expectations) in an autonomy-supportive way (e.g., through adopting their perspectives, providing them with choice, and accepting their negative affect).

The studies include among autonomy-supportive language communication strategies termed Open communication, Providing choice, Acknowledging and accepting students' negative feelings, Taking students' perspectives, Invitational language, Discussing the rules with students, Hinting, and Clear expectations ([Figure 2](#)). More specifically, in four studies, authors emphasize the importance of providing students with the desired amount of choice ([Assor et al., 2002](#); [Vansteenkiste et al., 2012](#); [Wallace et al., 2014](#); [Cheon et al., 2019](#)). [Vansteenkiste et al. \(2012\)](#) and [Cheon et al. \(2019\)](#) use the term Inviting language to talk about language, which promotes choice and volition (you might, you can) instead of control (you should, you must, you need to). Further, the authors emphasize the importance of giving students freedom of mobility, opportunities to collaborate, allowing them to work in their own way, or giving them the responsibility to use their own solutions to solve any issue without teacher control. [Romi et al. \(2009\)](#) use the term Hinting to talk about communication strategy which informs students that they are breaking some rule, without telling them explicitly what to do; thus, providing them with autonomy to solve the situation themselves. In contrast, [Assor et al. \(2002\)](#) conclude that the role of choice and freedom of action is less important than the extent to which students perceive the requests, rules, or task as personally meaningful. Therefore, it might be crucial to combine the desired amount of choice with supporting students to find their

The name of the strategy		Example of the strategy
1.	Taking students' perspectives	"Are you okay? What do you need?" "I understand that this material is new and unfamiliar and a bit harder"
2.	Active Empathic Listening	"Jane, you seem annoyed or frustrated by this activity and that is perfectly reasonable and ok."
3.	Acknowledging and accepting students' negative feelings	"Some of you look pretty anxious and stressed. Is that right?" "Peter, you seem frustrated by the reading quizzes. Is it too difficult for you?"
4.	Provide an explanatory rationale	"Let's follow the rule to not interrupt the person who is currently speaking. It helps us create a safe and supportive classroom for everyone."
5.	Invitational language	"You might like to take a break. It may help you concentrate."
6.	Non-verbal immediacy	Maintaining an open body posture, using a friendly and calm tone of voice, smiling, establishing eye contact with students
7.	Relevant self-disclosure	"As an illustration of that rule, I'll give you an example that happened to me. . ."
8.	Discussing the rules with the students	"What can we do to make our classroom a safe place for learning?"
9.	Providing meaningful choice	"You may work by yourself, with a partner, or in a small group, but please, make sure you do not interrupt the other students by your talking."
10.	I-statements	"When you are speaking when I am giving the instructions, I cannot concentrate on what I am saying. I feel then annoyed."
11.	Non-controlling informational language	"I have noticed you are having some difficulty concentrating today. Do I understand that correctly? Do you know why that might be?"
12.	Hinting	"There is a lot of noise in the classroom. We agreed that one person speaks at a time."
13.	Discussing the impact of student behavior on others	"When you are talking, some students cannot hear the instructions."
14.	Fostering relevance	"If we use respectful language, we help to create a classroom that is welcoming, safe, and supportive for everyone. How would you like our classroom to be like?"
15.	Remaining adaptive in practice	"I like the way you are working, there is still a lot of noise in the classroom. Is it that you need more time to talk about the topic?"
16.	Promoting independent thinking	"I want you to think about what would help you to concentrate. Do you need a break? Or perhaps you need to change the activity?"
17.	Using open communication	"I know you might be upset because the material is new and difficult. I gave it to you because... Now I would like you to..."
18.	Questions	"Are you okay? What do you need?" "What are you doing?"
19.	Clear expectations	"Staying respectful might be difficult all the time, let's use respectful language, because we want a classroom that is welcoming, safe, and supportive for everyone."
20.	Confirmation	"Okay, you think that this activity is boring. I appreciate to hear your opinion. I am sure we can find a way to go through this activity to make it more relevant for you."
21.	Autonomy support	"I understand. You think that this task is too difficult and that makes you upset. We still need to do a bit more practice. Do you want to continue working on this task—or switch to another?"
22.	Discussing conflicts with students	"Can you help me understand what happened in the class today?"
23.	Staying composed and respectful	"That there is a lot of noise in the classroom and we cannot hear each other. Let's either stay in the groups using moderate voice or work individually."
24.	Managing out of class	"You can come to me after the lesson and we can find some solution."

FIGURE 3
Strategies identified throughout the review.

own perspectives, values, and relevance of the expected behavior, for instance when discussing the rules with students (Aelterman et al., 2018).

4.2.3. Non-directive non-judgmental language

The non-directive and non-judgmental language was frequently mentioned and contrasted with evaluative, controlling, pressuring, and coercive teacher talk. Jang et al. (2010) state that by using non-controlling informational language, teachers provide the students with a sense of ownership, choice, and responsibility

for solving the problem and regulating their behavior. Likewise, Vansteenkiste et al. (2012) and Cheon et al. (2019) use the term Invitational language to speak about teachers' communication responses that minimize pressure on students by incorporating options and choices instead of commands and directives.

The review showed that most of the studies emphasized non-directive and non-judgmental language as part of the communication strategies listed (e.g., Active Listening, I-messages, Hinting, Non-controlling language, and Autonomy-supportive language). Specifically, Romi et al. (2009) and

Roache and Lewis (2011) use in their research a communication strategy called Hinting: a non-judgmental non-directional description of student behavior. They argue that unlike controlling teachers' statements, such as commands, hints offer students space for their responsibility to deal with their problems. Roache and Lewis (2011) distinguish various types of hints, such as specific hints (addressing a specific behavior), general hints (describing a problem situation in a general way), a restatement of expectations (re-emphasizing rules), I-messages, and direct questions. Wallace et al. (2014) also include Questioning as a communication strategy in their study and argue that it offers students space to respond or change their behavior before offering any comments, directives, or suggestions. Further, Weger (2017) points to Active listening as a key communication strategy and emphasizes that its most important part is that it is non-judgmental. Similarly, Peterson et al. (1979) point to the importance of non-judgmental informational language. They contrast I-messages with You-messages, claiming that I-messages contain no judgment and no directive for the students. Using I-messages, teachers describe students' disruptive behavior, state their feelings in response to the student's behavior, and indicate why students' behavior negatively affects the teacher (Peterson et al., 1979; Gordon and Burch, 2003).

4.2.4. Explanations and expectations

Providing students with explanations and expectations was another broad category of communication practices. When dealing with student behavior, teachers can communicate to students expectations, rules, and explanations for requested behavior; this way, teachers introduce their requests and rules by explaining their value and importance so that the students will see them as personally meaningful (Wallace et al., 2014; Cheon et al., 2019; Chesebro and Lyon, 2020). Romi et al. (2009) and Roache and Lewis (2011) conclude that when teachers set expectations and work out the rules for appropriate behavior with the students, they feel more responsible for their behavior. Aelterman et al. (2018) and Cheon et al. (2019) support this idea and argue Providing rationale is one of the communication strategies that help students feel not being pressured or controlled; instead, students can perceive behavioral requests as more personally meaningful and are likely to internalize the value. Further, Miller et al. (2014) connect explanatory rationale with teachers' life stories or experiences. They state that teachers can add relevant, friendly, and personal disclosure (e.g., their personal or professional experiences) to the explanation of rules or expected behavior to show their relevance.

Authors point to the importance of explaining and explicitly describing the impact of student behavior on others. Peterson et al. (1979) and Romi et al. (2009) describe I-messages and Hinting as communication strategies which explain why student behavior interferes with teachers' or students' needs and allow students to respond to the problem situation with no pressure or directive from the teacher. Wallace et al. (2014) and Chesebro and Lyon (2020) further note that when teachers explain why certain behavior interferes with their needs or with the learning process and what impact it can have, students may choose to behave differently for they find it personally meaningful.

Besides explanations, Vansteenkiste et al. (2012) state that teachers need to provide clear expectations for desirable classroom behavior. When giving expectations, teachers state the rule, instruction, or target behavior and transfer responsibility for the

action to students, leaving them accountable for their learning and behavior. However, they add that teachers may communicate the expectations in an autonomy-supportive way and in a non-judgmental manner, for instance, by providing explanatory rationale or using non-controlling language. Wallace et al. (2014) support this idea by suggesting that it is essential to give clear expectations and steps to follow when requesting a rule or action, however, it should always be accompanied by an explanatory rationale.

4.2.5. Teacher immediacy and form of the message

All the communication strategies listed differed in the form of the messages sent and in the closeness or immediacy of teacher behavior. The ways of communicating immediacy include appropriate use of eye contact, gestures, voice tone, smile, movement, physical contact, or physical closeness. It is any behavior that indicates physical or psychological closeness (Weger, 2017). Wallace et al. (2014), Cheon et al. (2019), and Chesebro and Lyon (2020) highlight the importance of teachers' voice tone when using Invitational language, asking questions, or using a Composed and respectful response. Invitational language is sent in a higher pitch to communicate understanding and support; Questions use a tone that conveys respect and implicit rationale of the message; Chesebro and Lyon (2020) stress the importance of sending the message in a tone that is composed, respectful and calm.

Eight main types of the teachers' messages were distinguished regarding the form of teacher talk: Asking questions (Romi et al., 2009; Wallace et al., 2014; Chesebro and Lyon, 2020), Offering advice (Wallace et al., 2014), Listening (Worley et al., 2007; Wallace et al., 2014; Weger, 2017; Cheon et al., 2019), Providing request (Vansteenkiste et al., 2012; Cheon et al., 2019; Chesebro and Lyon, 2020), Giving a statement (Peterson et al., 1979; Romi et al., 2009; Roache and Lewis, 2011), and Discussing or facilitating privately or out of class (Wallace et al., 2014; Aelterman et al., 2018; Chesebro and Lyon, 2020).

The two broadest categories were Questions and Listening. The authors identified various aims of asking the questions. Teachers ask questions to get clarification from the students about any issue, behavior, or affect they experience (Wallace et al., 2014; Weger, 2017). Questions can also point to problem behavior in the classroom, asking the students about the rules in the classroom or the expected behavior (Roache and Lewis, 2011). Wallace et al. (2014) suggest that teachers can help students think through problems by asking them questions, which help them to see things from different perspectives and find individual ways to solve an issue. The form of the question is always descriptive (What happened?) or corrective (What can we do with that?).

Listening to students was another broad category, which involves attentively listening to students' feelings and ideas. During active listening, teachers take students' perspectives and try to understand students' points of view. Teachers paraphrase what the students are saying to clarify understanding and send students a message that they understand and accept their feelings. The authors list active listening as a communication strategy to use during conflict resolutions or when involving students in decision-making (Worley et al., 2007; Weger, 2017). Romi et al. (2009) and Roache and Lewis (2011) include active listening as an essential component when teachers discuss the impact of student behavior on others.

5. Discussion and conclusion

The present research review aimed at providing a deep characterization of student-centered communication strategies, which lead to lower behavior problems in the classroom. Sixteen studies were identified that demonstrated the positive impact of student-centered communication strategies on student behavior. Five main characteristics that these communication strategies have in common were found. The prevailing characteristics of the student-centered communication practices were that they are non-judgmental and non-controlling, they recognize and accept students' perspectives, provide choice, if possible, give clear expectations for behavior, explain the reason for any request or rule, and are sent in a calm manner. These characteristics can be seen as principles that can help formulate student-centered responses to student behavior. Further, with the help of the authors of the reviewed studies, a table with specific twenty-four communication strategies with their brief descriptions and sample statements was created.

Notably, the studies' analysis results further revealed that using student-centered communication strategies in the classroom leads to fewer behavior problems or greater behavioral engagement; they further promote student motivation, learning achievements, responsibility, self-regulation, and interpersonal climate. These results align with those obtained by [Cornelius-White \(2007\)](#), who emphasizes the positive impact of a student-centered approach in the classroom. This reinforces the theory of [Fogelgarn et al. \(2020\)](#), suggesting that communication of the teacher is a crucial classroom management skill that directly impacts student behavior, autonomy, and classroom climate.

Most studies focused on verbal communication strategies; however, four studies also focused on non-verbal communication practices, such as voice tone, eye contact, smile, and calmness. Non-verbal aspects of communication are less discussed, even in theoretical works on classroom management. [LaBelle et al. \(2013\)](#) emphasize that teachers' responses inevitably include non-verbal behaviors which influence the statement that the teacher sends. Similarly, [Larrivee \(2005\)](#) concludes that there needs to be a congruence between teacher's non-verbal and verbal messages. This supported the study of [Jang et al. \(2010\)](#) and [Weger \(2017\)](#); if teachers send a message describing student behavior in a calm manner, with a smile, warm voice, and patience, students are more likely to adjust their behavior and respond to the teacher's request.

One interesting aspect that emerged from the analysis is that most of the student-centered communication strategies gathered can be used directly in response to student behavior. These results support the theory of [Madden and Senior \(2017\)](#) that responsive strategies do not always negatively impact student behavior and can be used in accordance with a student-centered approach to help students with their behavioral issues. For instance, teachers can use I-messages or Hinting to describe student behavior in a non-judgmental, non-directive manner. Similarly, providing an explanation of a teacher's request or a rule can be used directly in response to student behavior while making it meaningful to the student and thus more likely that the students will internalize the value and adapt their behavior.

Further, the descriptions of the communication practices showed that student-centered communication strategies derive

only slightly from teacher-directed methods. For instance, student-centered questions ([Wallace et al., 2014](#)) are sent in a tone that conveys respect, aiming to get additional information, understand student affect and behavior, or support students in their independent thinking (What happened? What can we do with that?) In comparison, [Gordon and Burch \(2003\)](#) and [Larrivee \(2005\)](#) emphasize the negative impact of questions, which are judging, threatening, or not accepting the students (Why did you...? Why are you...?). Similarly, the statement "You must take notes" is directive. [Porter \(2014\)](#) suggests that these statements lead to feelings of resentment and frustrate students' need for autonomy. On the other hand, using invitational language ("You might take notes") changes the statement into a non-directive, student-centered response. The studies which used this communication strategy reported increased prosocial behavior, interpersonal climate, and need satisfaction ([Cheon et al., 2019](#)). Thus, applying the five principles of student-centered communication strategies when responding to student behavior could potentially make the responses more relevant to students and might lead to a safe learning climate, increased student motivation, need satisfaction, learning achievements, responsibility, and self-regulation, and lower behavior problems in the classroom.

5.1. Implications for teaching and learning

The findings from this study have several implications for teaching and learning. Teachers worldwide spend a significant amount on managing student behavior, claiming they have insufficient skills to solve these situations ([Kwok, 2020](#)). The findings of this study may support teachers and student teachers to gain competencies to manage student behavior effectively. The results show the importance of teachers' student-centered responses to student behavior, define student-centered communication in its complexity and show specific communication strategies with their definition and sample responses. Teacher training should focus on these responses, illustrating to students the multiple ways to respond to student behavior effectively and the impact these communication practices may have on students. These materials can serve as a professional tool for teachers, who can learn about the communication strategies supported by research and use them in their practice. Further, the findings may inform teacher educators who can use them in their classroom management courses. The study provides qualitative examples and characteristics of both verbal and non-verbal responses for teachers and student teachers to use when dealing with problem situations in the classroom to solve them successfully with a positive impact on students and their behavior.

5.2. Limitations and further research

Whereas [Freiberg et al. \(2020\)](#) found a growing trend away from the teacher-directed approach, the current database search showed a rather small number of studies focusing on the student-centered behavior management approach or on student-centered

communication strategies specifically. This may be partly due to the focus of the study and the choice of criteria: the search focused on studies that concentrated specifically on teachers' communication strategies rather than on their classroom management strategies in general and included only studies conducted in the educational context. Future research could explore the topic of student-centered or humanistic communication in a broader context. On the other hand, the low number of studies included in the exploration may also indicate the strong need for further research on student-centered behavior management communication strategies in the classroom.

The inclusion and exclusion criteria were defined precisely and in advance to decide which study to include in the current review. All attempts were made to perform a quality selection of the studies; three individual searching methods were used and the results were constantly consulted with the experts in the field; however, it is acknowledged that even with two independent reviewers, the selection of the studies is at risk for subjectivity.

It should also be noted that the studies were conducted in different countries and cultures. The linkage of the findings of each study to a certain cultural context (and the highly culturally bound issue of classroom management strategies) is acknowledged. When interpreting the results, this must be taken into consideration.

Also, the main aim of the study was to concentrate on the characteristics of student-centered communication strategies gathered throughout the review. Further analysis is needed for a more detailed picture of the included studies and their research designs and methods. Besides, it was only briefly commented on the impacts of communication strategies on students. Additional research might explore these impacts in more detail. Some of the reviewed studies used a set of communication strategies and researched their impact on students. It is acknowledged that the specific communication strategies that had the greatest potential to impact student behavior cannot be distinguished. When interpreting the findings, this is important to note. Further research could be focused on the impact of each specific communication characteristic on student behavior.

5.3. Conclusion

This study makes several contributions. First, it characterizes the complexity of student-centered communication that is used to respond to student behavior. Next, it presents specific communication strategies with definitions and sample statements. It also describes the impact of these communication practices on student behavior, engagement, and other areas. Overall, the study reaffirms the importance of student-centered communication responses to problem behavior in the classroom; it defines the communication that can teachers and student teachers use to handle problem situations in the classroom successfully with a positive impact on students and their behavior.

Data availability statement

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

Author contributions

JK contributed to the conception and design of the study, wrote the first draft of the manuscript, and conducted the systematic review data gathering using the PRISMA protocol. JN wrote the sections of the manuscript and revised it critically. JN and JK contributed to the analysis of the data and their interpretation. Both authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This research was partially supported by Aktuální výzvy v profesi učitele 21. Století (Current challenges in the 21st Century Teaching Profession) (MUNI/A/1481/2020).

Acknowledgments

We would like to express our gratitude to Emeritus Laureate Professor John Hattie for his encouragement, advice, and feedback throughout the research process. We thank Shuzhan Li, Ph.D. for his valuable feedback on the final draft of the manuscript, and Dr. Louise Porter for her insights and feedback on the manuscript. Their contributions have been invaluable to the completion of this work. We would also like to express our gratitude to all the authors of the studies included in this review for their valuable contributions to the field and for allowing us to use their work in this review. We would like to extend our special appreciation to the authors who provided additional support and advice throughout this review. Their insights and expertise have been invaluable in shaping the findings of this study.

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/feduc.2023.1156530/full#supplementary-material>

References

- Aelterman, N., Vansteenkiste, M., and Haerens, L. (2018). Correlates of students' internalization and defiance of classroom rules: A self-determination theory perspective. *Br. J. Educ. Psychol.* 89, 22–40. doi: 10.1111/bjep.12213
- Alcruz, J., and Blair, M. (2022). *Student-centered classrooms: Research-driven and inclusive strategies for classroom management*. Lanham, MD: Rowman & Littlefield Publishers.
- Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., and Trautwein, U. (2018). Student misbehavior and teacher well-being: Testing the mediating role of the teacher-student relationship. *Learn. Instr.* 58, 126–136. doi: 10.1016/j.learninstruc.2018.05.006
- Assor, A., Kaplan, H., and Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *Br. J. Educ. Psychol.* 72, 261–278. doi: 10.1348/000709902158883
- Auditor General Western Australia, (2014). *Behaviour management in schools* (no. 4). Available online at: https://audit.wa.gov.au/wp-content/uploads/2014/03/report2014_04-BehaviourMgt-Amendment072015.pdf (accessed September 20, 2022).
- Booth, A., Sutton, A., and Papaioannou, D. (2016). *Systematic approaches to a successful literature review*, 2nd Edn. Los Angeles, CA: SAGE.
- Brophy, J. (2006). "History of research on classroom management," in *Handbook of classroom management: Research, practice, and contemporary issues*, eds C. M. Evertson and C. S. Weinstein (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 17–43.
- Buckner, M. M., and Frisby, B. N. (2015). Feeling valued matters: An examination of instructor confirmation and instructional dissent. *Commun. Stud.* 66, 398–413. doi: 10.1080/10510974.2015.1024873
- Burden, P. (2016). *Classroom management: Creating a successful K-12 learning community*, 6th Edn. New York, NY: Wiley.
- Chesebro, J. L., and Lyon, A. (2020). Instructor responses to disruptive student classroom behavior: A study of brief critical incidents. *Commun. Educ.* 69, 135–154. doi: 10.1080/03634523.2020.1713386
- Cheon, S. H., Reeve, J., and Ntoumanis, N. (2019). An intervention to help teachers establish a prosocial peer climate in physical education. *Learn. Instr.* 64:101223. doi: 10.1016/j.learninstruc.2019.101223
- Clunies-Ross, P., Little, E., and Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educ. Psychol.* 28, 693–710. doi: 10.1080/01443410802206700
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Rev. Educ. Res.* 77, 113–143. doi: 10.3102/003465430298563
- Eisenman, G., Edwards, S., and Cushman, C. A. (2015). Bringing reality to classroom management in teacher education. *Prof. Educ.* 39, 1–12.
- Emmer, E. T., and Evertson, C. M. (2012). *Classroom management for middle and high school teachers*, 9th Edn. London: Pearson.
- Estrapala, S., Rila, A., and Bruhn, A. L. (2020). A systematic review of tier 1 PBIS implementation in high schools. *J. Pos. Behav. Interv.* 23, 288–302. doi: 10.1177/1098300720929684
- Fogelgarn, R. K., Burns, E. A., and Lewis, R. (2020). Hinting as a pedagogical strategy to promote prosocial behaviour. *Educ. Action Res.* 29, 1–17. doi: 10.1080/09650792.2020.1743333
- Freiberg, H. J., and Lamb, S. M. (2009). Dimensions of person-centered classroom management. *Theory Pract.* 48, 99–105. doi: 10.1080/00405840902776228
- Freiberg, J. H., Oviatt, D., and Naveira, E. (2020). Classroom management meta-review continuation of research-based programs for preventing and solving discipline problems. *J. Educ. Stud. Placed Risk* 25, 319–337. doi: 10.1080/10824669.2020.175454
- Gokalp, G., and Can, I. (2021). Evolution of pre-service teachers' perceptions about classroom management and student misbehavior in an inquiry-based classroom management course. *Action Teach. Educ.* 44, 70–84. doi: 10.1080/01626620.2021.1939194
- Goodboy, A. K., and Myers, S. A. (2008). The effect of teacher confirmation on student communication and learning outcomes. *Commun. Educ.* 57, 153–179. doi: 10.1080/03634520701787777
- Gordon, T., and Burch, N. (2003). *Teacher effectiveness training: The program proven to help teachers bring out the best in students of all ages*. New York, NY: Crown.
- Hargreaves, D. (2007). "Teaching as a research-based profession: Possibilities and prospects (The teacher training agency lecture 1996)," in *Educational research and evidence-based practice*, ed. M. Hammersley (Milton Keynes: Open University Press), 3–17.
- Hart, R. A. (2010). Classroom behaviour management: Educational psychologists' views on effective practice. *Emot. Behav. Difficulties* 15, 353–371. doi: 10.1080/13632752.2010.523257
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*, 1st Edn. London: Routledge.
- Hepburn, L., and Beamish, W. (2019). Towards implementation of evidence-based practices for classroom management in Australia: A review of research abstract. *Austral. J. Teach. Educ.* 44, 82–98. doi: 10.14221/ajte.2018v44n2.6
- Hong, Q. N., Pluye, P., Fábregues, S., Bartlett, G., Boardman, F., Cargo, M., et al. (2018). *Mixed methods appraisal tool (MMAT), version 2018. Registration of copyright (# 1148552)*. Gatineau, QC: Canadian Intellectual Property Office.
- Jang, H., Reeve, J., and Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *J. Educ. Psychol.* 102, 588–600. doi: 10.1037/a0019682
- Korpershoek, H., Harms, T., de Boer, H., van Kwijk, M., and Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev. Educ. Res.* 86, 643–680. doi: 10.3102/0034654315626799
- Kunter, M., Baumert, J., and Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learn. Instr.* 17, 494–509. doi: 10.1016/j.learninstruc.2007.09.002
- Kwok, A. (2020). Pre-service teachers' classroom management beliefs and associated teacher characteristics. *Educ. Stud.* 47, 609–626. doi: 10.1080/03055698.2020.1717932
- LaBelle, S., Martin, M. M., and Weber, K. (2013). Instructional dissent in the college classroom: Using the instructional beliefs model as a framework. *Commun. Educ.* 62, 169–190. doi: 10.1080/03634523.2012.759243
- Larribee, B. (2005). *Authentic classroom management: Creating a learning community and building reflective practice*, 3rd Edn. London: Pearson.
- Lewis, R., Romi, S., Qui, X., and Katz, Y. J. (2005). Teachers' classroom discipline and student misbehavior in Australia, China and Israel. *Teach. Teach. Educ.* 21, 729–741. doi: 10.1016/j.tate.2005.05.008
- Madden, L. O. B., and Senior, J. (2017). A proactive and responsive bio-psycho-social approach to managing challenging behaviour in mainstream primary classrooms. *Emot. Behav. Diffic.* 23, 186–202. doi: 10.1080/13632752.2017.1413525
- Miller, A., Ferguson, E., and Byrne, I. (2000). Pupils' causal attributions for difficult classroom behaviour. *Br. J. Educ. Psychol.* 70, 85–96. doi: 10.1348/000709900157985
- Miller, A. N., Katt, J. A., Brown, T., and Sivo, S. A. (2014). The relationship of instructor self-disclosure, nonverbal immediacy, and credibility to student incivility in the college classroom. *Commun. Educ.* 63, 1–16. doi: 10.1080/03634523.2013.835054
- Ming-Tak, H., and Wai-Shing, L. (2008). *Classroom management: Creating a positive learning environment*. Hong Kong: Hong Kong University Press.
- OECD (2019). *TALIS 2018 results (volume I): Teachers and school leaders as lifelong learners*. Paris: TALIS, OECD Publishing. doi: 10.1787/1d0bc92a-en
- Oliver, R. M., Wehby, J. H., and Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive student behavior. *Campbell Syst. Rev.* 7, 1–55. doi: 10.4073/csr.2011.4
- Omoteso, B. A., and Semudara, A. (2011). The relationship between teachers' effectiveness and management of classroom misbehaviours in secondary schools. *Psychology* 2, 902–908. doi: 10.4236/psych.2011.29136
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71
- Paramita, P. P., Sharma, U., Anderson, A., and Laletas, S. (2021). Factors influencing Indonesian teachers' use of proactive classroom management strategies. *Int. J. Incl. Educ.* 1–19. doi: 10.1080/13603116.2021.1916107
- Paramita, P., Anderson, A., and Sharma, U. (2020). Effective teacher professional learning on classroom behaviour management: A review of literature. *Austral. J. Teach. Educ.* 45, 61–81. doi: 10.14221/ajte.2020v45n1.5
- Peterson, R. F., Loveless, S. E., Knapp, T. J., Loveless, B. W., Basta, S. M., and Anderson, S. (1979). The effects of teacher use of I-Messages on student disruptive and study behavior. *Psychol. Rec.* 29, 187–199. doi: 10.1007/bf03394606
- Porter, L. (2014). *A comprehensive guide to classroom management: Facilitating engagement and learning in schools*, 1st Edn. London: Routledge.
- Randolph, J. (2009). A guide to writing the dissertation literature review. *Res. Eval.* 14, 1–13.
- Roache, J., and Lewis, R. R. (2011). Teachers' views on the impact of classroom management on student responsibility. *Austral. J. Educ.* 55, 132–146. doi: 10.1177/000494411105500204
- Rogers, C. (1969). *Freedom to learn*, 1st Edn. Columbus, OH: C. E. Merrill Publishing.

- Romi, S., Lewis, R., and Katz, Y. J. (2009). Student responsibility and classroom discipline in Australia, China, and Israel. *Compare* 39, 439–453. doi: 10.1080/03057920802315916
- Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Future Child* 27, 137–155. doi: 10.1353/foc.2017.0007
- Silverman, D. (2020). *Qualitative research*. New York, NY: SAGE Publications.
- Sniltveit, B., Oliver, S., and Vojtkova, M. (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *J. Dev. Effect* 4, 409–429. doi: 10.1080/19439342.2012.710641
- Steinberg, M. P., and Lacoë, J. (2017). What do we know about school discipline reform? *Educ. Next* 17, 44–52.
- Talvio, M., Lonka, K., Komulainen, E., Kuusela, M., and Lintunen, T. (2014). The development of teachers' responses to challenging situations during interaction training. *Teach. Dev.* 19, 97–115. doi: 10.1080/13664530.2014.979298
- Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., et al. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learn. Instruct.* 22, 431–439. doi: 10.1016/j.learninstruc.2012.04.002
- Vongalis-Macrow, A. (2009). The simplicity of educational reforms: Defining globalization and reframing educational policies during the 1990's. *Int. J. Educ. Pol.* 3, 62–80.
- Wallace, T. L., Sung, H. C., and Williams, J. D. (2014). The defining features of teacher talk within autonomy-supportive classroom management. *Teach. Teach. Educ.* 42, 34–46. doi: 10.1016/j.tate.2014.04.005
- Weger, H. (2017). Instructor active empathic listening and classroom incivility. *Int. J. Listen.* 32, 49–64. doi: 10.1080/10904018.2017.1289091
- Weinstein, C. S., and Romano, M. (2018). *Elementary classroom management: Lessons from research and practice*, 7th Edn. New York, NY: McGraw-Hill Education.
- Worley, D., Titsworth, S., Worley, D. W., and Cornett-DeVito, M. (2007). Instructional communication competence: Lessons learned from award-winning teachers. *Commun. Stud.* 58, 207–222. doi: 10.1080/10510970701341170



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Inês Silva,
Instituto Politécnico de Viana do Castelo,
Portugal
M. Teresa Anguera,
University of Barcelona, Spain
María Pérez Sánchez,
University of A Coruña, Spain

*CORRESPONDENCE

Verónica Violant-Holz
✉ vviolant@ub.edu
Manuel J. Rodríguez
✉ marodriguez@ub.edu

SPECIALTY SECTION

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

RECEIVED 23 February 2023

ACCEPTED 30 March 2023

PUBLISHED 24 April 2023

CITATION

Violant-Holz V, Rodríguez-Silva C and
Rodríguez MJ (2023) Preschool teachers
display a flexible pattern of pedagogical actions
in promoting healthy habits in children.
Front. Psychol. 14:1172460.
doi: 10.3389/fpsyg.2023.1172460

COPYRIGHT

© 2023 Violant-Holz, Rodríguez-Silva and
Rodríguez. 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.

Preschool teachers display a flexible pattern of pedagogical actions in promoting healthy habits in children

Verónica Violant-Holz^{1,2,3*}, Carlota Rodríguez-Silva^{1,2,3} and
Manuel J. Rodríguez^{4,5*}

¹Department of Didactics and Educative Organization, Faculty of Education, Universitat de Barcelona, Barcelona, Spain, ²International Observatory in Hospital Pedagogy, Universitat de Barcelona, Barcelona, Spain, ³Research Group and Innovation in Designs (GRID), Technology and Multimedia and Digital Application to Observational Designs, Thematic Core, Hospital Pedagogy, Barcelona, Spain, ⁴Department of Biomedical Sciences, Institute of Neurosciences, School of Medicine and Health Sciences, Universitat de Barcelona, Barcelona, Spain, ⁵Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

The school represents the optimal setting for promoting the physical, emotional, and social health of children, especially during the first years of life. Understanding the pedagogical actions of teachers to address health education is an important first step in promoting healthy behaviors in children. We inhere analyzed the pedagogical action patterns in the preschool teaching of healthy habits from a holistic health perspective. We used photography as a strategy for data collection and applied a Chi-square automatic interaction detection (CHAID) classification tree, a data mining procedure, to generate a pattern model. We found that the school space and the learning playfulness strategies for the development of executive functions, classified according to the exercise, symbolic, assembly, rules (ESAR) model, were the main factors that influence the pedagogical actions fostering healthy habits. By contrast, the school and the pedagogical resources of the classroom are factors with a much smaller impact on working with healthy habits. This pedagogical action pattern is flexible, since teachers conduct a multiplicity of pedagogical actions through different strategies, in different school spaces, at any time. In conclusion, our results unmask the interdependent relationships between the different factors that determine the teacher's actions at the preschool. It also contributes to the understanding of the teacher's practices in fostering healthy habits in a healthy learning environment.

KEYWORDS

preschool teachers, teaching strategies, health education, healthy school environment, ESAR model, CHAID analysis

1. Introduction

The sustainable development goals of the United Nations proposed education as a crucial factor for improving individual quality of life (World Health Organization, 2003). UNESCO is implementing a global strategy on Education for Health and Well-being (UNESCO, 2022). Early childhood represents a critical period for intervention because healthy lifestyle habits are developed during this period (Goldfield et al., 2012). The school represents the optimal setting to promote physical, emotional, and social health, especially during the first years of life

(Violant-Holz et al., 2021). Consequently, teachers and all the stakeholders involved in education have been requested to systematically consider how educational approaches could increase the healthy behaviors of children (Pulimeno et al., 2020), understanding them as daily healthy routines.

As a consequence, some studies encourage preschool teachers to foster healthy habits within children, such as promoting mental health (Haraldsson et al., 2017; Blewitt et al., 2020), invigorating healthier nutrition and feeding (Whiteside-Mansell and Swindle, 2019), animating physical activity (Tyler et al., 2020), inciting hygiene (Bonnesen et al., 2015), or reassuring emotional health (Ocasio et al., 2015). Recent research encourages preschool teachers to integrate all those health dimensions in their educational approach in classrooms (Violant-Holz et al., 2021), following the emerging concept of health education (Gavidia, 2016). This means understanding education, and therefore also health education, as a complex phenomenon (Morin et al., 2002), integrating physical, mental, and social health with the frailty concept (Boers and Cruz Jentoft, 2015). However, to our knowledge, there are no studies using data mining to analyze the characteristics of the pedagogical actions and the patterns of interaction with children deployed by teachers to address these education-for-health approaches.

Photography is a classical visual method in social science research, including the education field (Deligny, 2006). For example, photography can be used as a method reflecting healthy habits implemented in real-time, in terms of educational spaces in the whole school (Lewallen et al., 2015). Currently, visual methods can be understood as tools for both eliciting data and as a source of data. In the first sense, we can enumerate research using visual methods for eliciting data for educative purposes. This includes cartographies (Padilla-Petry et al., 2021); photo-elicitation (Dworkin et al., 2021), which was also successfully used regarding health concepts with children and adolescents with disabilities, such as health and well-being, physical activity participation and leisure activity (Lumia et al., 2021); photo-interview (Kolb, 2008); or photo-voice (Shaw, 2021), which may also constitute an inclusive didactic strategy (Parrilla et al., 2017).

In the second sense, photographs are used as an empirical tool for recording information in educational research wherever the child is. Thus, photo evidence has effectively been used as a tool to approach healthy eating within families through remote food photography (Bekelman et al., 2019; McCloskey et al., 2019) and proven a feasible methodology to capture parent-packed lunches for preschoolers (Elliott et al., 2021). Other studies applied photo evidence to assess the eating behaviors of scholars. It has been used as a source of data to analyze healthy eating within schools (Bélanger et al., 2016), to compare the nutritional quality of morning snacks (Bucher Della Torre et al., 2021), to measure what students at elementary schools select and consume from school cafeteria meals (Swanson, 2008), and to evaluate what items children chose from the salad bar at schools (Moreno-Black and Stockard, 2018). Photography has also been applied to examine the characteristics of effective drinking water access in schools (Patel et al., 2020).

Beyond nutrition, Kontak et al. (2017) collected photographic data to examine school-based health environments in an enlightening study. They collected photographs to analyze different health dimensions, which include promotion, access, availability of opportunities for healthy eating and physical activity, healthy school

climate and safety, and school accessibility. All in all, photography allows us to see what was happening at a particular point in time (Cleland and MacLeod, 2021), and it has been proven an effective method to capture the reality of the schools in education research, for example by the perspectives of children (Kyrönlampi et al., 2021).

Data mining is an analysis method that extracts information from large data sets, enables the construction of predictive and classification models, and detects interactions within subsets of variables (Collins, 2021). In the education field, several data mining strategies such as classification trees, neural networks, and logistic regression have been implemented to generate classification models of students' profiles and teachers' performance (See Khan and Ghosh, 2021 for a review). The decision tree for classification and prediction is considered as an effective quantitative method for the analysis of condensed and categorized qualitative data (Collins, 2021). The predictive ability of this classification method makes it a popular choice for student performance analysis (Qasrawi et al., 2020; Khan and Ghosh, 2021), and educational quality performance assessment (Cervero, 2020). For example, decision trees have provided sufficient accuracy in predicting education tasks for sports learning (Feu et al., 2019). Information unmasked by education data mining procedures helps researchers to identify new patterns, which are valuable in evaluating pedagogical actions (Rodrigues et al., 2018).

Decision tree algorithms have been used for both the classification and prediction of factors in relationship with children's health. For example, they have been used for the analysis of the underlying causes of child hunger among households accessing food pantries (Wetherill et al., 2021), for the prediction of risk factors for obesity among children aged 24 to 80 months (Lee et al., 2019), and for the identification of success indicators in the implementation in public schools of a physical activity program (Nickelson et al., 2011). Some other studies used decision tree analysis to identify predictors that allow us to implement measures related to psychological health in schools, such as the prevention of homophobic victimization (Fisher and Tanner-Smith, 2016), and bullying (Brewer et al., 2018). This analysis has also been used to detect children's unsafety in families with parental or child developmental problems (van der Put et al., 2016). However, to our knowledge, no studies used data mining analysis to classify the healthy habits-related pedagogical actions and the patterns of interaction with children deployed by preschool teachers.

Our objective and main goal in the present study was to quantitatively define the pattern and trends of the pedagogical actions (understanding patterns as a way of interacting with children in the school) deployed by preschool teachers to promote healthy habits among children. To do that, we performed a novel mixed-method analysis that used photography as a strategy to collect qualitative data and analyzed the patterns of pedagogical actions by using a quantitative classification tree algorithm based on Chi-square automatic interaction detection (CHAID) analysis.

2. Methods

2.1. Design of the study

This study is part of a larger research assessing the impact of a co-created educational project to analyze the health-related preschool

teacher's actions and classroom routines and to improve children's healthy habits (Violant-Holz et al., 2021). In the present study, we used a mixed-methods research design that combined observational methodology (Anguera et al., 2020), and selective methodology (Anguera and Violant, 2015) to collect qualitative data, analyze it, and integrate it into quantitative data (Portell et al., 2015; Anguera et al., 2021; Collins, 2021). We implemented the photography method as a strategy of data collection based on overt indirect nonparticipating naturalist systematic observation, guided by a deductive checklist (Anguera-Argilaga, 1979; Anguera et al., 2020). According to these authors, "Overt" means that we observed the school daily life, being openly and visibly observing; "Indirect" means that we collected photographs; "Nonparticipating" refers to minimizing interaction with participants; "Naturalistic" means that we conducted the observations in the natural environment; and "Systematic" means that researchers entered the field with an observation instrument and that this instrument guided them to perform the observations and take pictures. In this study, photographs function as empirical data collected by the researcher (Poveda, 2018). After qualitative data condensation and categorization, health-related pedagogical actions and routines at preschool were analyzed by data mining procedures.

2.2. Participants

Participants were recruited from four public schools in Barcelona (Spain). For enrollment of teachers, schools were approached in collaboration with the Consorci d'Educació de Barcelona, which is the main educational authority for the city's organization of primary education centers. These four public schools were selected by non-probabilistic convenience sampling, according to their availability and willingness to participate in the study, and their allocation in different city neighborhoods, thus seeking the inclusion of participants from different socioeconomic backgrounds of the city. Teachers were selected by non-probabilistic convenience sampling from those four schools voluntarily. From a total of 21 invited teachers, 16 teachers (12 female and 4 male) and 328 children (3–6-year-old) ultimately participated throughout the entire study.

2.3. Instruments and procedure

For qualitative analysis of the teachers' actions, observations were performed to capture photographs in different school spaces (Anguera et al., 2018), with a deductive checklist (Supplementary Table 1) that implied partial perceptibility. Seven trained observers took the photographs. Two of these observers trained the others in the checklist use. During the actual data collection, a coordinator provided feedback on data quality and assessment criteria (Fernández-Ballesteros, 1992). To analyze the health-related pedagogical actions of teachers, data collection was performed during different periods of the scholar course: in June and October. Observers followed the deductive checklist in the observation and taking photographs (deductive procedure). They also captured actions related to healthy habits that emerged from being at the schools (inductive procedure). These moments were analyzed following the checklist. Teachers and children knew in advance the dates and times when observers would be present to collect data. To minimize their interactions with teachers

or children, observers stood at the back of the classroom during the data collection process. The first session was conducted at each school in June 2019, before the teachers' attendance at our health-promotion project, while the second session was conducted in October 2019, after the teachers attended the project. Photographs were taken from 09.00 am until 1.30 pm.

Four hundred and fifty-two photographs were taken and reviewed to assess if they matched the criteria to be included in the analysis. These inclusion criteria were (1) to be taken in the school spaces: classrooms, corridors, playgrounds, and bathrooms; (2) to be related to the health dimensions: nutrition, physical activity, hygiene, and emotional health, and (3) to be related with pedagogical actions directly or indirectly promoted by teachers. By contrast, the criteria to exclude photographs from the study were: (1) to be repeated photographs, considered such as those photographs that did not provide new information; (2) to be a blurred picture and the actions photographed could not be distinguished in the image, and (3) to show a pedagogical action promoted by non-teacher staff, such as canteen monitors. A final number of two hundred and twenty-two photographs were included in the analysis. Two hundred and twenty photographs were excluded, because either they were repeated, had the same information class, with the same children and class, or were blurred.

2.4. Data classification

Each photograph was categorized into 6 nominal variables: (1) Health Dimension, (2) Educational Resource, (3) ESAR Facets, (4) Spaces, (5) Month, and (6) School, using consensus agreement between two researchers (Arana et al., 2016). As photographs are a multidimensional instrument (Cristancho and Helmich, 2019), can give information about different variables, according to the items analyzed. We controlled the mutual exclusivity criterion (Arnau et al., 1990). These categories included in each nominal variable are explained below.

2.4.1. Health dimension

This nominal variable referred to the health dimensions defined by parents, teachers, and health professionals in three independent co-creation sessions (Violant-Holz et al., 2021). These dimensions defined four categories: (1) *Physical Activity*, (2) *Hygiene*, (3) *Nutrition*, and (4) *Emotional Health*. In concordance with the objectives of the current study, we selected this variable as the dependent one.

2.4.2. Educational resource

This variable referred to the resources that teachers mobilized during their pedagogical actions. Following the criteria published by Torre et al. (2004) on the types of educational resources, this nominal variable had eight categories: (1) *Body*, involvement of any part of the body as a resource; (2) *Experimental materials*, such as experimentation with real, symbolic objects, technological objects and pieces; (3) *Food*, involvement of any food; (4) *Free play*, involvement of non-directed activity, wherever the child is; (5) *Informational materials*, with elements that become vehicles for learning, such as the use of photographs, a book or an image; (6) *Natural environment*, with elements found in nature and environment, such as water; (7) *Social*

environment, with elements involving interpersonal relationships, such as conducting an assembly, group work, and lining up, and (8) *Worksheets*, actions involving papers and pencils.

2.4.3. ESAR facets

This variable referred to the learning playfulness strategies of the Exercise, Symbolic, Assembly, Rules (ESAR) model. This model is based on childhood learning from the perspective of the child's psychological development, which teachers addressed with their pedagogical action (Garón et al., 1996). We selected the ESAR model because it is based on pedagogical principles regarding the use of games and playfulness as learning strategies for the development of both cognitive function and personality at early ages (Piaget, 1989). Every one of the six facets of the ESAR model was considered as a category of this nominal variable: (1) *Affective behavior*, the action involved affective-emotional development in its different levels: confidence, autonomy, initiative, etc. (2) *Cognitive behavior*, the action was represented by cognitive development and sensory-motor and symbolic-intuitive behaviors; (3) *Functional ability*, the action was represented by the expression of cognitive behaviors and basic tools for living in their different forms: discovery, reproduction, efficiency, and combination; (4) *Language ability*, the action was represented by receptive and productive language, oral and written; (5) *Ludic activity*, the action was represented by play and playful behavior; and (6) *Social activities*, the action was represented in social or group context.

2.4.4. Spaces

We defined this variable to assess the putative influence of the school rooms on the teacher's pedagogical actions. This nominal variable included five categories: (1) Classrooms; (2) Corridors; (3) Playgrounds; (4) Bathrooms; and (5) Psychomotricity rooms.

2.4.5. Month

This variable referred to the month when the photographs were taken. It provided information on the effect of the health promotion project on the teacher's pedagogical actions. This nominal variable included two categories. (1) *June*, for photographs taken in June 2019, before the teachers participated in the project, and (2) *October*, for photographs taken in October 2019, after the teacher's participation in the project.

2.4.6. School

This variable referred to the school where the picture was taken. It provided information on the putative differences in pedagogical actions between the centers where the data was collected. As photographs were collected in four different schools, this nominal variable had four categories: (1) School 1; (2) School 2; (3) School 3; and (4) School 4.

By way of clarification, Figure 1 shows the classification of illustrative photographs into the different categories of the Health Dimension, Educational Resource, ESAR Facets, and Spaces variables.

2.5. Data analysis

The variables presenting the strongest impact on the pedagogical actions addressed to foster the children's healthy habits were determined using the data mining method of the CHAID decision

tree, based on the data extracted from photographs. The categorical dependent variable was Health Dimension. Firstly, we performed a descriptive exploratory analysis of every variable that defined the pedagogical actions according to the Health Dimensions. We present the number and percentage of items in each variable category in contingency tables. We then performed an inferential analysis to identify the relationships among the study variables. We performed Pearson's Chi-squared test (χ^2) to analyze the independence between the categorical variables. We assessed the association strength by calculating the Cramer's V coefficient (ϕ_c). In all analyses, we provide the exact value of p and considered value of $p < 0.05$ significant.

Finally, we generated a classification tree as a predictive model to classify and segment the relationships among variables (Collins, 2021). We used the CHAID algorithm to generate the classification tree because most variables were nominal and not binary (Kass, 1980). We used the exhaustive CHAID method, which allows us to analyze with precision all possible results, independently of the categories of the variables (Biggs et al., 1991). We applied the following criteria for the exhaustive CHAID tree growing: statistical test for determining node splitting, Pearson's χ^2 ; significance level $p < 0.05$ with Bonferroni adjustment; maximum tree depth of 5 levels; a minimum of 80 cases in the parental node; a minimum of 40 cases in the filial node; and cross-validation. All analyses were performed with the SPSS Statistics v26 (IBM Corp. USA) statistical package.

3. Results

3.1. Descriptive analysis of the model

The final number of 232 selected photographs provided a total of 267 unities of analyzed data reflecting the pedagogic actions on healthy habits at schools. Comparisons of data revealed differences in frequency proportions between health dimensions ($\chi^2 = 52.44$; d.f. = 3; $p = 0.00001$). Emotional Health presented a higher proportion of items collected (42.32%) than expected for uniformly distributed proportions (25% each), while Hygiene and Physical Activity presented lower proportions than expected (16.10 and 14.61%, respectively). The proportion of items from Nutrition was as expected (26.97%).

We then established the strength of the relationship between the variables Educational Resource, ESAR Facets, Spaces, Month, and School with the Health Dimension variable. With this analysis, we established a rank list of the variables according to their relevance for further analysis in the classification tree. We, thus, examined differences in the Health Dimension distribution among groups of Educational Resources, ESAR Facets, Spaces, Month, and School by the Pearson χ^2 test (Table 1). We found differences in Health Dimension frequency distributions among all the five variables analyzed. The variable with the most statistically significant effect on the Health Dimension frequency distribution was Spaces ($p = 1.7 \times 10^{-25}$), with a moderate degree of association between the variable categories ($\phi_c = 0.314$). Educational Resources also presented a highly significant effect ($p = 2.8 \times 10^{-22}$), with a moderate degree of association between categories ($\phi_c = 0.439$), and the same was true for ESAR Facets ($p = 5.15 \times 10^{-16}$; $\phi_c = 0.366$). We observed a lower significant effect of the variables Month ($p = 0.001$) and School ($p = 0.020$) in the Health Dimension distribution frequencies, both



Categorization table

Photograph	Health dimension	Educational resource	ESAR facets	Spaces
A	Nutrition	Experimental material	Ludic activity	Classroom
B	Emotional health	Social environment	Language ability	Classroom
C	Emotional health	Worksheets	Cognitive behavior	Classroom
D	Physical activity	Free play	Social Activity	Playground
E	Emotional health	Social environment	Affective behavior	Corridor
F	Hygiene	Experimental material	Functional ability	Bathroom

FIGURE 1
 Classification of photographs into the different categories of four variables of the study. (A–F) Illustrative photographs taken in the study; Written informed consent to take photographs at the school was obtained from the individuals participating in this study and /or their representatives; people's faces have been blurred to ensure privacy. The categorization table shows the classification of these six photographs into the categories of the health dimension, educational resource, ESAR Facets, and Spaces variables. See text for details.

with a low degree of association between categories ($\varphi c = 0.262$ for Month, and $\varphi c = 0.155$ for School). Nevertheless, as Health Dimension presented a significant dependence relationship with all five variables (all p -values were lower than 0.05), we included all of them as predictors in the further classification tree analysis.

3.2. Pattern analysis by Chi-square automatic interaction detection

To obtain a model for explaining the patterns of pedagogical actions addressed to foster the children's healthy habits, we performed a classification tree analysis using Health Dimension as the dependent variable and the variables Educational Resource, ESAR Facets, Spaces, Month, and School as predictors. The model generated by CHAID analysis presented a classification tree with 8 nodes, 5 terminal nodes (nodes 2, 3, 5, 6, and 7), and a depth level of 3 (Figure 2).

The factor that best predicted the pedagogical actions and school routines on the four health dimensions was Spaces, which split Node 0 into 2 lines. As defined in terminal Node 2, 72.5% of the pedagogical actions carried out in the playgrounds and psychomotricity rooms were addressed to work Physical Activity, while the other health dimensions were mainly addressed in the classrooms, corridors, and bathrooms (Node 1). The learning facets classified according to the

ESAR system presented the second level of prediction and segmented data from Node 1 into 3 new lines. At this level, pedagogical actions addressed to develop Functional Ability were mainly displayed to work the Hygiene dimension (46.4%; terminal Node 3), while the facets of Social Activity and Affective Behavior, were mainly addressed to promote Emotional Health (83.7%; terminal Node 5). At this level, the ESAR learning facets Language Ability, Cognitive Behavior, and Ludic Activity were mainly addressed to work Emotional Health and Nutrition dimensions (Node 4). Finally, the Month of data collection presented the third level of prediction and segmented data from Node 4 into 2 terminal nodes. In June, the pedagogical actions of teachers mainly addressed the Emotional Health dimension (73.8%; terminal Node 7), whereas these actions mainly addressed the Nutrition dimension in October (53.8%, terminal Node 8). The predictors School and Educational Resource did not significantly contribute to the model (adj. p -value > 0.05 or a number of cases in the node lower than 40) and did not appear in the classification tree.

The generated CHAID model presented a Risk of an estimated 0.363 ± 0.029 based on the estimations of the classification table (Table 2), indicating that 63.7% of the pedagogical actions were correctly classified in the tree. Nutrition is the health dimension with the lowest positive predictive value (a sensitivity of 59.7%) followed by the Hygiene dimension with a predictive value of 60.5% and the

TABLE 1 Cross-tabulation of the study independent variables with the health dimensions.

		Health dimension (%)				Total				
Variable	Domain	Physical activity	Hygiene	Nutrition	Emotional health	<i>n</i>	%	χ^2	<i>p</i> -value	φ_c
Educational resources								154.15	2.8×10^{-22}	0.439
	Worksheets	0 (0.0)	0 (0.0)	2 (2.8)	15 (13.3)	17	6.4			
	Body	5 (12.8)	8 (18.6)	0 (0.0)	12 (10.6)	25	9.4			
	Experimental material	1 (2.6)	19 (44.2)	20 (27.8)	20 (17.7)	60	22.5			
	Food	0 (0.0)	0 (0.0)	14 (19.4)	4 (3.5)	18	6.7			
	Free play	22 (56.4)	0 (0.0)	2 (2.8)	16 (14.2)	40	15.0			
	Informational material	11 (28.2)	10 (23.3)	29 (40.3)	28 (24.8)	78	29.2			
	Natural environment	0 (0.0)	6 (14.0)	4 (5.6)	5 (4.4)	15	5.6			
	Social environment	0 (0.0)	0 (0.0)	1 (1.4)	13 (11.5)	14	5.2			
	Total	39 (100)	43 (100)	72 (100)	113 (100)	267	100			
ESAR facets								107.36	5.15×10^{-16}	0.366
	Affective behavior	1 (2.6)	4 (9.3)	1 (1.4)	28 (24.8)	34	12.7			
	Cognitive behavior	5 (12.8)	3 (7.0)	9 (12.5)	17 (15.0)	34	12.7			
	Functional ability	3 (7.7)	17 (62.8)	18 (25.0)	13 (11.5)	61	22.8			
	Language ability	8 (20.5)	6 (14.0)	24 (33.3)	27 (23.9)	65	24.3			
	Ludic activity	19 (48.7)	2 (4.7)	19 (26.4)	13 (11.5)	53	19.9			
	Social activity	3 (7.7)	1 (1.4)	1 (1.4)	15 (13.3)	20	7.5			
	Total	39 (100)	43 (100)	72 (100)	113 (100)	267	100			
Month								18.33	0.001	0.262
	June	11 (28.2)	19 (44.2)	18 (25.0)	61 (54.0)	109	40.8			
	October	28 (71.8)	24 (55.8)	54 (75.0)	52 (46.0)	158	59.2			
	Total	39 (100)	43 (100)	72 (100)	113 (100)	267	100%			
School								19.34	0.020	0.155
	School 1	5 (12.8)	10 (23.3)	24 (33.3)	18 (15.9)	57	21.3			
	School 2	10 (25.6)	6 (14.0)	13 (18.1)	28 (24.8)	57	21.3			
	School 3	4 (10.3)	10 (23.3)	13 (18.1)	32 (28.3)	59	22.1			
	School 4	20 (51.3)	17 (39.5)	22 (30.6)	35 (31.0)	94	35.2			
	Total	39 (100)	43 (100)	72 (100)	113 (100)	267	100			
Space								147.62	1.7×10^{-25}	0.429
	Bathrooms	0 (0.0)	8 (18.6)	3 (4.2)	5 (4.4)	16	6.0			
	Classrooms	8 (20.5)	28 (65.1)	58 (80.6)	88 (77.9)	182	68.2			
	Corridors	2 (5.1)	4 (9.3)	9 (12.5)	14 (12.4)	29	10.9			
	Playgrounds	24 (61.5)	0 (0.0)	2 (2.8)	5 (4.4)	31	11.6			
	Psychomotricity room	5 (12.8)	3 (7.0)	0 (0.0)	1 (0.9)	9	3.4			
	Total	39 (100)	43 (100)	72 (100)	113 (100)	267	100			

Emotional Health dimension with a predictive value of 63.7%. This is consistent with some of the proportions in the terminal nodes of the classification tree. Thus, according to terminal Node 3,

pedagogical actions addressed to develop Functional Ability in classrooms, corridors, and bathrooms were mainly displayed to work the Hygiene dimension (46.4%), but it should be noted that 32.1% of

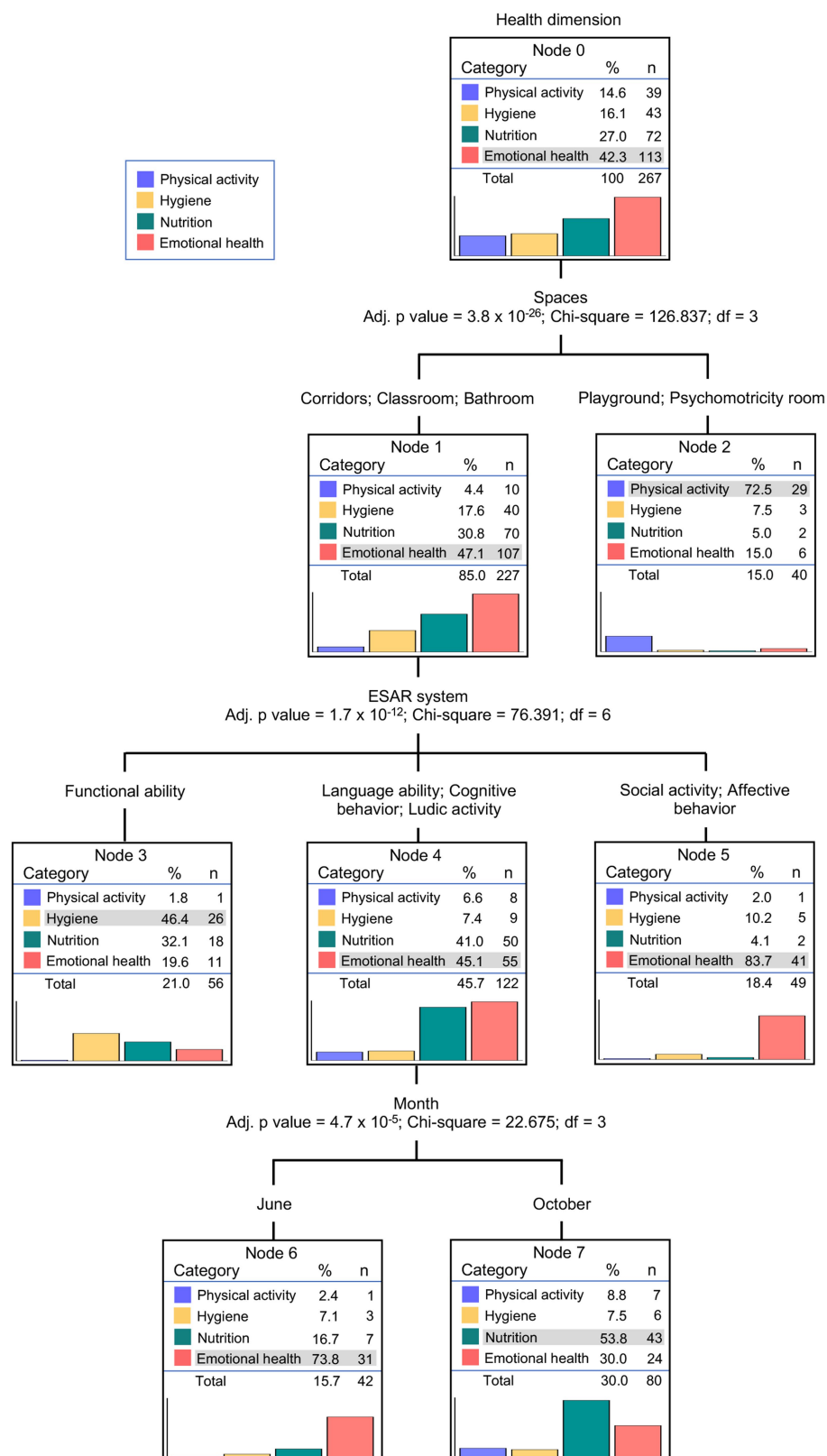


FIGURE 2

Classification tree using the exhaustive Chi-square automatic interaction detection (CHAID) algorithm to predict the patterns of the pedagogical actions of teachers in the promotion of children's healthy habits.

these actions were instead addressed to work the Nutrition dimension. Similarly, terminal Node 8 shows that, in October, learning facets of Language Ability, Cognitive Behavior, and Ludic Activity approached in classrooms, corridors, and bathrooms were mainly addressed to the Nutrition dimension (53.8%), while 30.0% of those pedagogical actions were focused on the Emotional Health dimension.

4. Discussion

In this article, we define for the first time the pattern of pedagogical actions and interaction trends of preschool teachers to foster healthy habits within children. To characterize this interaction pattern, we first used photography as a strategy for data collection and then applied a data mining analysis by using a CHAID classification tree. With this innovative mixed-methods research approach, we found that the school space and the learning playfulness strategies for the development of executive functions, classified according to the ESAR model, were the two main factors influencing the pedagogical actions of preschool teachers fostering healthy habits.

As a strategy of qualitative data collection, images can offer valuable data on the school activities and environment. In our hands, photography provided evidence about how teachers are working, about their activity patterns and interactions with children, and even captured the school environment providing information on whether it is healthy or not. Photographs emerge as a useful tool to observe the school's reality from the educational action point of view (Britsch, 2019). Furthermore, they also have a leading role in documenting the children's voices (Einarsdottir, 2005). Many voices, including teachers, principals, children, and families, can give us the reflection of the health factors generating the environment in the school. But a walk in the classroom can also provide information about the degree of tranquility in the classroom as well as the teacher's skills. Our study emphasizes determining this environment using photographs. It also provides new evidence on photography capturing data about healthy habits being in-real-time implemented, both in terms of all educational spaces in the whole school (Lewallen et al., 2015), and in terms of the emerging concept of health education integrating physical, mental, and social health with the frailty concept (Boers and Cruz Jentoft, 2015; Gavidia, 2016).

In our methodological approach, we have transformed qualitative information from photographs into quantitative data about the teachers' pedagogical actions, suitable to be processed statistically in a CHAID analysis. This crossover mixed-method analysis allowed us

to derive conclusions and meta-inferences unlikely to reach without this type of analysis (Hitchcock and Onwuegbuzie, 2020). To our knowledge, this is the first study implementing this approach to define the interaction trends of preschool teachers addressed to promote healthy habits within children. Under our conditions, we found a significant pattern of teacher preferences when performing their pedagogical actions. The robustness of our classification model relies on both the criteria of the exhaustive CHAID analysis performed (Önder and Uyar, 2017; Feu et al., 2019; Rodríguez-Sabiote et al., 2021) and on the very low Bonferroni-adjusted p -values obtained in the analysis, which are a measure of the compatibility of the entire model of data analysis with the structure of the collected data (Greenland et al., 2016).

School is an excellent setting to promote physical, emotional, and social well-being among students, which also has academic benefits (Pulimeno et al., 2020). Because of that, we have analyzed different educational spaces, such as classrooms, corridors, playgrounds, and bathrooms, following new approaches to health education such as the whole school, whole community, and whole child model (Lewallen et al., 2015). We are struck by how clearly our results point to the weight of spaces of the school in shaping how healthy habits are stimulated, such as those regarding the Physical Activity dimension, which is mostly addressed in playgrounds and psychomotricity rooms. This result illustrates the educational tendency stating that the educator's role during the children's free play is mainly as a "facilitator" and then "observer" but not really implicated in the activity (Kalpogianni, 2019). According to new tendencies in education at the playgrounds, such as active and inclusive playgrounds (Angulo et al., 2022), it would be interesting to promote a more involved preschool teacher in the children's free play. This implies not leaving physical activity only in the playground or in the psychomotricity classroom, but integrating it into the ordinary classroom, since this age represents a critical period to intervene (Baidal and Taveras, 2012; Goldfield et al., 2012). Furthermore, since our study points out that it is beneficial to teach "how to do it," especially at this young age, it would be enriching if teachers become significant role models in healthy habits alongside families (Natale et al., 2014).

These results also highlight the need to continue working on the widely accepted importance of physical activity in relation to health. Specifically, being physical activity an important factor in the obesogenic environment and a determinant of childhood obesity, highlighting its role in promoting health in recreational or physical activity facilities (Mei et al., 2021). Ideally, physical activity needs to be highly promoted within classrooms and other school spaces. Because it not only contributes to the development of gross motor

TABLE 2 CHAID classification table.

Observed	Predicted				
	Physical activity	Hygiene	Nutrition	Emotional health	Percent correct
Physical activity	29	1	7	2	74.4%
Hygiene	3	16	6	8	60.5%
Nutrition	2	18	43	9	59.7%
Emotional health	6	11	24	72	63.7%
Overall percentage	15.0%	21.0%	30.0%	34.1%	63.7%

skills, but also coordination through dance and play, learning to get the body moving, and acquiring healthy routines and habits (Santos-Beneit et al., 2022).

Our results also revealed what learning strategies for the development of executive functions of children are addressed by teachers to promote healthy habits in their educational practice. Furthermore, that pedagogical strategy is the second-best predictor of teacher interventions in the resulting CHAID classification tree. Previous studies report improvement in the ability of executive functions in preschoolers after based intervention (Keown et al., 2020). Thus, the development of executive functions through the promotion and practice in the classroom from early ages both at school and at home is essential. Understanding the way that teachers address these functions is a first step to appropriately defining the educational approaches that best increase children's healthy behaviors. The ESAR model systematically categorizes these executive actions, which facilitates the comprehension of children's development (Garon, 1992). Our results revealed this model as a valid categorization of executive actions when classifying the actions carried out by teachers for health promotion in their educational practice. Specifically, the Hygiene dimension of health was mainly promoted in the classroom by addressing the Functional Ability of children, while Social Activity and Affective Behavior were the functions mobilized to address the Emotional Health dimension. Finally, Language Ability, Cognitive Behavior, and Ludic Activity were addressed by teachers mainly to promote Nutrition and also Emotional Health dimensions.

Related to these different pedagogical approaches we also observed a focus of teachers' actions on addressing Emotional Health, which has a distinct position compared to the other analyzed health dimensions—i.e., Nutrition, Physical Activity, and Hygiene. These results are in line with previous findings showing that the educative community tends to identify health with appropriate Nutrition and considers Emotional health as the most important dimension of health to be developed among preschoolers (Violant-Holz et al., 2021). Taken altogether, our results reflect that teachers address the different health dimensions in separate pedagogical actions in specific school spaces. This points out the pedagogical necessity to develop multidimensional health educational programs (Santos-Beneit et al., 2022), because there is still room for implementing healthy lifestyle habits, and in doing it understanding health promotion as a complex phenomenon (Morin et al., 2002). Aligned with this, we would also like to stress how important is the teacher's education to develop appropriate teaching skills and strategies (Ferreira et al., 2021), and to see themselves as role models and promoters of healthy habits (Ho and Funk, 2018). Furthermore, especially regarding teachers' education, we can also see how there is still room for creativity and innovation regarding the didactics of healthy habits at the school (Gómez-García et al., 2020).

On the contrary, we want to draw attention to the fact that the variables Schools and Pedagogical resources do not appear in our CHAID classification tree. Thus, compared to the other factors analyzed, the school and the pedagogical resources of the classroom are factors with a smaller impact on working with healthy habits. This result may be related to a possible homogeneity of the response in the classroom regardless of the teacher, the school, and even each student. Considering that the

trend is to move toward an inclusive school model with personalized learning (Generalitat de Catalunya, Departament d'Educació, 2015), this result should encourage the educational community to reflect and debate about health pedagogy in the schools of today and tomorrow.

In addition to its efficiency and accuracy in evaluating potential predictor variables (Collins, 2021), the CHAID analysis also provides a Risk estimate that specifies the proportion of cases classified incorrectly (Mardikyan and Badur, 2011). In our study, the relatively elevated Risk estimate may indicate that these patterns of pedagogical actions are flexible. Research indicates that preschool teachers display differentiated intervention levels, adapt their actions depending on the circumstances of the moment, and provide autonomy support reinforcing children's self-determination (Sagastui et al., 2020). Under this point of view, the Risk estimate may quantify the multiplicity of pedagogical actions that teachers conduct to address health education through different strategies, in the different school spaces at any time. This is consistent with the whole school, whole community, and whole child model of approaching educational health (Lewallen et al., 2015). Indeed, in our model, teachers' pedagogical actions addressed the development of Functional Ability in classrooms were mainly displayed to work the Hygiene dimension, but also in no insignificant proportion to work the Nutrition dimension. More interestingly, we observed teachers' interventions on the Emotional Health dimension mostly in classrooms, corridors, and bathrooms by implementing the Social Activity and Affective Behavior facets of the ESAR model. However, we also documented a non-negligible percentage of actions addressed to this health dimension in other school spaces and ESAR facet interventions. This is in accordance with the high importance that preschool teachers give to the promotion of emotional health among preschoolers (Violant-Holz et al., 2021).

To sum up, in our study, the combination of photography as a resource to collect data with the CHAID classification tree as a quantitative statistical analysis unmask the inter-dependence relationships between the different factors that determine the teacher's actions at the school. In our analysis, the school spaces and the learning playfulness strategies for the development of executive functions are the two main factors influencing the pedagogical actions of preschool teachers in fostering healthy habits. Thus, Physical Activity was mainly addressed in the playgrounds and psychomotricity rooms, while it was hardly addressed in regular classrooms. By contrast, we found that the school and the pedagogical resources of the classroom are factors with a much smaller impact on working with healthy habits. These results also teach us that researchers need to be creative in approaching and analyzing what happens in the classroom. Indeed, innovation appears crucial whether we want to improve pedagogical practice based on healthy learning.

4.1. Some limitations and possible future lines of research

This study presents some limitations. First, during data collection, the fact that teachers were warned of the presence of observers in advance may influence their practice. However, this was done

following the ethics standards in naturalistic observation. Second, data collection in this study is based on a single qualitative method. Thus, data transformation into quantitative variables to be analyzed statistically may be biased due to the interpretive subjectivity of researchers. To minimize the impact of this bias, the classification criteria were first defined, and then, data transformation was conducted by consensus agreement, until mutual exclusion was reached (Arana et al., 2016). Finally, even though this is small and local research, it provides significant results. It would be interesting to replicate this research on a larger scale. In this sense, more locations and more centers would be needed to draw more general classification patterns. It would also be interesting to perform more observations and longer registers.

From the methodological point of view, photography brings us the possibility of analyzing the influence of healthy and unhealthy environments on psychological health (Taylor et al., 1997). This can be true not only for the obesogenic environment (Mei et al., 2021) but also, as shown in our study, for more health dimensions, with a holistic analysis of the environment (Sagala and Munawar, 2018). All in all, photography allows us to see what is happening at a particular time point (Cleland and MacLeod, 2021) and has proven to be an effective method to capture the reality of the schools in education research, for example by the children's perspectives (Kyrönlampi et al., 2021). In our hands, photography, as a visual method to reflect teacher strategies in the daily life of preschool, and CHAID analysis, as a tool to classify data and unmask relationships, have been proven valuable methods to define a pattern of teachers' strategies in the daily life of preschool. It will be interesting in the future to combine the use of photography with other observation instruments and data analysis procedures that would complement and consolidate our results.

4.2. Conclusion

The present study reaches a series of conclusions. First, we provide new evidence on the value of photography as a visual method of doing social research. Especially regarding its capacity to show an effective way of capturing the school environment and the promotion of healthy habits in the school's daily life, at the very moment they are being promoted by preschool teachers. Secondly, the ESAR model systematically categorizes the learning playfulness strategies for executive function development, which facilitates the comprehension of children's development. Thirdly, the school space and the learning playfulness strategies for the development of executive functions, classified according to the ESAR model, are the main factors that influence the pedagogical actions fostering healthy habits. This pedagogical action pattern is flexible, since teachers conduct a multiplicity of pedagogical actions through different strategies, in different school spaces, at any time. Finally, promoting healthy habits and routines in schools requires the integration of multiple variables, such as spaces for pedagogical action, the learning strategies promoted by teachers in these spaces, and the moment at which these habits are being promoted. Applying action pattern analyses will help to better define the factors that influence the promotion of such habits among preschoolers.

Data availability statement

The datasets generated and analyzed during the current study are not publicly available due to the need for informed consent for participation and bioethics agreement, but they are available from the corresponding author upon reasonable request.

Ethics statement

Ethical approval of the methods and experimental protocols of this study was granted by the Universitat de Barcelona (Spain). Institutional Review Board approval number-IRB00003099. The study followed the regulations established by the European Union (EU) 2016/679 of the European Parliament and the Council from April 27th on the protection of natural persons with regard to the processing of personal data and free movement of such data, and the Spanish Ley Orgánica 3/2018, from December 5th on the protection of personal data and digital rights. Teachers participated voluntarily in the study and were willing to take photographs and be observed in the classroom in different sessions, at the observer's choice. All participants and/or legal representatives gave their written informed consent to participate in the study in paper format.

Author contributions

VV-H conceived the study. All authors collected the data. VV-H and CR-S anonymized data and performed the qualitative analysis. MR performed the analysis of quantitative data. All authors contributed to the article and approved the submitted version.

Funding

This study is part of all@once and schools@once European research projects (Project numbers: 600137 and 210582, respectively) that received funding from EIT Health. EIT Health is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union that receives support from the European Union's Horizon 2020 Research and Innovation Programme.

Acknowledgments

The authors warmly thank all the marvelous teachers and kids who kindly and selflessly participated in this study. We are very grateful to M. Carol, M. Ellina, and A. Codina (Lifelong Learning Institute-IL3-of the Universitat de Barcelona, Barcelona, Spain), P. Cañal (Hospital Sant Joan de Déu, Barcelona, Spain), and A. Velasco (Consorci d'Educació, Barcelona City Hall, Spain) for their participation in data collection. The authors also acknowledge the support of the Department of Research and Universities of the Generalitat de Catalunya to the Group of Research and Innovation in Designs (GRID) Technology and multimedia and digital application to observational designs (Code: 2021 SGR 00718; VV-H and CR-S) and to the group of

Molecular mechanisms and therapeutical approaches in brain pathologies (Code 2021 SGR 01086; MR). They also would like to thank the International Observatory in Hospital Pedagogy, University of Barcelona, Spain.

Conflict of interest

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

The reviewer MA declared a shared affiliation with the authors to the handling editor at the time of review.

References

- Anguera, M. T., Blanco-Villaseñor, A., Losada, J. L., and Sánchez-Algarra, P. (2020). Integración de elementos cualitativos y cuantitativos en metodología observacional. *Ámbitos. Revista Internacional de Comunicación* 49, 49–70. doi: 10.12795/ambitos.2020.149.04
- Anguera, M. T., Portell, M., Chacón-Moscoso, S., and Sanduvete-Chaves, S. (2018). Indirect observation in everyday contexts: concepts and methodological guidelines within a mixed methods framework. *Front. Psychol.* 9, 1–20. doi: 10.3389/fpsyg.2018.00013
- Anguera, M. T., Portell, M., Hernández-Mendo, A., Sánchez-Algarra, P., and Jonsson, G. K. (2021). “Diachronic analysis of qualitative data” in *The Routledge Reviewer's Guide to Mixed Methods Analysis*. eds. R. B. J. Anthony and J. Onwuegbuzie (New York: CRC Press).
- Anguera, M. T., and Violant, V. (2015). *Investigación en Pedagogía Hospitalaria*. Málaga: Aljibe.
- Anguera-Argilaga, T. (1979). Observational typology. *Qual. Quant.* 13, 449–484. doi: 10.1007/BF00222999
- Angulo, I., Vizcarra, M. T., Gamito, R., and López-Vélez, A. L. (2022). ¿Cómo hacer más coeducativos los patios de recreo? Una propuesta de aprendizaje-servicio. *RIDAS. Revista Iberoamericana de Aprendizaje-Servicio* 13, 60–79. doi: 10.1344/ridas2022.13.4
- Arana, J., Lapresa, D., Anguera, M. T., and Garzón, B. (2016). Ad hoc procedure for optimising agreement between observational records. *Anales de Psicología* 32:589. doi: 10.6018/analesps.32.2.213551
- Arnau, J., Anguera, M. T., and Gómez, J. (1990). *Metodología de la Investigación en Ciencias del Comportamiento*. Murcia: Universidad de Murcia.
- Baidal, J. A. W., and Taveras, E. M. (2012). Childhood obesity: Shifting to focus to early prevention. *Arch. Pediatr. Adolesc. Med.* 166, 1179–1181. doi: 10.1001/2013
- Bekelman, T. A., Bellows, L. L., McCloskey, M. L., Martin, C. K., and Johnson, S. L. (2019). Assessing dinner meals offered at home among preschoolers from low-income families with the remote food photography method. *Pediatr. Obes.* 14, 49–58. doi: 10.1111/ijpo.12558
- Bélanger, M., Humbert, L., Vatanparast, H., Ward, S., Muhajarine, N., Chow, A. F., et al. (2016). A multilevel intervention to increase physical activity and improve healthy eating and physical literacy among young children (ages 3–5) attending early childcare centres: the healthy start-Départ Santé cluster randomised controlled trial study protocol. *BMC Public Health* 16, 1–10. doi: 10.1186/s12889-016-2973-5
- Biggs, D., De Ville, B., and Suen, E. (1991). A method of choosing multiway partitions for classification and decision trees. *J. Appl. Stat.* 18, 49–62. doi: 10.1080/02664769100000005
- Blewitt, C., Morris, H., Jackson, K., Barrett, H., Bergmeier, H., O'Connor, A., et al. (2020). Integrating health and educational perspectives to promote preschoolers' social and emotional learning: development of a multi-faceted program using an intervention mapping approach. *Int. J. Environ. Res. Public Health* 17, 1–24. doi: 10.3390/ijerph17020575
- Boers, M., and Cruz Jentoft, A. J. (2015). A new concept of health can improve the definition of frailty. *Calcif. Tissue Int.* 97, 429–431. doi: 10.1007/s00223-015-0038-x
- Bonnesen, C. T., Plauborg, R., Denbæk, A. M., Due, P., and Johansen, A. (2015). Process evaluation of a multi-component intervention to reduce infectious diseases and improve hygiene and well-being among school children: the hi five study. *Health Educ. Res.* 30, 497–512. doi: 10.1093/her/cyv019
- Brewer, S. L., Brewer, H. J., and Kulik, K. S. (2018). Bullying victimization in schools: why the whole school, whole community, whole child model is essential. *Int. J. School Health* 88, 794–802. doi: 10.1111/josh.12686
- Britsch, S. (2019). Exploring science visually: science and photography with pre-kindergarten children. *J. Early Child. Lit.* 19, 55–81. doi: 10.1177/1468798417700704
- Bucher Della Torre, S., Fabbri, S., and Carrard, I. (2021). Healthy snacks in school: how do regulations work? A mixed-design study. *J. Sch. Health* 91, 697–705. doi: 10.1111/josh.13061
- Cervero, A. (2020). Evaluation of educational quality performance on virtual campuses using fuzzy inference systems. *PLoS One* 15, e0232802–e0232816. doi: 10.1371/journal.pone.0232802
- Cleland, J., and MacLeod, A. (2021). The visual vernacular: embracing photographs in research. *Persp. Med. Educ.* 10, 230–237. doi: 10.1007/s40037-021-00672-x
- Collins, K. (2021). “Chi-square automatic interaction detection analysis of qualitative data” in *The Routledge Reviewer's guide to mixed methods analysis*. eds. A. J. Onwuegbuzie and R. B. Johnson (New York: CRC Press).
- Cristancho, S. M., and Helmich, E. (2019). Rich pictures: a companion method for qualitative research in medical education. *Med. Educ.* 53, 916–924. doi: 10.1111/medu.13890
- Generalitat de Catalunya, Departament d'Educació. (2015). De l'escola inclusiva al sistema inclusiu. Una escola per a tothom, un projecte per a cadascú. Generalitat de Catalunya. https://repositori.educacio.gencat.cat/bitstream/handle/20.500.12694/246/de_l_escola_inclusiva_al_sistema_inclusiu_una_escola_per_a_tothom_un_projecte_per_a_cadascu_2015?sequence=1&isAllowed=y
- Deligny, F. (2006). *Permitir, Trazar, Ver*. Barcelona: MACBA.
- Dworkin, M., Akintayo, T., Calem, D., Doran, C., Guth, A., Kamami, E. M., et al. (2021). Life during the pandemic: an international photo-elicitation study with medical students. *BMC Med. Educ.* 21:244. doi: 10.1186/s12909-021-02684-x
- Einarsdottir, J. (2005). Playschool in pictures: Children's photographs as a research method. *Early Child Dev. Care* 175, 523–541. doi: 10.1080/03004430500131320
- Elliott, S., McCloskey, M. L., Johnson, S. L., Mena, N. Z., Swindle, T., and Bellows, L. L. (2021). Food photography as a tool to assess type, quantity, and quality of foods in parent-packed lunches for preschoolers. *J. Nutr. Educ. Behav.* 53, 164–173. doi: 10.1016/j.jneb.2020.10.007
- Fernández-Ballesteros, R. (1992). *Introducción a La Evaluación Psicológica*. Madrid: Pirámide.
- Ferreira, M., Reis-Jorge, J., and Batalha, S. (2021). Social and emotional learning in preschool education-a qualitative study with preschool teachers. *Int. J. Emot. Educ.* 13, 51–66.
- Feu, S., García-Rubio, J., Gamero, M. D. G., and Ibáñez, S. J. (2019). Task planning for sports learning by physical education teachers in the pre-service phase. *PLoS One* 14:e0212833. doi: 10.1371/journal.pone.0212833
- Fisher, B. W., and Tanner-Smith, E. E. (2016). Examining school security measures as moderators of the association between homophobic victimization and school avoidance. *J. Sch. Violence* 15, 234–257. doi: 10.1080/15388220.2014.983644
- Garon, D. (1992). *Classificação e análise de materiais lúdicos: o sistema ESAR*. São Paulo: O Direito de Brincar: a Brinquedoteca.
- Garón, D., Filion, R., and Doucet, M. (1996). *El sistema ESAR: Un método de análisis psicológico de los juguetes*. Ibi (Alicante): AIJU, Instituto Tecnológico del Juguete.
- Gavidia, V. (2016). *Los ocho ámbitos de la Educación para la Salud en la Escuela*. Valencia: Tirant lo Blanch. <https://www.uv.es/comsal/pdf/librocomsal.pdf>
- Goldfield, G. S., Harvey, A., Grattan, K., and Adamo, K. B. (2012). Physical activity promotion in the preschool years: a critical period to intervene. *Int. J. Environ. Res. Public Health* 9, 1326–1342. doi: 10.3390/ijerph9041326

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1172460/full#supplementary-material>

- Gómez-García, G., Marín-Marín, J. A., Romero-Rodríguez, J.-M., Ramos Navas-Parejo, M., and Rodríguez Jiménez, C. (2020). Effect of the flipped classroom and gamification methods in the development of a didactic unit on healthy habits and diet in primary education. *Nutrients* 12, 12. doi: 10.3390/nu12082210
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., et al. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur. J. Epidemiol.* 31, 337–350. doi: 10.1007/s10654-016-0149-3
- Haraldsson, K., Isaksson, P., and Eriksson, M. (2017). 'Happy when they arrive, happy when they go home'—focusing on promoting children's mental health creates a sense of trust at preschools. *Early Years* 37, 386–399. doi: 10.1080/09575146.2016.1191442
- Hitchcock, J. H., and Onwuegbuzie, A. J. (2020). Developing mixed methods crossover analysis approaches. *J. Mixed Methods Res.* 14, 63–83. doi: 10.1177/1558689819841782
- Ho, J., and Funk, S. (2018). Preschool: promoting young Children's social and emotional health. *YC Young Child.* 73, 73–79.
- Kalpagianni, D. E. (2019). Why are the children not outdoors? Factors supporting and hindering outdoor play in Greek public day-care centres. *Int. J. Play* 8, 155–173. doi: 10.1080/21594937.2019.1643979
- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Appl. Stat.* 29:119. doi: 10.2307/2986296
- Keown, L. J., Franke, N., and Triggs, C. M. (2020). An evaluation of a classroom-based intervention to improve executive functions in 4-year old children in New Zealand. *Early Childhood Educ. J.* 48, 621–631. doi: 10.1007/s10643-020-01023-x
- Khan, A., and Ghosh, S. K. (2021). *Student performance analysis and prediction in classroom learning: A review of educational data*, vol. 26 Education and Information Technologies, 205–240 doi: 10.1007/s10639-020-10230-3
- Kolb, B. (2008). Involving, sharing, Analysing—potential of the participatory photo interview. *Forum: Qual. Soc. Res.* 9, 1–25. doi: 10.17169/fqs-9.3.1155
- Kontak, J. C. H., McIsaac, J. L. D., Penney, T. L., Kuhle, S., and Kirk, S. F. L. (2017). The picture of health: examining school-based health environments through photographs. *Health Promot. Int.* 32, 322–330. doi: 10.1093/heapro/daw027
- Kyrönlampi, T., Uitto, M., and Puroila, A. M. (2021). Place, peers, and play: Children's belonging in a preprimary school setting. *Int. J. Early Childhood* 53, 65–82. doi: 10.1007/s13158-021-00285-9
- Lee, I., Bang, K. S., Moon, H., and Kim, J. (2019). Risk factors for obesity among children aged 24 to 80 months in Korea: a decision tree analysis. *J. Pediatr. Nurs.* 46, e15–e23. doi: 10.1016/j.pedn.2019.02.004
- Lewallen, T. C., Hunt, H., Potts-Datema, W., Zaza, S., and Giles, W. (2015). The whole school, whole community, whole child model: a new approach for improving educational attainment and healthy development for students. *J. Sch. Health* 85, 729–739. doi: 10.1111/josh.12310
- Lumia, C. M., Arbour-Nicitopoulos, K. P., Moola, F. J., and McPherson, A. C. (2021). Using photo-elicitation to explore health promotion concepts with children and adolescents with disabilities: a rapid scoping review. *Disabil. Rehabil.* 44, 5708–5718. doi: 10.1080/09638288.2021.1945690
- Mardikyan, S., and Badur, B. (2011). Analyzing teaching performance of instructors using data mining techniques. *Inf. Educ.* 10, 245–257. doi: 10.15388/infedu.2011.17
- McCloskey, M. L., Johnson, S. L., Bekelman, T. A., Martin, C. K., and Bellows, L. L. (2019). Beyond nutrient intake: use of digital food photography methodology to examine family dinnertime. *J. Nutr. Educ. Behav.* 51, 547–555.e1. doi: 10.1016/j.jneb.2019.01.020
- Mei, K., Huang, H., Xia, F., Hong, A., Chen, X., Zhang, C., et al. (2021). State-of-the-art of measures of the obesogenic environment for children. *Obes. Rev.* 22:e13093. doi: 10.1111/obr.13093
- Moreno-Black, G., and Stockard, J. (2018). Salad bar selection patterns of elementary school children. *Appetite* 120, 136–144. doi: 10.1016/j.appet.2017.08.034
- Morin, E., Roger, E., and Motta, R. (2002). *Educación en la era Planetaria: El Pensamiento Complejo Como Método de Aprendizaje en el Error y la Incertidumbre Humana*. Quito: Grupo Santillana 2005. UNESCO.
- Natale, R. A., Messiah, S. E., Asfour, L., Uhlhorn, S. B., Delamater, A., and Arheart, K. L. (2014). Role modeling as an early childhood obesity prevention strategy: effect of parents and teachers on preschool children's healthy lifestyle habits. *J. Dev. Behav. Pediatr.* 35, 378–387. doi: 10.1097/DBP.0000000000000074
- Nickelson, J., Alfonso, M. L., McDermott, R. J., Bumpus, E. C., Bryant, C. A., and Baldwin, J. A. (2011). Characteristics of "tween" participants and non-participants in the VERB™ summer scorecard physical activity promotion program. *Health Educ. Res.* 26, 225–238. doi: 10.1093/her/cyq089
- Ocasio, K., Van Alst, D., Koivunen, J., Huang, C. C., and Allegra, C. (2015). Promoting preschool mental health: results of a 3 year primary prevention strategy. *J. Child Fam. Stud.* 24, 1800–1808. doi: 10.1007/s10826-014-9983-7
- Önder, E., and Uyar, Ş. (2017). CHAID analysis to determine socioeconomic variables that explain students' academic success. *Univ. J. Educ. Res.* 5, 608–619. doi: 10.13189/ujer.2017.050410
- Padilla-Petry, P., Hernández-Hernández, F., and Sánchez-Valero, J. A. (2021). Using cartographies to map time and space in teacher learning in and outside school. *Int. J. Qual. Methods* 20, 160940692199290–160940692199211. doi: 10.1177/1609406921992906
- Parrilla, Á., Raposo-Rivas, M., Martínez-Figueira, E., and Doval, M. I. (2017). Materiales didácticos para todos: el carácter inclusivo de fotovoz. *Educación. Siglo XXI* 35:17. doi: 10.6018/j/308881
- Patel, A. I., Podrabsky, M., Hecht, A. A., Morris, S., Yovanovich, S., Walkinshaw, L. P., et al. (2020). Development and validation of a photo-evidence tool to examine characteristics of effective drinking water access in schools. *J. Sch. Health* 90, 271–277. doi: 10.1111/josh.12873
- Piaget, J. (1989). *Le langage et la pensée chez l'enfant: études sur la logique de l'enfant: 10ème éd.* Paris: Delachaux et Niestlé.
- Portell, M., Anguera, M. T., Chacón-Moscoso, S., and Sanduvete-Chaves, S. (2015). Guidelines for reporting evaluations based on observational methodology. *Psicothema* 27, 283–289. doi: 10.7334/psicothema2014.276
- Poveda, D. (2018). Photographs as a research tool in child studies: some analytical metaphors and choices. *Qual. Res. Educ.* 7, 170–196. doi: 10.17583/qre.2018.3350
- Pulimeno, M., Piscitelli, P., Miani, A., Colazzo, S., Mazza, A., and Colao, A. (2020). Narrative pedagogy to promote health and wellbeing in school setting: an approach proposed by UNESCO chair on health education and sustainable development. *Health Promot. Perspect.* 10, 1–2. doi: 10.15171/hpp.2020.01
- Qasrawi, R., Abdeen, Z., Taweel, H., Younis, M. A., and Al-halawa, D. A. (2020). Data mining techniques in identifying factors associated with schoolchildren science and arts academic achievement. *IEEE* 78–83. doi: 10.1109/ICPET51420.2020.00023
- Rodrigues, M. W., Zárate, L. E., and Isotani, S. (2018). Educational data mining: a review of evaluation process in the e-learning. *Telematics Inform.* 35, 1701–1717. doi: 10.1016/j.tele.2018.04.015
- Rodríguez-Sabiote, C., Álvarez-Rodríguez, J., Álvarez-Ferrandiz, D., and Zurita-Ortega, F. (2021). Using chi-squared automatic interaction detection modelling to identify student opinion profiles regarding same-sex couples as a family structure. *Heliyon* 7:e06469. doi: 10.1016/j.heliyon.2021.e06469
- Sagala, A. C. D., and Munawar, M. "Analysis of holistic environmental approach to health, nutrition, and safety in early childhood education." in *Proceedings of the International Conference of Early Childhood Education (ICECE 2017)*, 169 (Icece 2017). (2018). Atlantis Press, 246–248.
- Sagastui, J., Herrán, E., and Anguera, M. T. (2020). A systematic observation of early childhood educators accompanying young Children's free play at Emmi Pikler nursery school: instrumental behaviors and their relational value. *Front. Psychol.* 11, 1–12. doi: 10.3389/fpsyg.2020.01731
- Santos-Beneit, G., Fernández-Jiménez, R., de Cos-Gandoy, A., Rodríguez, C., Carral, V., Bodega, P., et al. (2022). Lessons learned from 10 years of preschool intervention for health promotion: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* 79, 283–298. doi: 10.1016/j.jacc.2021.10.046
- Shaw, P. A. (2021). Photo-elicitation and photo-voice: using visual methodological tools to engage with younger children's voices about inclusion in education. *Int. J. Res. Method Educ.* 44, 337–351. doi: 10.1080/1743727X.2020.1755248
- Swanson, M. (2008). Digital photography as a tool to measure school cafeteria consumption. *J. Sch. Health* 78, 432–437. doi: 10.1111/j.1746-1561.2008.00326.x
- Taylor, S. E., Repetti, R. L., and Seeman, T. (1997). Health Psychology: what is an unhealthy environment and how does it get under the skin? *Annu. Rev. Psychol.* 48, 411–447. doi: 10.1146/annurev.psych.48.1.411
- Torre, S., Oliver, C., Violant, V., and Tejada, J. (2004). Metodología de estudio de casos y perfil de estrategias docentes. *Contextos Educativos* 65–86. doi: 10.18172/con.529
- Tyler, E. C., Brazendale, K., Hunt, E., Rafferty, A., Beets, M. W., and Weaver, R. G. (2020). Physical activity opportunities of low-income elementary school-aged children during the segmented school day. *J. Sch. Health* 90, 787–793. doi: 10.1111/josh.12939
- UNESCO (2022). Unesco strategy on education for health and well-being: education for health. UNESCO, 8–12.
- van der Put, C. E., Hermanns, J., van Rijn-van Gelderen, L., and Sondeijker, F. (2016). Detection of unsafety in families with parental and/or child developmental problems at the start of family support. *BMC Psychiatry* 16:15. doi: 10.1186/s12888-016-0715-y
- Violant-Holz, V., Rodríguez-Silva, C., Carol, M., and Rodríguez, M. J. (2021). Impact of cocreation training capsules for preschool teachers on children's healthy habits: a pilot study conducted in Barcelona, Spain. *BMC Public Health* 21, 1–13. doi: 10.1186/s12889-021-12160-2
- Wetherill, M. S., Hartwell, M. L., Williams, M. B., White, K. C., Harriat, A. W., Proffitt, S., et al. (2021). Beyond groceries: an analysis of referral needs to address underlying causes of child hunger among households accessing food pantries. *Soc. Work Public Health* 36, 732–748. doi: 10.1080/19371918.2021.1943099
- Whiteside-Mansell, L., and Swindle, T. M. (2019). Evaluation of together we inspire smart eating: pre-school fruit and vegetable consumption. *Health Educ. Res.* 34, 62–71. doi: 10.1093/her/cyy048
- World Health Organization (2003). Skills for health: skills-based health education including life skills: an important component of a child-friendly/health-promoting school. *World Health Org.* 12:39. doi: 10.7748/nm.12.7.39.s23



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Cristina Honrubia,
Universidad Isabel I de Castilla, Spain
María Martínez López,
University of Castilla La Mancha, Spain

*CORRESPONDENCE

Omar Abdullah Alsamani
✉ o.alsamani@uoh.edu.sa

RECEIVED 23 January 2023

ACCEPTED 14 April 2023

PUBLISHED 02 May 2023

CITATION

Alsamani OA, Alsamiri YA and Alfaidi SD (2023)
Elementary school teachers' perceptions of the
characteristics of twice-exceptional students.
Front. Educ. 8:1150274.
doi: 10.3389/feduc.2023.1150274

COPYRIGHT

© 2023 Alsamani, Alsamiri and Alfaidi. 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.

Elementary school teachers' perceptions of the characteristics of twice-exceptional students

Omar Abdullah Alsamani^{1*}, Yasir Ayed Alsamiri¹ and
Seham Defallah Alfaidi²

¹Department of Special Education, Hai'il University, Ha'il, Saudi Arabia, ²Department of Psychology, Hafar Al Batin University, Hafar Al Batin, Saudi Arabia

The current study explored the perceptions and understanding of Saudi elementary school teachers regarding the characteristics of twice-exceptional (2e) students. This qualitative phenomenological study recruited 14 elementary school teachers at schools in Medina City in Saudi Arabia. Data were gathered through in-depth semi-structured interviews to gain insight into participants' perceptions of this student population. Findings indicated that teachers associate certain characteristics with the giftedness trait of 2e students, including excellent communication skills, problem-solving ability, strong retention, curiosity, creativity, abstract thinking, and leadership skills. Moreover, Saudi elementary school teachers perceive that 2e students experience academic challenges related to their disabilities in reading, writing, and mathematics, as well as are aware that such students tend to struggle with low self-esteem and maintaining concentration. Overall, there is a tendency to view these students as having two independent conditions—giftedness and disabilities that cause academic challenges—rather than identifying these individuals as having a single condition of being twice exceptional that comprises traits that impact their learning in combination.

KEYWORDS

twice-exceptional students, gifted students with learning disabilities, elementary teachers' perceptions, gifted students characteristics, gifted education, Saudi Arabia

1. Introduction

Although gifted students with disabilities, commonly known as twice-exceptional (2e) learners, have diverse characteristics and abilities, they are often overlooked in identification for support services. While these students have specific talents—higher-level intellectual abilities, superior vocabulary, exceptional comprehension of abstract ideas and concepts, and high levels of creativity—at the same time, they tend to exhibit poor reading and writing skills, lack organizational and study skills, and have low self-esteem (Buică-Belciu and Popovici, 2014; Lee and Ritchotte, 2019). This category of students, ones whose giftedness is accompanied by another condition that requires special education support, includes those with such conditions as autism spectrum disorder (ASD), attention-deficit hyperactivity disorder (ADHD), and learning disabilities like dysgraphia, dyslexia, and dyscalculia (Levi, 2019; Weiss, 2022).

Ironically, school system processes often identify 2e students as having learning or other types of disabilities but regularly overlook their giftedness, due to the challenges related to conceptualizing students that might possess both traits (Hays, 2016; Wellisch, 2016). Moreover, teacher identification for gifted programs is often the elementary method of identification in

the early years because it is seen as difficult to test for giftedness in the K-3 years. Therefore, if teachers do not have a comprehensive perception of what giftedness comprises, including that it might involve special education needs, they might not be able to identify children who are twice exceptional and may face substantial challenges in understanding, identifying, and teaching these students (Chimhenga, 2016; Amran and Majid, 2019; Cody et al., 2022).

1.1. Importance of teacher identification of the 2e

The literature highlights the significant role of elementary school teachers in the identification and education of students with both giftedness and learning disabilities (Alsamirani and Aljohani, 2018; Firat and Bildiren, 2022). In this regard, teachers' awareness of the characteristics of students with this combination of traits facilitates early identification and intervention for these students (Firat and Bildiren, 2022). Yet there is limited research on the perceptions of teachers during the elementary school period when the identification and education of students with giftedness and learning disabilities are of critical importance (Akinbote et al., 2017; Alsamirani, 2019). If teachers understand that students can be simultaneously gifted and have special education needs, they are better equipped to identify 2e students, recognize their unique learning needs, and modify teaching styles and curricula to accommodate them (Alsamirani and Aljohani, 2018; Maddocks, 2020). Teachers need comprehensive knowledge of the general characteristics of these students to identify them and provide interventions when necessary (Firat and Bildiren, 2022).

One of the prominent characteristics of 2e students that has been discussed in past research is the discrepancy between their strengths and their weaknesses, characterized by their high potential and low attainment (Ritchotte and Zaghlawan, 2019; Mohammad, 2020). According to Hopwood (2019), most 2e students fall through the cracks of the education system, due to the lack of professional development programs and university preparation available to teachers. In addition, the high ability of 2e students often masks their special education needs in the early years and so interventions and accommodations are not implemented; this can result in these children facing significant and prominent challenges at the middle and high school levels (Baldwin et al., 2015; Lee and Ritchotte, 2019; Ritchotte and Zaghlawan, 2019; Mohammad, 2020).

1.2. Training and experience of teaching 2e

Elementary school teachers in Saudi Arabia are inadequately prepared to recognize and support 2e students (Alsamirani, 2018). However, many of those teachers are willing to help 2e students when relevant forms of training and support are offered. 2e students in Saudi Arabia are placed in general education, where they spend their time in a regular classroom with a general education teacher. Training about teaching and understanding 2e students is not officially provided for general education teachers in pre-services teachers preparations (Alsamirani, 2018). However, there are regular education teachers who are voluntarily involved in some in-services courses related to gifted education and special education where they developed a better understanding of the characteristics of those 2e students. In regard to the

lack of training of general education teachers, Alsamirani (2019) suggested that elementary school teachers should be supported by gifted and special education educators to better recognize and support 2e students.

Studies have shown that the teachers' professional development regarding twice-exceptionality can significantly impact their 2e students' educational progress (Cody et al., 2022). For example, Coolahan (2004) and Chessman (2005) studied how receiving postgraduate training in gifted education can positively affect the understanding of 2e students' needs that is reflected in providing better education experiences for those students. Research also demonstrated that experienced teachers have a deeper understanding of the needs of gifted students including 2e students (Alsamirani, 2019). Teachers' perspectives of 2e students are associated with their experiences and training in gifted and special education (Ramsay-Cohen, 2016). Teachers who have experience in teaching 2e students are less likely to be biased against them compared to teachers without such expertise (Alkhunaini, 2013; Ramsay-Cohen, 2016; Alsamirani, 2018). Many inexperienced teachers are reluctant to refer 2e students to special education regarding their learning disabilities as they would not expect that a gifted child may have a learning disability (Wormald et al., 2015). Bianco and Leech (2010) argue that inexperienced teachers and inadequate teacher training in twice-exceptionality interfere with effectively determine 2e students and then refer them to support programs (Shaunessy, 2007). Generally, only a small proportion of general education teachers have the capability and are prepared to recognize and support 2e students (Alkhunaini, 2013).

1.3. Characteristics of 2e students

Educational researchers have noted certain consistent traits in this population, including poor organizational skills, lack of focus in class, difficulty with basic spelling/reading skills, poor peer relationships, and low self-esteem—all of which may result in the individual experiencing academic failure or demonstrating a dislike of school. At the same time, studies have found these students may have exceptional vocabulary, analytical/problem-solving and comprehension skills, and show extraordinary interest or talent in a particular area (Coleman and Gallagher, 2015; Roberts et al., 2015; Mayes and Moore, 2016; Wright and Ford, 2017).

Gifted students who also have learning disabilities may face academic challenges that directly affect their learning journey. They might have difficulty handling activities that require sequencing and memorization, and some might have spelling, phonics, and computation difficulties (Barnard-Brak et al., 2015; Lee and Ritchotte, 2019; Maddocks, 2020). A significant number of 2e students have poor organizational and handwriting skills (Firat and Bildiren, 2022). Additionally, they often perform poorly on timed tests and may fail to achieve learning objectives and experience difficulties engaging in activities that require focus—including reading and writing (Wormald et al., 2015).

1.4. Twice-exceptional students in the Saudi context

There are only a few studies regarding Saudi elementary school teachers' awareness of this student population and their characteristics

(Weiss, 2022). The understanding and perceptions of elementary school teachers regarding these children becomes critical when considering the importance of early intervention for both the gifted and those with disabilities that affect academic achievement (Wellisch, 2016; Alsamirani and Aljohani, 2018). Early identification of 2e students is integral to successful intervention; thus, elementary school teachers should possess good understanding and accurate perceptions of them (Firat and Bildiren, 2022). This situation grows even more complex when these teachers' lack of knowledge about gifted students is coupled with the difficulties of identifying learning disabilities at the K-6 level (Wellisch, 2016).

1.5. Purpose of the study

No in-depth analyses have been conducted in Saudi Arabia on the perceptions of elementary school teachers regarding gifted students with learning and other disabilities that affect educational achievement. For this reason, and those previously described as barriers for this population of students, this research was designed to explore the perceptions and understanding of Saudi elementary school teachers regarding the characteristics of gifted students with learning disabilities. Understanding how these teachers perceive and conceptualize these students is a wise addition to the literature. In this study, we endeavored to develop a clearer picture of the views of this population of teachers and to identify the gaps in their understanding of the attributes and challenges of these students. Such an investigation is critical as the delay in understanding and identifying 2e students may lead to their failure to achieve to their academic potential.

Examining elementary school teachers' understanding and perceptions of 2e students and their characteristics is vital to better supporting these children. We therefore engaged in this qualitative study to deeply explore participants' perceptions and understanding of the twice exceptional through an interview designed to obtain data on the teachers' perceptions of both the strengths and weaknesses of these children, by first establishing the following research questions.

RQ1. How do Saudi Arabian elementary school teachers perceive the gifted characteristics of twice-exceptional students?

RQ2. How do Saudi Arabian elementary school teachers perceive the academic challenges twice-exceptional students experience?

2. Research design

Qualitative research is considered the most appropriate design when examining a social or human problem, as it allows the researcher to create a multifaceted picture of an understudied phenomenon (Creswell and Poth, 2018). Qualitative design was employed for this research since the goal was to acquire a deep understanding of elementary school teachers' perceptions of 2e students (Merriam and Tisdell, 2016).

2.1. Participants

After receiving approval to conduct the study, we first contacted elementary schools in Medina city via email explaining our study and the inclusion criteria for recruiting the participants. The first inclusion

criterion was that only certified current elementary teachers from Medina city are invited to participate in the study. Second, as there is only a small portion of teachers who have sufficient understanding of 2e students, we invited only elementary teachers who identified themselves as aware of the existence of 2e students in their classrooms and have successfully taught them. The third inclusion criterion was that only teachers who had at least 3 years of experience teaching 2e students in public schools. Teachers with less than 3 years of experience were excluded and those with no experience teaching 2e students were also eliminated.

We received a total of 27 responses to our initial inquiry, of which we eliminated 13 teachers as they did not meet one or more of the inclusion criteria. Therefore, we had a final group of 14 teachers to participate in the study and be interviewed regarding their perceptions. Before the interviews were conducted, we obtained a signed consent form from each teacher. This study involved 14 elementary school teachers, seven of whom were female and seven of whom were male. The age of the participants range between 28 and 45 years old; their range of teaching experience was 5–14 years. Most of the participants ($n = 12$; 85.71%) had a bachelor's degree; two had a master's degree. The participants were coded as follows: FT=female teacher and MT=male teacher (e.g., "MT2" stands for the second male teacher interviewed).

2.2. Instrument

The research instrument was a semi-structured, in-depth interview based on a list of eight questions, to obtain data from the participants regarding their perceptions and lived experiences (Vanden Berg, 2005). There were follow-up questions based on the interviews' directions. We began with a list of 10 questions, which were then reviewed by two experts in special education who teach special education curriculum at a Saudi university. Based on the responses of these experts, we eliminated two questions that were deemed unnecessary and made some adjustments to four questions so that they would be more understandable to our proposed participant population. This resulted in the 8 questions that were incorporated into each interview. During the interviews, based on the individual participant's responses, additional questions were asked to obtain clarification or to encourage the participant to expand upon their answer. The interview questions focused primarily on obtaining an understanding of the participants views on the giftedness characteristics and the learning challenges of 2e students. The interviews were conducted in a private room within the individual interviewee's school. Each interview was 45–60 min in length.

2.3. Procedure

To conduct the study, we first obtained permission from the Ministry of Education of Saudi Arabia. Due to the gender segregation of schools in Saudi Arabia, male teachers are not permitted to teach in girls' schools. Therefore, a female elementary school teacher with a postgraduate degree in special education was recruited to conduct the interviews with the female teachers. As previously mentioned, each interview was conducted in-person and was audio-recorded so it could be transcribed later. The transcriptions were created by the first

author who then had each interviewee review their interview transcript to confirm its accuracy. Each interviewer also took handwritten notes during the interviews for use during the data analysis process. The interviews were conducted in Arabic and the transcriptions were also done in Arabic.

2.4. Data analysis

A comprehensive approach to thematic analysis consisting of six stages that were followed during the data analysis of this research (Braun and Clarke, 2006). This approach works well with analyzing qualitative interview data (Braun and Clarke, 2006). First, all interviews were recorded and transcribed. Second, the transcription of each interview was read line-by-line to identify the initial codes. The first and second authors created their codes separately for each interview to ensure reliability and compare the results. Rigorous discussions between the researchers were conducted to identify any inconsistencies and disagreements and to ensure agreement. This helped develop a guideline according to which the rest of the transcribed data were coded. The results of the open-coding phase were repeatedly checked and read to ensure that the coded information was accurate.

Next, the researchers identified the emerging categories in the coded information, which were later merged to develop themes and subthemes (Braun and Clarke, 2006). Both researchers developed and reviewed the themes separately to ensure reliability. Then, the researchers work on assessing the themes utilizing a two-step process and jointly reflected on these data sources to ensure that the reality and coherence of the participants that were accurately reflected in the findings. This process enabled peer-checking and investigator triangulation (Creswell and Poth, 2018). Next, the themes and categories were evaluated and studied and make sure all codes are coalesced with their themes and had a profound connection (Creswell and Poth, 2018). To strengthen the transferability and trustworthiness of this study, the actual words of the participants are frequently presented. Finally, a thorough evaluation was conducted concerning the “narrative” of each theme and how each theme narrative is related to the overall narrative of the entire research data that answer research questions concerning the study questions, and then the final report was developed (Braun and Clarke, 2006, p. 92).

3. Results

Through the semi-structured interview questions that focused on exploring the participants’ perceptions and understanding of the characteristics of 2e individuals, two main themes were identified: (a) Giftedness Characteristics of Twice-Exceptional Students, and (b) Academic Challenges of Twice-Exceptional Students. Each main theme had subthemes, as shown in Table 1.

3.1. Theme 1: giftedness characteristics of twice-exceptional students

In the interviews, the teachers described the strengths of elementary school-aged twice-exceptional students in various ways

(see Table 1) and discussed gifted characteristics of these children they had observed in their personal experience. The related themes that emerged from their insights were: (a) excellent communication skills, (b) problem-solving skills, (c) strong retention, (d) curiosity, (e) abstract thinking, (f) creativity, and (g) leadership skills.

Communication skills were identified by most of the participants ($n = 12$; 85.71%) as a clear strength of 2e students. The participants indicated that these students have well-developed communications skills, emphasizing aspects such as their students’ ability to apply language effectively. One teacher, for example, stated “One of my students is very good in how [they] speak during class and has advanced oral presentation skills” (FT3); another said, “The strengths of twice-exceptional students is they speak well, ... I believe that they have good speaking skills [when called upon] in class” (MT6).

Many participants emphasized cognitive characteristics of 2e students and linked these to problem-solving skill. This theme emerged among both the male and female teachers. For example, FT5 stated, “Some students in my class were very good and highly-skilled in problem-solving”; and MT2 said, “Problem-solving is one of the strengths of twice-exceptional students, because they can identify the causes of [an] issue and understand it fully.”

In addition, strong retention and curiosity were identified by 10 participants (71.43%) as gifted characteristics of 2e students, which was referenced by FT1 with the following statement: “From my working experience as a teacher, twice-exceptional students display strong skills in areas such as strong retention, and some twice-exceptional students have high levels of curiosity and an eagerness to learn.” This was also remarked upon by MT4, who said: Twice-exceptional students demonstrate curiosity for learning new subjects...also they are curious about their peers and have the desire to learn everything new... [In general, they have] curiosity, uninhibited expression of their advanced ideas, and strong insights into highly complex topics.

The majority of the participants ($n = 12$; 85.71%) cited abstract thinking as another characteristic of these students, with MT3 mentioning, “When I teach them, I can see that they can learn abstract concepts very quickly.” Similarly, MT1 stated, “I am sure twice-exceptional students [are] abstract thinkers; it is one of their strengths. I have several students in my class that have this gift, but school should help them to develop such potential.” One participant, FT7, noted that she had identified a twice-exceptional student as a gifted underachiever. The teacher had decided that relocating the student within the classroom might make a difference, stating:

I had a student in Grade 3 who was underachieving in school ... I put this student in the front row and found that this caused her to become a more active participant in the class. I came to realize [after this move] that [this 2e student] has excellent abstract-thinking skills.

Quite a few of those interviewed ($n = 11$; 78.57%) stated they consider creativity to be an observed gifted characteristic of 2e students. They emphasized that these children tend to exhibit many aspects of creativity and that this is a trait of the twice exceptional (FT4; MT5; MT7). Similarly, leadership skills were cited by 11 (78.57%) of the teachers as a common strength of their 2e students, as described in statements such as: “[A] gifted characteristics of twice-exceptional students is that they are leaders, and they like to take responsibility for tasks beyond their chronological age” (MT4). This was also mentioned by FT2 who

TABLE 1 Themes and subthemes for each research question regarding the twice exceptional.

RQ	Themes	Sub-themes	<i>n</i>	%	Representative quotes
RQ1	Giftedness characteristics of 2e	Excellent communication skills	12	85.71	"They speak well."
		Problem-solving	13	92.86	"Some [twice-exceptional] students in my class were very good and highly skilled in problem-solving."
		Strong retention	12	85.71	"Twice-exceptional students display strong skills in areas such as strong retention."
		Curiosity	10	71.43	"Twice-exceptional students demonstrate curiosity for learning new subjects."
		Creativity	11	78.57	"There are many gifted characteristics of twice-exceptional students, such as being highly creative."
		Abstract thinking	12	85.71	"I had a student in Grade 3 who was underachieving in school ... I put this student in the front row, and I found her to be an excellent abstract thinker."
		Leadership skills	11	78.57	"They are leaders, and they like to take responsibility for tasks beyond their chronological age."
RQ2	Academic challenges of 2e	Poor reading skills	14	100.00	"A weakness of twice-exceptional students is poor reading skills; they struggle with reading when asked to read."
		Poor writing skills	14	100.00	"One of my best-performing students in terms of creativity, who is twice-exceptional, struggles significantly with writing skills."
		Math is challenging	11	78.57	"The math teacher in the school told me that several twice-exceptional students in our school have weak mathematics skills."
		Difficulties in concentration	10	71.43	"One of my students performs exceptionally in many subjects but still has an issue where he cannot concentrate well."
		Low self-esteem	9	64.29	"I can say that one of the difficulties twice-exceptional students have is feelings of low self-esteem."

said, "Twice-exceptional students are leaders, whether in class projects or sports, and are popular and persuasive." Another comment that referenced this was, "Twice-exceptional students tend to be popular and persuasive, and to function as leaders during class" (FT6).

3.2. Theme 2: academic challenges of twice-exceptional students

Under the theme regarding challenges of the twice-exceptional, several subthemes emerged, which were: (a) Poor Reading Skills, (b) Poor Writing Skills, (c) Math Is Challenging, (d) Difficulty Concentrating, and (e) Low Self-Esteem.

Almost all the participants ($n = 13$; 92.86%) highlighted poor reading skills as a distinguishing characteristic of 2e students, with comments such as, "A weakness of twice-exceptional students is poor reading skills in that when called upon to read aloud, they tend to struggle" (MT7). This was also noted specifically by FT4.

All respondents indicated that poor writing ability is a common challenge for 2e students. However, because of their learning disabilities, some twice-exceptional students struggle with motor skills related to their handwriting, which might interfere with their ability to complete assignments that require hand-written responses. Examples of related statements on this topic were as follows: "One of my best-performing twice-exceptional students in creativity struggles

significantly with writing skills" (FT2); and "There are several weaknesses of twice-exceptional students, such as difficulties with writing" (MT5).

Another sub-theme that emerged under that of the challenges of 2e students was "Math Is Challenging." Participants noted these students do poorly in mathematics (FT6), and also mentioned that this is something other teachers mention to them when they have discussions regarding their 2e students. For example, one teacher referenced asynchronous development when stating, "The math teacher in the school told me that several twice-exceptional students in our school have weak mathematics skills" (MT1).

A large number of participants ($n = 10$; 71.43%) also mentioned that some 2e students have trouble concentrating, noting they may "performs exceptionally in many subjects but still have an issue where [they] cannot concentrate" (MT3). A female teacher similarly said, "Some [2e students] are unable to concentrate well, and it is the responsibility of teachers to [help them] address this problem" (FT1). Participants also pointed to these students having limited attention spans, stating they are easily distracted during class activities (MT2). MT6 also referenced this sub-theme by saying, "I think the twice-exceptional student in my class is easily distracted and has difficulty completing assignments."

The issue of low self-esteem is another trait that the teacher participants identified as a perceived characteristic of 2e students (MT4). Similarly, a female teacher indicated these students not only

have low self-esteem, but that this trait causes them to struggle in school (FT5).

In general, participants were aware that twice-exceptional students exist, meaning these teachers were able to perceive that these students are gifted despite their different disabilities. However, although participants were able to describe certain negative and positive characteristics of 2e students, the quality and depth of their knowledge and perceptions of the traits of 2e students were limited; for example, they exhibited little knowledge of such characteristics as problem-solving, strong retention, curiosity, abstract thinking, creativity, and leadership characteristics.

4. Discussion

This study focused on exploring the perceptions and understanding of Saudi elementary school teachers regarding the characteristics of 2e students. In general, we found the teachers to be aware of the fact that some students can present with both giftedness and learning disability at the same time.

One of our main findings was that the elementary school teachers believe that 2e students are more developed than their peers in terms of language and communication skills. This finding echoes those of previous studies, which have argued that 2e students are different from other students in their level of language use (Wright and Ford, 2017; Alsamir, 2019). Chimhenga (2016) found that in contrast with their peers, these students may possess extensive vocabularies, have high listening comprehension skills, and fluent speech.

We also found that Saudi elementary school teachers believe 2e students possess advanced problem-solving skills, which is also consistent with the findings of other researchers, including Coleman and Gallagher (2015) and Roberts et al. (2015). Moreover, in comparison with their peers who solely have disabilities that impact learning, Ritchotte and Zaghawan (2019) found that 2e students significantly positively differ in their problem-solving skills, particularly in area of the coping strategies they develop to deal with difficult school situations.

As was found by other researchers, our teacher participants also identified strong retention as a strength of 2e students (Baldwin et al., 2015; Alsamir and Aljohani, 2018). The Saudi elementary school teachers in our study also indicated that twice-exceptional students have high levels of curiosity and an eagerness to learn, which was a trait also identified by Lee and Ritchotte (2019).

Unsurprisingly, most of the teachers interviewed noted abstract-thinking skills as a characteristic of 2e students in Saudi schools, which again comports with the findings of past research (Levi, 2019; Firat and Bildiren, 2022). In agreement with our findings, Wellisch (2016) stated strong abstract thinking skills is commonly associated with giftedness in 2e students.

Consistent with the literature, this study also found that according to our participants, creative thinking is a distinguished characteristic of 2e students (Ritchotte and Zaghawan, 2019; Mohammad, 2020). Renzulli (2005) described two groups of gifted students, the academically gifted and the creatively gifted. As with the creatively gifted, 2e students possess advanced creative thinking skills (Alsamir, 2020). Nielsen and Higgins (2015) found, as we did, that leadership skills are another characteristic of the 2e in terms of their relationships

with their peers and the difficulty of the tasks they are willing to take on.

On the other hand, this research identified certain characteristics of these students that contribute to the academic challenges they often experience; these traits, as noted by the study participants, include poor reading comprehension, poor writing skills, challenges with mathematics, poor concentration, and low self-esteem (Firat and Bildiren, 2022). However, despite the teachers noting these traits, few of them indicated that they had referred 2e students to services and support for their learning disabilities. This lack of referral for special education support was also mentioned by Ritchotte and Zaghawan (2019).

Several of the teachers in the study noted they had students who showed a high level of curiosity but displayed poor verbal and writing skills in different subjects, such as Arabic studies; previous studies have obtained similar findings (Mayes and Moore, 2016; Ramsay-Cohen, 2016). However, although the participants noted these issues, they did not draw conclusions about the implications of these conflicting findings and did not provide more in-depth information on them. We found, as did Mayes and Moore (2016), that teachers identified the following as weaknesses of 2e students in the general area of writing: difficulty in spelling and poor handwriting; the absence of organizational skills; and low scores on tests and assessments. This might be related to the limited training provided for elementary school teachers regarding 2e students and their unique traits (Al-Hroub and Whitebread, 2019).

One trait our participants mentioned that has received significant attention in educational research literature, such as Weiss (2022), is that 2e students perform poorly in mathematics compared to their peers. Amran and Majid (2019) also found that some 2e require educators consider accommodations and best teaching strategies for those with learning disabilities when instructing this population in mathematics to address issues with focus.

Regarding the trait of low self-esteem, which our participants noted as a common trait of 2e students, our findings are supported by those of Amran and Majid (2019), which noted low self-esteem—as well as a lack of motivation, depression, and stress—in this population.

5. Conclusion

Saudi elementary school teachers have an important role in identifying 2e students and ensuring these children receive the full range of support they require. As a result, it is necessary they be fully aware of all the traits such children possess, both in terms of strengths and weaknesses. The present study found that the teachers are aware of various characteristics 2e students exhibit in the classroom and are able to identify them as related to their giftedness traits (e.g., communication skills, problem-solving, strong retention, curiosity, abstract thinking, creativity, and leadership) and as related to disability traits that can impact learning outcomes (e.g., poor reading comprehension, poor writing skills, difficulties with mathematics, poor memory and focus, and low self-esteem).

6. Limitations

In the current study, we explored the perceptions of Saudi elementary school teachers regarding the characteristics of 2e

students using the qualitative method of in-depth, semi-structured interview. While this allowed us to obtain deep insights that could not have been achieved through quantitative methods, this does make our findings less generalizable to other contexts, such as other regions of the country. Since we only obtained data from 14 individuals, this is also a factor that limits generalizability to the larger population of all Saudi elementary school teachers. Moreover, our subjects were relatively young and therefore might be more open to newer trend in education in Saudi, such as accommodations for students with special education needs and the importance of teacher awareness regarding the identification processes for students, whether gifted, challenged in some way, or both. This study was also limited to only elementary school teachers. Perceptions of other types of educators, such as special and gifted education teachers, are also important to capture.

7. Implications for practice

From our findings, while Saudi elementary school teachers are aware of the common traits of 2e students, it appears they are not well-trained in how to best support students who possess both gifted traits and also struggle with learning due to different types of disabilities. Moreover, our study participants indicated that they only infrequently refer identified 2e students for special education services. To address this critical gap in the support of twice-exceptional students, it is necessary that stakeholders—especially the Ministry of Education—invest in training providers and administrators in developing programs to better serve these students as well as in professional development to better educate teachers regarding the special needs of 2e students. Since teachers are often the first professionals to identify gifted learners for accelerated programs, they are uniquely positioned to identify when a gifted learner also has a disability and/or academic challenges. However, if teachers are not better trained, preferably during their university studies, in the true characteristics of 2e students as well as in the special education services these students require, this population might only be served in terms of their gifted traits and not receive necessary support for their learning disabilities, leading to chronic academic underachievement and even failure.

8. Considerations for future research

The current study focused only on the perceptions of elementary school teachers. Other types of educators play important roles in the education of 2e students, including special education teachers, gifted education teachers, school personnel, and parents. Obtaining qualitative data on the views of these other populations is an important

area for future research. Moreover, in future studies researchers should directly obtain information from 2e students themselves regarding their feelings and perceptions of their needs, learning experiences, and the difficulties they face in education. In the future, we should also explore whether teachers at all educational levels perceive 2e students as a separate and unique population requiring special education support that possesses a unique set of traits, rather than obtaining data from participants based on what they perceive as strictly gifted traits and what they perceive as traits specifically connected to the learning challenges of this group of 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 Standing Committee for Research Ethics at Hail University-Research Number: H-2022-394. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YA reviewed the literature, collected the data, and collaborated in data analysis and discussion. OA wrote the methodology, result, and discussion and collaborated in data analysis. SA drew the conclusion and collaborated in data analysis and the literature review. All authors helped to draft the manuscript, read, and approved the final manuscript.

Conflict of interest

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

Publisher's note

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

References

- Akinbote, R. O., Olowe, P. K., and John, N. I. (2017). Pre-elementary school teachers' knowledge of characteristics and nurturing of gifted children in Ondo West Local Government Area. *Afr. J. Educ. Res.* 21, 28–38.
- Al-Hroub, A., and Whitebread, D. (2019). Dynamic assessment for identification of twice-exceptional learners exhibiting mathematical giftedness and specific learning disabilities. *Roeper Rev.* 41, 129–142. doi: 10.1080/02783193.2019.1585396
- Alkhunani, N. A. (2013). Twice-exceptional learners: Saudi teachers' awareness and referral of gifted students with disabilities. unpublished master's thesis. Elon: University of Elon.
- Alsamani, O. A. (2020). Fostering creative productivity for creatively gifted students. *Teaching for High Potential Journal.* 19–22.
- Alsamiri, Y. A. (2018). Teachers' perspectives of identifying students of Saudi primary students with giftedness and learning disabilities. doctoral thesis, School of Education, UNSW.
- Alsamiri, Y. A. (2019). Twice-exceptional students in Saudi Arabia: a review of the research literature from 1990 to 2018. *Int. J. Learn. Manag. Sys.* 7, 63–70. doi: 10.21608/ijlms.2019.200140

- Alsamiri, Y. A., and Aljohani, S. A. (2018). Learning disabilities teachers' attitudes about professional development to address the needs of students with gifted and learning disabilities (SGLD): a qualitative study. *Global J. Health Sci.* 11, 81–91. doi: 10.5539/gjhs.v11n1p81
- Amran, H. A., and Majid, R. A. (2019). Learning strategies for twice-exceptional students. *Int. J. Spec. Educ.* 33, 954–976.
- Baldwin, L., Omdal, S. N., and Pereles, D. (2015). Beyond stereotypes: understanding, recognizing, and working with twice-exceptional learners. *Teach. Except. Child.* 47, 216–225. doi: 10.1177/0040059915569361
- Barnard-Brak, L., Johnsen, S. K., Pond Hannig, A., and Wei, T. (2015). The incidence of potentially gifted students within a special education population. *Roeper Rev.* 37, 74–83. doi: 10.1080/02783193.2015.1008661
- Bianco, M., and Leech, N. (2010). Twice-exceptional learners: effects of teacher preparation and disability labels on gifted referrals. *Teach. Educ. Div. Council. Except. Child.* 33, 319–334. doi: 10.1177/0888406409356392
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp0630a
- Buică-Belciu, C., and Popovici, D.-V. (2014). Being twice exceptional: gifted students with learning disabilities. *Procedia Soc. Behav. Sci.* 127, 519–523. doi: 10.1016/j.sbspro.2014.03.302
- Chessman, A. (2005). Policy development and practice: the New South Wales experience. *Int. Educ. J.* 6, 156–163.
- Chimhenga, S. (2016). The challenges of intervention practices for gifted children with learning disabilities in elementary schools of Zimbabwe. *Glob. J. Adv. Res.* 3, 764–770.
- Cody, R. A., Boldt, G. T., Canavan, E. J., Gubbins, E. J., Hayden, S. M., Bellara, A. P., et al. (2022). Teachers' reported beliefs about giftedness among twice exceptional and culturally, linguistically, and economically diverse populations. *Front. Psychol.* 13:953059. doi: 10.3389/fpsyg.2022.953059
- Coleman, M. R., and Gallagher, S. (2015). Meeting the needs of students with 2e: it takes a team. *Gift. Child Today* 38, 252–254. doi: 10.1177/1076217515597274
- Coolahan, J. (2004). The historical development of teacher education in the Republic of Ireland. *Teach. Educ. Repub. Ireland*, 3, 3–11
- Creswell, J. W., and Poth, C. N. (2018). *Qualitative inquiry and research design: choosing among five approaches*. 4th Edn. Thousand Oaks: Sage.
- Firat, T., and Bildiren, A. (2022). The characteristics of gifted children with learning disabilities according to preschool teachers. *Early Years*, 1–17. doi: 10.1080/09575146.2022.2034755
- Hays, E. A. (2016). Academic outcomes in higher education for students screened as twice-exceptional: gifted with a learning disability in math or reading. doctoral dissertation, University of Tennessee. TRACE: Tennessee research and creative exchange. Available at: https://trace.tennessee.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=4822&context=utk_graddiss
- Hopwood, K. A. (2019). Twice-exceptionality: Teachers' awareness and training on their effects on the academic, social and emotional outcomes of students (publication no. 27668759). doctoral dissertation, Concordia University. ProQuest Dissertations and Theses Global. Available at: <https://www.proquest.com/openview/c3e110bd104b8269b663e24aee9381ca/1?pq-origsite=scholar&cbl=18750&diss=y>
- Lee, C.-W., and Ritchotte, J. A. (2019). A case study evaluation of the implementation of twice-exceptional professional development in Colorado. *J. Educ. Gift.* 42, 336–361. doi: 10.1177/0162353219874440
- Levi, R. (2019). Supporting twice exceptional students within the general education classroom: strategies for general education teachers [capstone project, Hamline university]. Available at: https://digitalcommons.hamline.edu/hse_cp/403/
- Maddocks, D. L. S. (2020). Cognitive and achievement characteristics of students from a national sample identified as potentially twice exceptional (gifted with a learning disability). *Gift. Child Q.* 64, 3–18. doi: 10.1177/0016986219886668
- Mayes, R. D., and Moore, J. L. I. I. (2016). The intersection of race, disability, and giftedness: understanding the education needs of twice-exceptional, African American students. *Gift. Child Today* 39, 98–104. doi: 10.1177/1076217516628570
- Merriam, S. B., and Tisdell, E. J. (2016). *Qualitative research: a guide to design and implementation*. 4th Edn. San Francisco, California: Jossey-Bass.
- Mohammad, A. A. (2020). The effects of disability labeling on teachers' referrals of twice-exceptional children to gifted programs in Saudi Arabia. doctoral dissertation, University of Northern Colorado. Scholarship and Creative Works @ digital UNC. Available at: <https://digscholarship.unco.edu/dissertations/657/>
- Nielsen, M. E., and Higgins, L. D. (2015). Into the wilderness...: the parents of a twice-exceptional child look back and offer hope. *Gift. Child Today* 38, 220–222. doi: 10.1177/1076217515597278
- Ramsay-Cohen, P. (2016). Where do they belong? Teacher experiences with twice-exceptional students in streamed gifted classes. unpublished master's thesis. University of Toronto. Available at: https://tspace.library.utoronto.ca/bitstream/1807/72275/1/RamsayCohen_Pamela_D_201606_MT_MTRP.pdf
- Renzi, J. S. (2005). "The three-ring conception of giftedness: a developmental model for creative productivity" in *Conceptions of giftedness*. eds. R. J. Sternberg and J. E. Davidson (Cambridge, England: Cambridge University Press), 53–63.
- Ritchotte, J. A., and Zaghlawan, H. Y. (2019). Coaching parents to use higher level questioning with their twice-exceptional children. *Gift. Child Q.* 63, 86–101. doi: 10.1177/0016986218817042
- Roberts, J. L., Pereira, N., and Knotts, J. D. (2015). State law and policy related to twice-exceptional learners: implications for practitioners and policymakers. *Gift. Child Today* 38, 215–219. doi: 10.1177/1076217515597276
- Shaunessy, E. (2007). Implications for gifted education. *Gift. Child Q.* 51, 119–135. doi: 10.1177/0016986207299470
- Van den Berg, H. (2005). Reanalyzing qualitative interviews from different angles: the risk of decontextualization and other problems of sharing qualitative data. *Hist. Soc. Res.* 6, 179–192. doi: 10.17169/fqs-6.1.499
- Weiss, C. H. (2022). Teacher practices, beliefs, and conceptual understanding of mathematics: a phenomenological case study of teachers instructing mathematically gifted and promising students [doctoral dissertation, the University of North Carolina at Charlotte]. Available at: <https://graduateschool.charlotte.edu/teacher-practices-beliefs-and-conceptual-understanding-mathematics-phenomenological-case-study>
- Wellisch, M. (2016). Gagné's DMGT and underachievers: the need for an alternative inclusive gifted model. *Aust. J. Gift. Educ.* 25, 18–30. doi: 10.21505/ajge.2016.0003
- Wormald, C., Rogers, K. B., and Vialle, W. (2015). A case study of giftedness and specific learning disabilities: bridging the two exceptionalities. *Roeper Rev.* 37, 124–138. doi: 10.1080/02783193.2015.1047547
- Wright, B. L., and Ford, D. Y. (2017). Untapped potential: recognition of giftedness in early childhood and what professionals should know about students of color. *Gift. Child Today* 40, 111–116. doi: 10.1177/1076217517690862



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Johanna Pöysä-Tarhonen,
University of Jyväskylä, Finland
Gerard Rayner,
Swinburne University of Technology, Australia

*CORRESPONDENCE

Ling Jiang
✉ ljiangcn@ccnu.edu.cn
Heng Luo
✉ luoheng@mail.ccnu.edu.cn

RECEIVED 17 February 2023

ACCEPTED 10 April 2023

PUBLISHED 05 May 2023

CITATION

Wang M, Jiang L and Luo H (2023) Dyads or quads? Impact of group size and learning context on collaborative learning.
Front. Psychol. 14:1168208.
doi: 10.3389/fpsyg.2023.1168208

COPYRIGHT

© 2023 Wang, Jiang and Luo. 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.

Dyads or quads? Impact of group size and learning context on collaborative learning

Mindi Wang, Ling Jiang* and Heng Luo*

Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, China

Collaborative learning has been widely used in both offline and online contexts to support deep learning, and its effectiveness may be adjusted by the size of the collaborative groups. To examine the effect of learning context and group size on collaborative learning, this study conducted two experiments with 62 third-year undergraduate students enrolled in the course named *Application of Modern Educational Technology* to compare learning outcomes, learning engagement, and collaborative experience between quad (four-person) and dyad groups in both face-to-face and online learning contexts. The results indicated that learning outcomes and collaborative experience were not significantly affected by group size and learning context, but for peer interaction, the dyad group showed more communication and interaction during the learning process. In general, the dyad group showed higher and more stable scores in all aspects, as well as being able to adapt to changes in learning contexts. Based on the research results, three practical implications were proposed to promote the implementation of collaborative learning in teaching.

KEYWORDS

collaborative learning, group size, learning context, learning outcome, learning engagement, collaborative experience

1. Introduction

Collaborative learning has been widely used in various disciplinary fields in both face-to-face and online learning contexts. The literature has affirmed the advantages of collaborative learning in traditional classrooms: it promotes improvement in students' higher-order thinking ability, such as problem-solving, critical thinking, and creative thinking (Segundo Marcos et al., 2020; Liang et al., 2021), and also provides students with improved learning experience and superior learning outcomes (Laal and Ghodsi, 2012). Many researchers have also reported the benefits of collaborative learning in online learning environments. For example, Dewiyanti et al. (2007) found that students had a more positive collaborative experience and greater satisfaction in an asynchronous online environment. Similarly, Strauß and Rummel (2020) proved that integrating collaborative learning into online courses is beneficial to students' learning performance and social presence. Synthesizing the results from many empirical studies, several meta-analyses have confirmed the positive impact of collaborative learning across various educational contexts (Ibrahim and Harun, 2015; Chen et al., 2018; Kaldırım and Tavşanlı, 2018; Talan, 2021).

While the literature supports the overall effectiveness of collaborative learning, its effectiveness is largely dependent upon specific instructional strategies such as task design, role scripts, prework activities, and group size (Pfister and Oehl, 2009; Luo et al., 2023). Group size merits particular attention due to its ease of implementation and great influence on social interaction (Kim, 2013; Wang et al., 2020). Generally, small groups consist of 14 or

fewer members, medium groups consist of 15–34 members, and large groups have a minimum size of 35 members (Benton et al., 2015; Parks-Stamm et al., 2017). The literature has identified benefits and limitations for each group size during collaboration: large or medium groups can produce more diverse perspectives but are susceptible to social loafing (Qiu and McDougall, 2015; Amichai-Hamburger et al., 2016), while small groups tend to promote dyadic communication and individual contributions but can be heavily influenced by group dynamics and relationships (Kim, 2013; Yang et al., 2022). Overall, the empirical evidence supports the superiority of small groups over large or medium groups in promoting collaborative learning behaviors, experiences, and achievements (Shaw, 2013; Luo et al., 2023).

However, the optimal group size for collaborative learning within the small range has not yet been identified in the literature. While some researchers have shown that students in dyad groups outperformed those in groups of three or four (Lohman and Finkelstein, 2000; Kim et al., 2020), other studies found that quad (four-person) groups functioned better than dyads in terms of learning performance and social discourse (Peltokorpi and Niemi, 2019; Corrége and Michinov, 2021). Additionally, the optimal group size might vary with different disciplinary domains. For instance, groups of two or three members were found to be superior in technology courses such as robotics and artificial intelligence (Bianco, 2014; Zhan et al., 2022) while groups of five demonstrated enhanced social interactions in English courses (Abuseileek, 2012). Furthermore, the findings regarding group size are known to be moderated by the type of classroom settings where collaborative learning occurs (Corrége and Michinov, 2021); however, few studies have systematically compared the effect of group size between the offline and online learning contexts, and little is known about the optimal group size for each contextual setting.

To address this research need, we conducted two experimental studies in an undergraduate course with 62 third-year students randomly assigned to quad and dyad groups. One experiment took place in a face-to-face offline learning context, while the other experiment occurred in an online learning context. The purpose of the study was to determine the impact of group size (four person vs. dyad) on collaborative learning in both offline and online contexts and to explore the moderating effect of learning context. More specifically, we sought to answer the following research questions:

1. Do collaborative learning outcomes differ between quad and dyad groups in offline and online learning contexts?
2. Does collaborative learning engagement differ between quad and dyad groups in offline and online learning contexts?
3. Do collaborative experiences differ between quad and dyad groups in offline and online learning contexts?
4. What are the moderating effects of learning context on students' learning outcomes, learning engagement, and collaborative experiences in both quad and dyad groups?

2. Literature review

2.1. Collaborative learning

Learning occurs in the social process of knowledge construction, not as a single effort (Vygotsky, 1978). Collaborative learning, as a learning approach, allows students to organize activities together in the form of groups, forming a small learning community in which everyone has shared responsibility and power to construct knowledge and meaning (Moreira, et al., 2016). Collaborative learning also provides teachers with activities and strategies to facilitate student-centered learning, and it is widely accepted and practiced in both face-to-face and online contexts. According to Dillenbourg (1999), collaborative learning describes “a situation in which particular forms of interaction among people are expected to occur, which would trigger learning mechanisms” (p. 5). Wang and Liao (2017) emphasized the constructive nature of collaborative learning and argued that it enables knowledge construction through mutual participation and motivation of group members. Gu et al. (2015) further identified two key conditions of effective collaborative learning: positive interdependence and individual accountability through which all members contribute to the meaning making and knowledge sharing process. Synthesizing the existing definitions in the literature, in this study, we denote collaborative learning as two or more students completing learning tasks and solve learning problems together through mutual participation, division of labor, and coordinated efforts.

In recent years, with the development of internet technology, collaborative learning has also been applied to the online learning environment. The advantages, challenges, and solutions in the application of collaborative learning in the new environment have formed a new research field: computer-supported collaborative learning (CSCL). The main differences between CSCL and traditional collaborative learning are location (online virtual environment) and time (both synchronous and asynchronous). However, even if the form of collaborative learning changes, its characteristics and purpose are to promote learning outcomes (Shen and Wu, 2011; Wu, 2018), to create better learning engagement (Gopinathan et al., 2022) and to provide a collaborative experience (Burgess et al., 2014; Liu et al., 2017). These three components are interrelated. Higher learning engagement and a positive collaborative experience promote improvements in learning outcomes, which in turn affect learning engagement and the collaborative experience. At present, many scholars have also used social network analysis to explore peer interaction in the process of collaboration, but we believe that interaction in the process of collaboration is not the purpose of collaboration, but rather a means to enable learners to obtain better results and experience.

2.2. Effect of group size in the offline context

In our studies, offline environment refers to face-to-face teaching in the real world. In this environment, peer discussion and social interaction play a crucial role in the process of collaboration. When carrying out collaborative learning, we usually ask, “How many people are in the optimally sized group?” No consensus has, however, been

Abbreviations: CSCL, computer-supported collaborative learning.

reached on the answer to this question. Research has shown that students prefer smaller groups, including peer teaching. For example, Kooloos et al. (2011) divided a 15-person group into three groups of five people; they achieved higher learning gains, greater goal achievement, and a more active sense of participation. Slavin (1987) found that the results of paired learning were better than those for groups of four or more people. Other studies have tried to distinguish the differences between dyads and three-person groups (Jensen and Wiley, 2006; Zhan et al., 2022). Jensen and Wiley (2006) showed that in the task of solving arithmetic problems, three-person groups showed better reasoning ability and more effective password problem solving ability than the dyads. In artificial intelligence courses, Zhan et al. (2022) determined that the dyads were better than the three-person groups in learning motivation and quality of collaborative problem solving over the 6-week course, and also imposed more cognitive load on the individual students. There have, however, also been studies that suggest a group of four people is the optimal group size (Alexopoulou and Driver, 1996; Sung et al., 2017).

2.3. Effect of group size in the online context

With the rapid advancement of internet and mobile technologies, online learning is becoming increasingly popular in higher education and has emerged as a major context for CSCL (Hmelo-Silver and Jeong, 2021; Gopinathan et al., 2022). The literature has also explored the effect of group size on online CSCL. A recent meta-analysis revealed a significant influence for group size on online problem-based learning (Saqr et al., 2019): larger groups tended to lead to weaker the team cohesion, worse individual performance, and less social interaction. Through asymmetric collaborative simulation, Rannastu et al. (2019) found that dyads reported better collaboration than quads in grade six. Similarly, several researchers have found that small groups of three members led to slightly better participation and performance in online discussion (Shaw, 2013; Luo et al., 2023).

Similar to the offline environment, there have also been studies that show that a group of four people is ideal (Sugai et al., 2018) for using social networks for collaborative argumentation. Compared with offline environment, however, online collaborative learning faces some challenges. Due to the lack of facial expressions, intonation, and gestures in the online environment, the dialogue between members is more likely to be misunderstood, and they need more time and energy to maintain communication effectiveness. The online environment is also more likely to foster social loafing, even as it provides convenience in time and place through asynchronous communication.

To sum up, we have reason to believe that group size in collaborative learning tends to be more collaborative in small groups in either offline or online environments, and this may be related to the relationship between individual responsibility and motivation. In group decision-making, individual input is often lower than in individual decision-making. Dyads have an equal relationship in the process of collaboration, have more opportunities to participate, and team members can concentrate on thinking without being influenced by other members (Dugosh et al., 2000). The quad group is still small in scale, but it can increase the diversity of views and mutual feedback while also minimizing the possibility of “free riding” (Shimazoe and Aldrich, 2010). Therefore, this study used the group sizes of two and

four people per group and compared the changes in offline and online learning contexts to determine the impact of different group sizes on collaborative learning in different learning contexts.

3. Methods

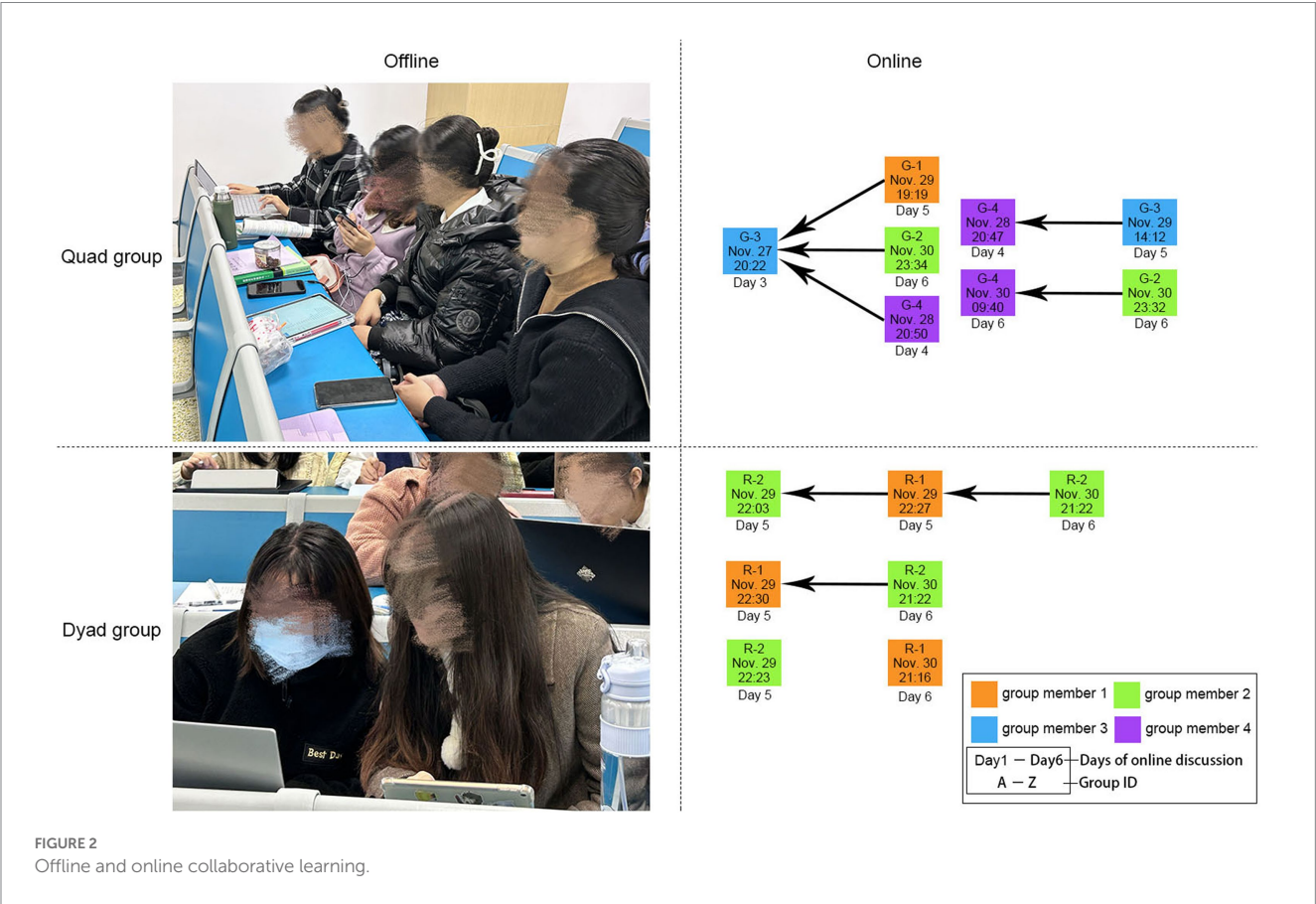
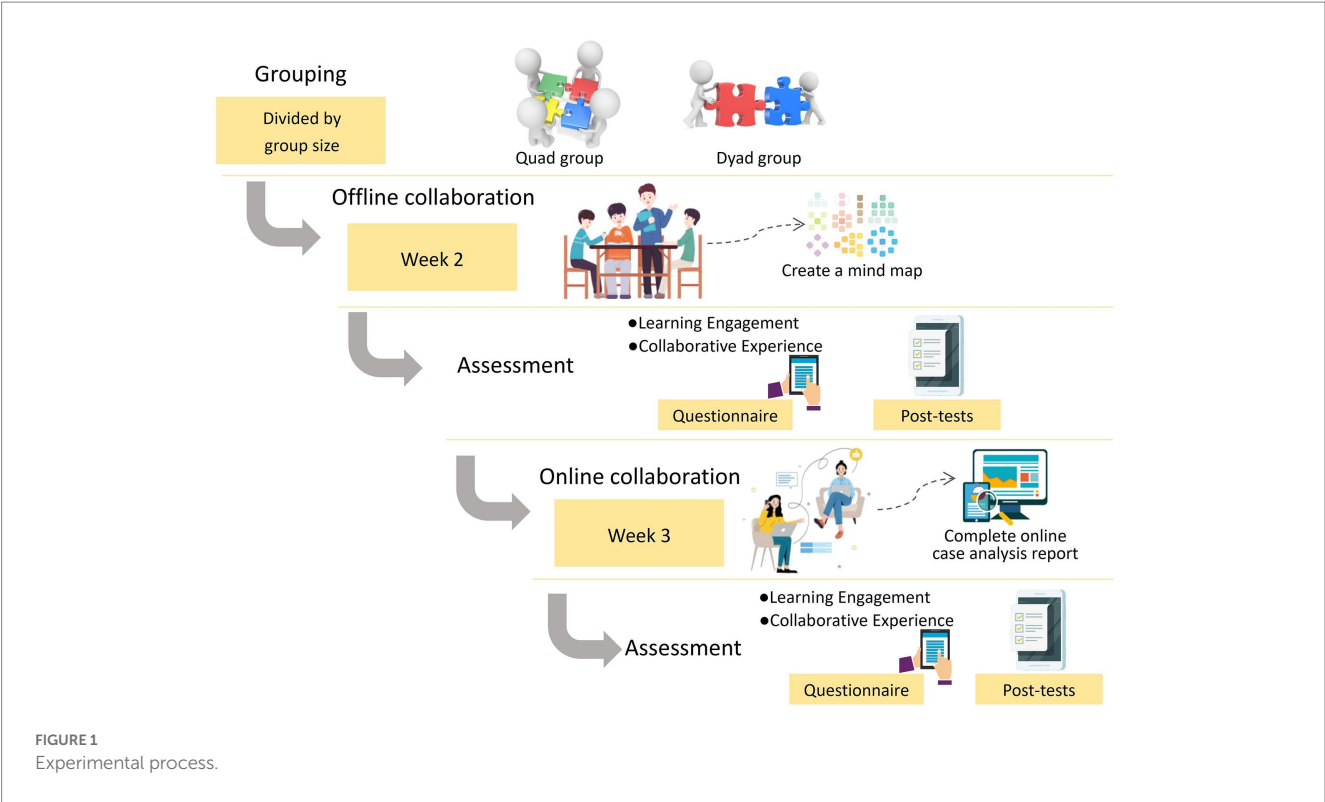
3.1. Participants

The present study was conducted in the first 4 weeks of a 7-week blended course named *Application of Modern Educational Technology* offered by a research university in central China. A total of 62 third-year students agreed to participate; participants were from the undergraduate programs of chemistry (teacher training) and mathematics (teacher training). All of the participants were randomly assigned into eight quad groups ($n=32$) and 15 dyad groups ($n=30$). All of the groups had participated in two sessions of collaborative learning in both offline and online learning contexts, each lasting for 2 weeks. The gender ratio was roughly 7.9:1 (55 female students and seven male students), and the mean age was 20.32 (min = 19, max = 23). The research protocols and instruments were reviewed and approved by the Institutional Review Board of Central China Normal University (CCNU-IRB-202103019, approved on 2021/03/16). Informed consent was obtained from all of the participants before implementing the study.

3.2. Procedure

The entire experimental process is shown in Figure 1. In the first week of the course, the teacher gave an overview of the course and explained the basic information such as the course requirements and assessment methods. A total of 62 students were randomly divided into eight quad groups ($n=32$) and 15 dyad groups ($n=30$). In the second week, all of the participants worked in groups offline to complete a collaborative learning assignment that involved the creation of a mind map while attending a face-to-face class. The offline collaboration session lasted about 45 min. In the third week, the participants switched to the online learning mode and worked together with their group members to study and discuss a learning design case in an online discussion forum. The online collaboration session was asynchronous and lasted for 6 days. Afterward, all of the participants were required to submit a case analysis report. Both the offline and online tasks tested the participants' comprehension of leaning and teaching theories and their higher-order thinking skills such as application, analysis, and creation. At the end of each collaborative learning session, the participants were required to complete a learning experience questionnaire and submit the task assignment (i.e., mind map or case analysis report) individually.

Figure 2 shows how the participants collaborated with their group members in both offline and online contexts. In the offline collaboration, the group members sit adjacently in one row for convenient communication, and used various electronic devices such as tablets, mobile phones, and laptops to jointly complete the mind map assignment. In the online collaboration, the social interactions among the group members took mainly the form of discussion threads. As seen in Figure 2, the majority of the messages were posted in the latter phase of the discussion period with few threads extended over three interactions; there were also many instances of “orphan” postings that received no replies.



It is important to note that, although this experiment has two independent variables: group size (quad group vs. dyad group) and learning context (offline vs. online), we decided to investigate their effect on collaborative learning separately rather than simultaneously through a two-by-two factorial design. There are two reasons behind this decision: First, synchronous face-to-face communication and asynchronous online communication are inherently different featured by varying activity length and interaction intensity, and the confounding influence of such differences cannot be eliminated from the two-by-two design. Second, the focus of the present study is on comparing the effectiveness of dyad and quad group, with learning contexts being examined as a possible moderating factor. Consequently, we believed that conducting two separate experiments in each learning context can better answer our research questions and provide more flexibility in research design such as choosing different collaboration task or collaboration length.

3.3. Data collection and instruments

The data were collected using four instruments: a learning experience questionnaire, a post-test on learning and teaching theory, grading criteria for the collaborative learning assignment, and a semi-structured interview protocol.

3.3.1. Learning experience questionnaire

The learning experience questionnaire comprised three parts. The first part included five items about personal information: name, student ID, gender, age, and group size. The second part was a subscale measuring learning engagement in terms of cognitive, social, and behavioral engagement; it included 17 items that were informed by the works of Fredricks et al. (2004), Gunuc and Kuzu (2015), and Ma and Zhou (2019). The third part was a subscale measuring collaborative experience, which included 14 items that were informed by the works of Barron (2000) and Miao et al. (2017). All of the subscale items were scored on a five-point Likert scale ranging from 1, “strongly disagree,” to 5 “strongly agree.” The complete questionnaire items are shown in [Supplementary Appendix A](#). During the research process, the subscales for learning engagement and collaborative experience were tested twice. The Cronbach's α values for learning engagement were 0.777 and 0.927, and for collaborative experience they were 0.867 and 0.950, which indicated good internal reliability.

3.3.2. Post-tests for collaborative learning

The research group developed two post-tests to measure students' collaborative learning outcomes. The first one was given at the end of the offline collaboration session focusing on learning theories (e.g., behaviorist, cognitive, and socio-cognitive theories), and the second one was given at the end of the online collaboration session focusing on teaching theories (e.g., nine events of instruction, first principles of instruction, and community of inquiry framework). Both tests consisted of 20 multiple-choice questions with a maximum total score of 100 points (five points for each question). Despite the different testing content, the two post-tests were considered equivalent in test format, structure, and difficulty.

3.3.3. Grading criteria for collaborative learning assignment

During the entire experiment, students from different groups completed two collaborative learning assignments in different learning

context—the production of mind maps and the writing of online case analysis reports. The former focused on learning theory and the latter on teaching theory. Students' learning outcomes could be measured through the analysis of the two assignments. The two collaborative learning assignments were independently scored by four researchers according to three aspects. The scoring criteria for the mind map were correctness (correctly expressing multiple concepts and their relationships related to the subject, without obvious intellectual errors); comprehensiveness (reflecting the relevant content of the subject as fully as possible); and rendering effect (concision or selecting and refining keywords, clarity of structure, and the rational use of symbols and diagrams). The scoring criteria for the case analysis reports were correctness (correct expression of the problems in the case without obvious intellectual errors); comprehensiveness (reflecting the teaching problems in the case as fully as possible); and quality of the description (proper number of words, in-depth analysis of the problems and solutions, and the level of detail of the description). The interrater reliability Spearman's rho was greater than 0.9.

3.3.4. Interview protocol

To understand how students' collaborative learning experience varied by group size and learning context, we conducted semi-structured interviews with 10 selected participants immediately after the second experiment to gather their in-depth opinions about the contextualized collaboration in terms of perceived benefits, challenges, peer- and self-evaluation, and improvement suggestions, as well as their preference for group size and learning context. A total of 110 min of interview data were recorded and transcribed for future analysis. Please refer to [Supplementary Appendix B](#) for details about the interview protocol.

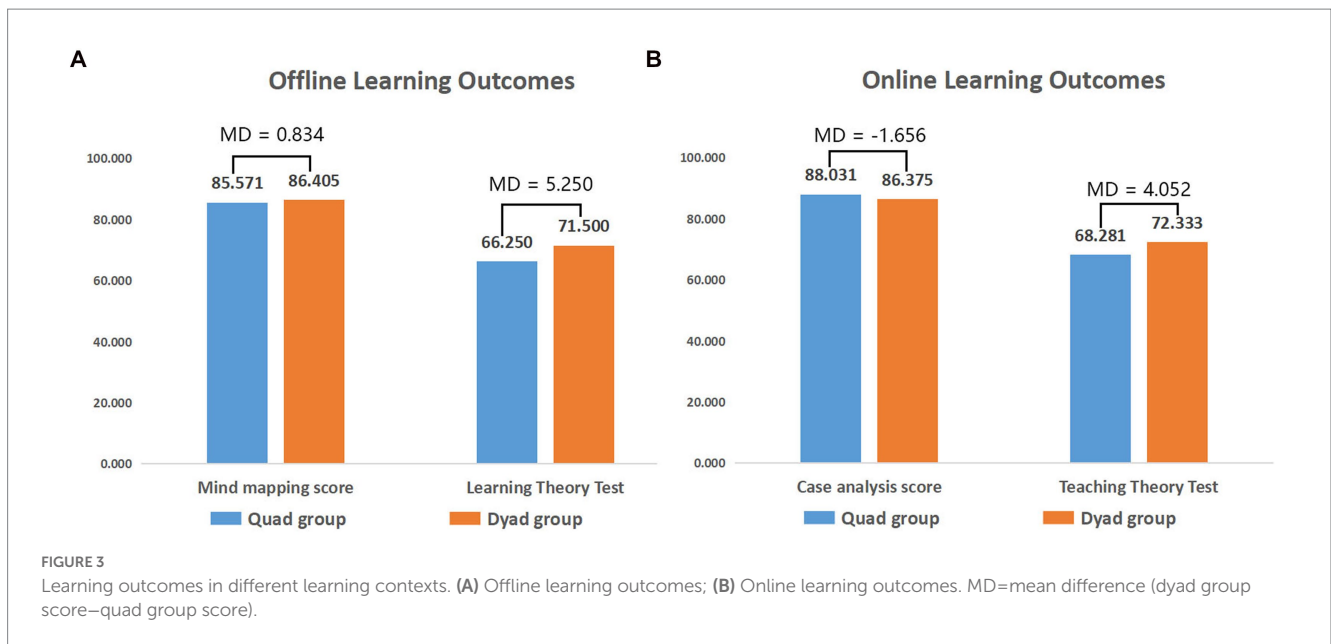
3.4. Data analysis

Two rounds of one-way ANOVA were conducted to investigate the differences in participants' collaborative learning outcomes, leaning engagement, and collaborative experience between quad and dyad groups in both offline and online learning contexts. The Kolmogorov–Smirnov test and Levene's test were conducted to ensure that the statistical assumptions of normality and homogeneity were met for ANOVA. Quantitative data were analyzed using IBM SPSS (version 23). Interview data were analyzed qualitatively to extract themed findings that would enable triangulation and meaningful interpretation of the statistical results.

4. Results

4.1. Learning outcomes

Students' learning outcomes were reflected by the average scores for the collaborative learning assignments (full score 100) and post-tests (full score 100). As shown in [Figure 3](#), the offline collaborative learning assignment was the production of a mind map, and the post-tests covered learning theory; the online collaborative learning assignments and post-tests were the online case analysis reports and teaching theory tests, respectively. The data in the figure illustrate that there seems to be no difference in the learning outcomes between the quad and dyad groups, in both online and offline learning. Although there was a gap of 5.25 and 4.052 in the post-tests—that is, the scores



for the dyad group were slightly higher—the *t*-test results showed that there was no significant difference. From the figure, we can also see that the quad group showed improvement in the online learning condition in terms of both the collaborative learning assignments and post-tests, but there was little difference in the scores of the dyad groups. It is possible that there were, however, more changes in the learning process; as one dyad collaborative student noted: “In fact, the final results may not be much different, but the process may be different.”

4.2. Learning engagement

We compared the learning engagement of students with different collaboration modes in the learning process and plotted the average scores in Figure 4. The difference was statistically significant and marked with an asterisk. The information in the figure indicates that there is little difference in learning engagement between students in the quad and dyad groups. However, after further analysis, as shown in Figure 4B, we found that although the dyad groups had fewer group members, they reported more communication and mutual assistance with peers during the offline face-to-face collaborative learning process than the quad groups ($MD=0.332$, $F=3.078$, $p=0.014<0.05$). This indicates that increasing the number of collaborators does not necessarily lead to more communication and exchange within the group. This also explains, to some extent, the reason why the dyads performed better in comprehensiveness ($MD=4.033$) in the production of the mind map for the offline collaborative learning assignments. In online learning, although the results showed no significant difference, the dyads still had more communication and mutual assistance than the quad groups ($MD=0.126$).

4.3. Collaborative experience

Figure 5 shows the average scores for the collaborative experience of students in different-sized groups in the learning process. In general, participants had a positive attitude toward collaborative

learning regardless of the size of their collaboration groups, as there was no significant difference in their perceived collaborative experience. Most participants believed that collaborative learning could promote understanding of course content (Item 33, $M=4.270$), stimulate their own interest in learning (Item 34, $M=4.035$), and expressed their willingness to participate in collaborative learning if they had the opportunity (Item 36, $M=4.040$). Interestingly, the item with the lowest score in the two questionnaires was students' interest in collaborative pedagogy (Item 20, $M=3.280$), which indicates that although participants had a positive attitude toward collaborative learning, it is still not very attractive to them.

4.4. Difference between offline and online collaboration

A change in learning context inevitably leads to differences between different-sized groups. We used a series of paired sample *t*-tests to further compare the data collected from the quad and dyad groups during the transition from the offline to online context. As shown in Figure 6, the change in learning context had little influence on the test scores and assignment scores for both the quad and dyad groups, as evidenced by the flat lines in the line chart. However, Figure 7 shows that the learning context had a moderating effect on the learning engagement of the participants from the quad groups, with online context associated with higher learning engagement. Thus far, the dyad groups witnessed no such moderating effect. In general, dyads were more stable and less likely to be influenced by the mode of learning context, while quad groups might be more suitable for online collaboration due to increased collaborative experience in such context.

We further examined participants' learning engagement in terms of six subscales, namely active learning, collaboration performance, self-management, peer interaction, deep learning, and affective engagement, and compared their variance between the offline and online collaboration contexts. Figure 8 shows that the quad groups demonstrated greater affective engagement in the online collaboration

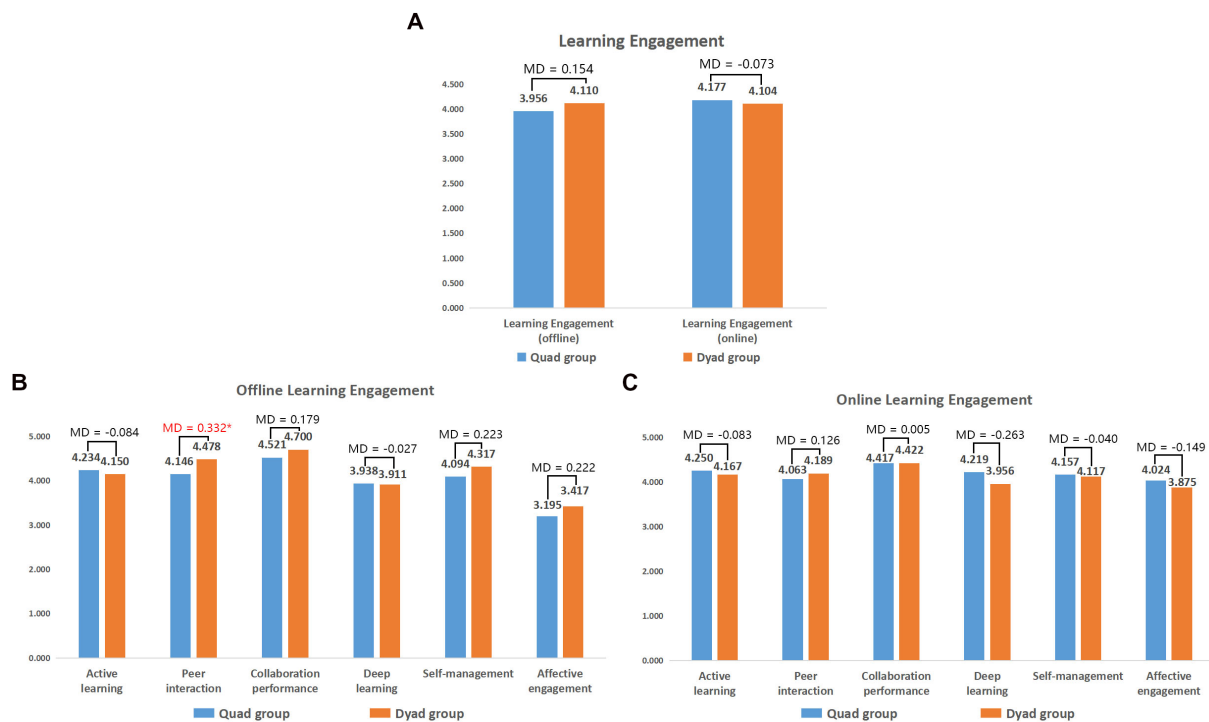


FIGURE 4 Learning engagement in different learning contexts. (A) Total score of offline and online learning engagement; (B) Offline learning engagement; (C) Online learning engagement. MD, mean difference (dyad group score–quad group score); * $p < 0.05$; ** $p < 0.01$.

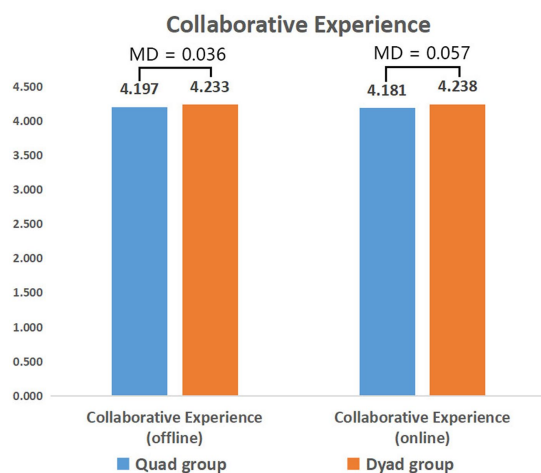


FIGURE 5 Collaborative experience in different learning context. MD=mean difference (dyad group score–quad group score); * $p < 0.05$; ** $p < 0.01$.

context ($MD = 0.829$, $p = 0.000 < 0.01$) than in the offline context, and a slight improvement was found in deep learning after the online transition ($MD = 0.281$, $p = 0.024 < 0.05$). Similarly, the dyads reported a significant increase in affective engagement after switching to the online context ($MD = 0.458$, $p = 0.010 < 0.05$). However, perceived peer interaction ($p = 0.001 < 0.01$), collaboration performance ($p = 0.004 < 0.01$), and self-management ($p = 0.043 < 0.05$) witnessed a

moderate decline. The results indicate that although the overall learning engagement score of dyads remained steady across the two learning contexts, variance existed in the subscales of engagement.

5. Discussion and conclusion

In this study, we explored the impact of changes in group size and learning context on collaborative learning, and the impact was reflected through learning outcomes, learning engagement, and collaborative experience. In general, in the same learning context, there was little difference in learning outcomes, learning engagement, and collaborative experience between quad and dyad groups. However, compared to the quad groups, dyads experienced few changes when switching contexts, which indicates that it is a more stable grouping strategy for collaborative learning.

The learning outcomes were primarily measured by collaborative tasks and post-tests. In terms of scores, except for the online case analysis report, the dyad group scored slightly higher, contrary to the findings of Shaw (2013) and Pfister and Oehl (2009). This indicates that, although the dyads were at a disadvantage in terms of numbers, they still achieved relatively better scores. One possible explanation for this is that, in the process of collaboration, each member of the dyad had more responsibility for group work. To better complete the learning task and obtain higher performance, they may have had stronger motivation, more active input, and greater participation in the process of collaboration, thus resulting in better learning performance. Contrarily, due to the increase in the number of members, the connection between individual contributions and final

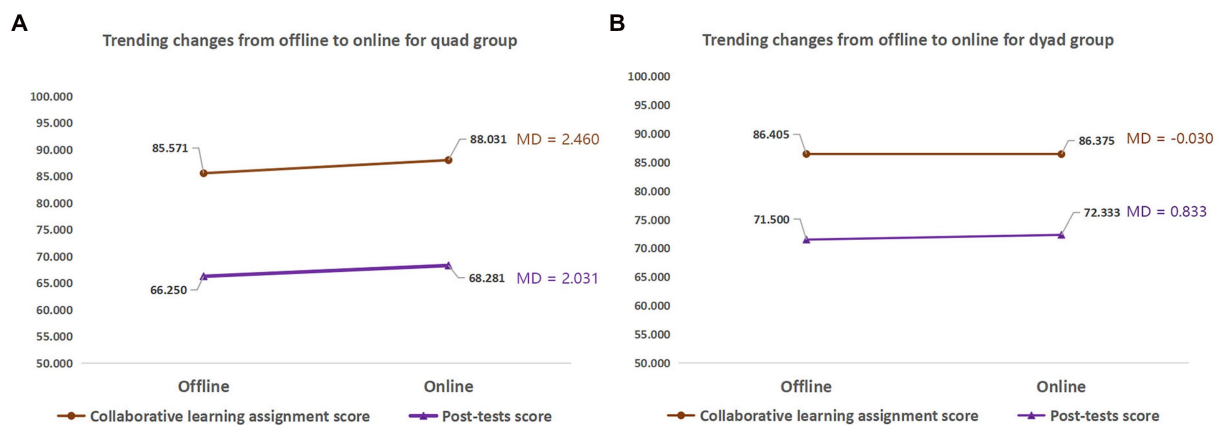


FIGURE 6

Trends of different sized groups in learning outcomes. (A) Quad groups; (B) Dyads. MD, mean difference (online score–offline score); * $p < 0.05$; ** $p < 0.01$.

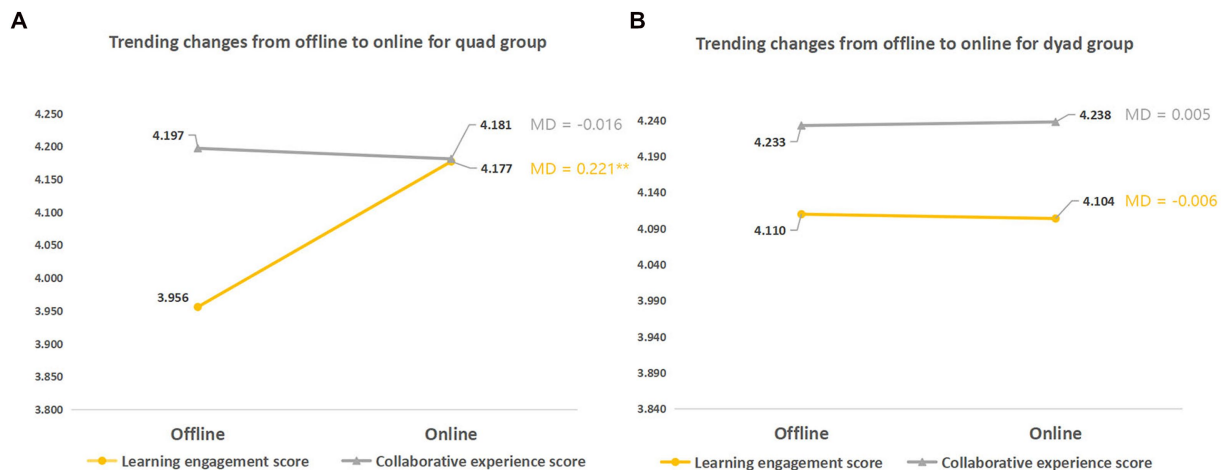


FIGURE 7

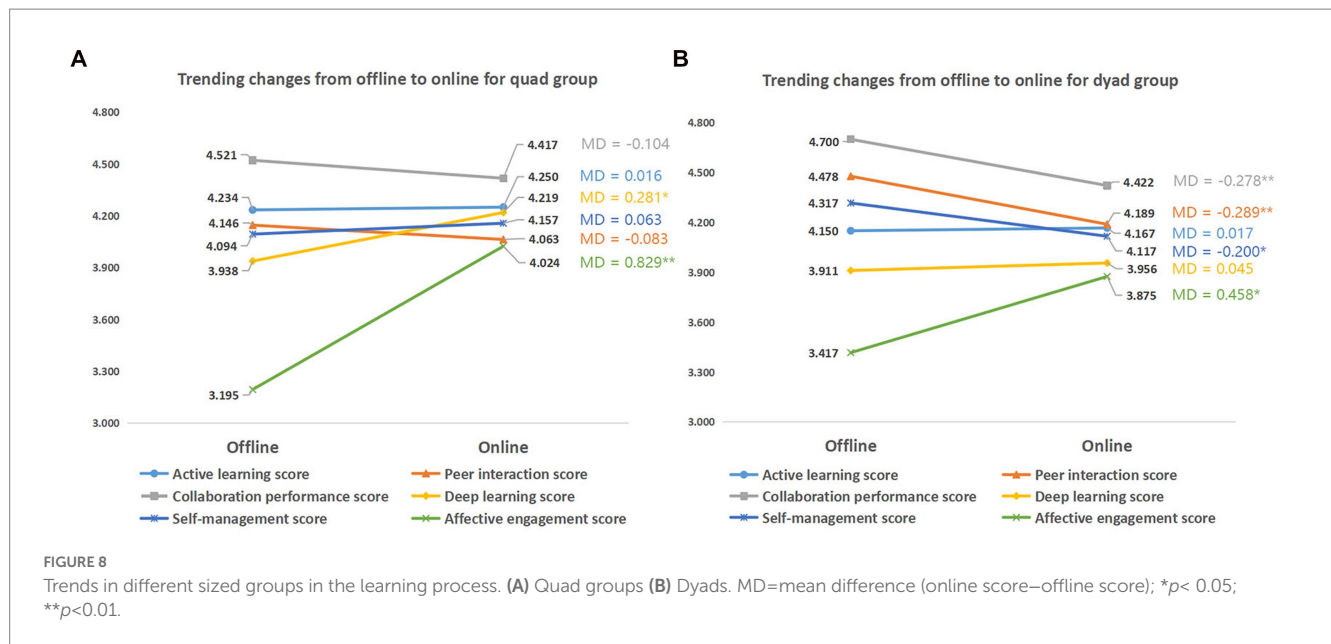
Trends of different sized groups in learning engagement and collaboration experience. (A) Quad groups; (B) Dyads. MD=mean difference (online score–offline score); * $p < 0.05$; ** $p < 0.01$.

results in the whole process of collaboration may have been reduced, and each member's responsibility was relatively small, which may lead to isolation and inactivity for some members (Saqr et al., 2019), as well as the phenomenon of social loafing (Aggarwal and O'Brien, 2008). The participants in the quad groups who were interviewed occasionally mentioned that a few individual members of the group did not contribute to the team's efforts, nor did they better master the learning knowledge.

For learning engagement, we found that in offline face-to-face collaborative learning, the dyads showed more active interaction than the quad groups, which was consistent with the conclusions of Siampou et al. (2014) and van der Meijden and Veenman (2005). However, in the online collaborative learning process, there was no difference between the quad and dyad groups, which may be related to the unique benefits of the different learning context (Lee and Tsai, 2011). Offline learning provides a real environment, and peers need to communicate face to face in real time, which enable rich

peer expressions, gestures, social relations, and other factors to positively affect the output and expression of personal views. In contrast, online learning generates a sense of distance; with the right level of social distance, peers are more likely to express their real ideas and regulate relationships within the group for the common goal.

To our surprise, the collaborative experience did not differ among dyad and quad groups in either the face-to-face or online contexts, contrary to the findings of Lai and Wong (2021). This shows that the change in learning context and group size had little impact on learners' cognition and emotion from collaboration, and high satisfaction was maintained. There are three possible reasons for this. First, strong learning motivation might overcome the differences in collaborative experience caused by group size and learning context (Deci et al., 1991; Ryan and Deci, 2000). If students are interested in collaborative tasks and collaborative content, they are willing to learn whether the group size is four or two, or whether the learning context is offline or online.



Second, cultural differences may also have played a part in these results (Gyasi et al., 2021); Chinese students may not have been willing to express their unhappy experiences in the questionnaire due to the influence of their own learning atmosphere. Third, the data collected by the questionnaire were in the form of self-report, which may not have been very accurate, which may make it difficult to determine the differences in the details of students' collaborative experience.

One interesting finding of this study is that, compared with the quad groups, the dyads were relatively stable in terms of learning outcomes, learning engagement, and collaborative experience, with little changes in their scores. It seems justified to assume that the dyad group was more adaptable to the change in learning context. One possible explanation for this is that the relationship between the two group members is more equal, resulting in greater intrinsic learning motivation that was less susceptible to contextual factors (Xie and Ke, 2011). The fact that online learning context promoted learning engagement for the quad groups is consistent with the existing literature; since the online context can afford students more autonomy and flexibility while reducing social anxiety, it can promote cognitive and affective engagement for larger groups (Walsh et al., 2003; Almahasees et al., 2021).

5.1. Implications for teaching practice

Based on the results of the research, this paper puts forward three suggestions for implementing collaborative learning. First, for teaching large classes with a large number of students, we suggest that teachers choose groups of four for collaborative learning. Because the effects of the quad and dyad groups remained the same across different contexts, the advantages of the quad group in management efficiency could reduce teacher workload. Second, if a collaborative learning activity is primarily offline and the focus is on promoting social interaction and dialogue, then teachers are encouraged to choose dyads instead of quad groups as the former can encourage greater peer interaction due to strong real-time offline communication and equal

relationships during collaboration. Finally, for the developers of online learning platforms, it is very necessary to develop a function that allows teachers to flexibly select group size and groupings according to specific teaching conditions, class size, learning tasks, and other factors.

5.2. Limitations and future research

Several limitations of this study need to be noted. First, the study was conducted for a short time (4 weeks), and students may have entered the next learning situation without fully grasping the characteristics of the current learning context. Therefore, in the follow-up research, the time frame for collaborative learning could be extended in the different learning contexts to improve students' sense of the context and ensure the accuracy of the collected data. Second, the participants in this study came from the same course offered by one university and were all third-year students. Future research should improve the universality of the findings by improving the diversity of the participant groups, collaborative learning tasks, and disciplinary domains, and investigate the impact of demographics such as learners' gender, age, academic achievements, and different learning fields on the results, and experience of collaborative learning. Third, we only collected the end results of collaborative learning in this study without in-depth analysis of the learning data generated during the learning process. Future researchers can report more information of the processes leading to the outcomes, which might provide rationale for the optimal group size. Finally, the data collected in this study were primarily quantitative data lacking in diversity, such as questionnaires and evaluation scores. However, it is difficult to triangulate and explain the statistical results without in-depth analysis of, among other aspects, learning behaviors, emotions, and attitudes in the different collaborative groups. Therefore, we suggest that qualitative surveys be used in the future to improve the credibility and interpretability of the results.

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 at: <https://doi.org/10.17632/6x56rzy2wt.1>.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board of Central China Normal University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HL: conceptualization and funding acquisition. LJ and HL: methodology, writing—review and editing, and supervision. MW and HL: formal analysis and investigation. MW: writing—original draft preparation and visualization. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Hubei Provincial Teaching and Research Project for Higher Education, Hubei, China, grant number 2021085; Teacher Education Specialized Grant of Central China Normal University, grant number CCNUTEIII 2021-10; and the “AI + Education” Innovative Teaching Research Grant of Central China Normal University, grant number 2022XY017.

References

- Abuseileek, A. F. (2012). The effect of computer-assisted cooperative learning methods and group size on the EFL students' achievement in communication skills. *Comput. Educ.* 58, 231–239. doi: 10.1016/j.compedu.2011.07.011
- Aggarwal, P., and O'Brien, C. L. (2008). Social loafing on group projects. *J. Mark. Educ.* 30, 255–264. doi: 10.1177/0273475308322283
- Alexopoulou, E., and Driver, R. (1996). Small-group discussion in physics: peer interaction modes in pairs and fours. *J. Res. Sci. Teach.* 33, 1099–1114. doi: 10.1002/(sici)1098-2736(199612)33:10<1099::aid-tea4>3.0.co;2-n
- Almahasees, Z., Mohsen, K., and Amin, M. O. (2021). Faculty's and students' perceptions of online learning during COVID-19. *Front. Educ.* 6:638470. doi: 10.3389/feduc.2021.638470
- Amichai-Hamburger, Y., Gazit, T., Bar-Ilan, J., Perez, O., Aharoni, N., Bronstein, J., et al. (2016). Psychological factors behind the lack of participation in online discussions. *Comput. Hum. Behav.* 55, 268–277. doi: 10.1016/j.chb.2015.09.009
- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *J. Learn. Sci.* 9, 403–436. doi: 10.1207/s15327809jls0904_2
- Benton, S. L., Li, D., Brown, R., Guo, M., and Sullivan, P. (2015). “Revising the IDEA student ratings of instruction system 2002–2011 data,” in *IDEA Technical Report No. 18*. IDEA Center, Inc. Available at: <https://eric.ed.gov/?id=ED573673>
- Bianco, A. S. (2014). Starting and teaching: Basic robotics in the classroom: Modern, engaging engineering in technology education. *Technol. Eng. Teach.* 73, 32–38.
- Burgess, A. W., McGregor, D. M., and Mellis, C. M. (2014). Applying established guidelines to team-based learning programs in medical schools: a systematic review. *Acad. Med.* 89, 678–688. doi: 10.1097/acm.0000000000000162
- Chen, J., Wang, M., Kirschner, P. A., and Tsai, C.-C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: a meta-analysis. *Rev. Educ. Res.* 88, 799–843. doi: 10.3102/0034654318791584
- Corrêge, J.-B., and Michinov, N. (2021). Group size and peer learning: peer discussions in different group size influence learning in a biology exercise performed on a tablet with stylus. *Front. Educ.* 6:733663. doi: 10.3389/feduc.2021.733663
- Deci, E., Vallerand, R., Pelletier, L., and Ryan, R. (1991). Motivation and education: the self-determination perspective. *Educ. Psychol.* 26, 325–346. doi: 10.1207/s15326985ep2603&4_6
- Dewiyanti, S., Brand-Gruwel, S., Jochems, W., and Broers, N. J. (2007). Students' experiences with collaborative learning in asynchronous computer-supported collaborative learning environments. *Comput. Hum. Behav.* 23, 496–514. doi: 10.1016/j.chb.2004.10.021
- Dillenbourg, P. (ed.) (1999). “What do you mean by 'collaborative learning'?” in *Collaborative-Learning: Cognitive and Computational Approaches* (Oxford: Elsevier), 1–15.
- Dugosh, K. L., Paulus, P. B., Roland, E. J., and Yang, H.-C. (2000). Cognitive stimulation in brainstorming. *J. Pers. Soc. Psychol.* 79, 722–735. doi: 10.1037/0022-3514.79.5.722
- Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059
- Gopinathan, S., Kaur, A. H., Veeraya, S., and Raman, M. (2022). The role of digital collaboration in student engagement towards enhancing student participation during COVID-19. *Sustain. For.* 14:6844. doi: 10.3390/su14116844
- Gu, X., Shao, Y., Guo, X., and Lim, C. P. (2015). Designing a role structure to engage students in computer-supported collaborative learning. *Internet High. Educ.* 24, 13–20. doi: 10.1016/j.iheduc.2014.09.002
- Gunuc, S., and Kuzu, A. (2015). Student engagement scale: development, reliability and validity. *Assess. Eval. High. Educ.* 40, 587–610. doi: 10.1080/02602938.2014.938019
- Gyasi, J. F., Zheng, L., and Long, M. (2021). Reflecting on the past to shape the future: a systematic review on cross-cultural collaborative learning from 2011 to 2020. *Sustain. For.* 13:13890. doi: 10.3390/su132413890
- Hmelo-Silver, C. E., and Jeong, H. (2021). Benefits and challenges of interdisciplinarity in CSCL research: a view from the literature. *Front. Psychol.* 11:579986. doi: 10.3389/fpsyg.2020.579986

Acknowledgments

The authors acknowledge the invaluable contributions of Zhili Zhang, Panpan Duan, Xiang Li, and Yanan Wang from Central China Normal University for their assistance in data collection. The authors also want to thank the participated students for their support in this study.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1168208/full#supplementary-material>

- Ibrahim, S.N.K.A., and Harun, J. (2015). "A meta-analysis on students' argumentative knowledge construction process in social collaborative learning environment" in 2015 International Conference on Learning and Teaching in Computing and Engineering. Taiwan: IEEE, 134–141.
- Jensen, M. S., and Wiley, J. (2006). When three heads are better than two. *Proc. Annu. Meet. Cog. Sci. Soc.* 28, 2375–2380.
- Kaldırım, A., and Tavşanlı, Ö. F. (2018). The effect of collaborative learning approach on students' academic achievement in Turkish courses in Turkey: a meta-analysis study. *Eğitim Bilim* 43, 185–205. doi: 10.15390/eb.2018.7553
- Kim, J. (2013). Influence of group size on students' participation in online discussion forums. *Comput. Educ.* 62, 123–129. doi: 10.1016/j.compedu.2012.10.025
- Kim, N. J., Belland, B. R., Lefler, M., Andreasen, L., Walker, A., and Axelrod, D. (2020). Computer-based scaffolding targeting individual versus groups in problem-centered instruction for STEM education: meta-analysis. *Educ. Psychol. Rev.* 32, 415–461. doi: 10.1007/s10648-019-09502-3
- Kooloos, J. G. M., Klaassen, T., Vereijken, M., Van Kuppeveld, S., Bolhuis, S., and Vorstenbosch, M. (2011). Collaborative group work: effects of group size and assignment structure on learning gain, student satisfaction and perceived participation. *Med. Teach.* 33, 983–988. doi: 10.3109/0142159x.2011.588733
- Laal, M., and Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia Soc. Behav. Sci.* 31, 486–490. doi: 10.1016/j.sbspro.2011.12.091
- Lai, X., and Wong, G. K. W. (2021). Collaborative versus individual problem solving in computational thinking through programming: a meta-analysis. *Br. J. Educ. Technol.* 53, 150–170. doi: 10.1111/bjet.13157
- Lee, S. W.-Y., and Tsai, C.-C. (2011). Students' perceptions of collaboration, self-regulated learning, and information seeking in the context of internet-based learning and traditional learning. *Comput. Hum. Behav.* 27, 905–914. doi: 10.1016/j.chb.2010.11.016
- Liang, H.-Y., Hsu, T.-Y., Hwang, G.-J., Chang, S.-C., and Chu, H.-C. (2021). A mandatory contribution-based collaborative gaming approach to enhancing students' collaborative learning outcomes in science museums. *Interact. Learn. Environ.* 1–15, 1–15. doi: 10.1080/10494820.2021.1897845
- Liu, Y.-H., Kwon, K., and Johnson, L. P. (2017). Exploration of factors in the early collaboration phase affecting virtual groups' overall collaborative learning experiences. *J. Educ. Comput. Res.* 56, 485–512. doi: 10.1177/0735633117715034
- Lohman, M. C., and Finkelstein, M. (2000). Designing groups in problem-based learning to promote problem-solving skill and self-directedness. *Instr. Sci.* 28, 291–307. doi: 10.1023/a:1003927228005
- Luo, H., Chen, Y., Chen, T., Koszalka, T. A., and Feng, Q. (2023). Impact of role assignment and group size on asynchronous online discussion: an experimental study. *Comput. Educ.* 192:104658. doi: 10.1016/j.compedu.2022.104658
- Ma, J., and Zhou, Q. (2019). An empirical study on the dimensions and factors of college student engagement in blended learning environment. *Educ. Res. Dev.* 39:12.
- Miao, J., Luo, S. F., and Wang, Q. (2017). Exploring collaborative learning experience in a teacher professional development MOOC. *Open Educ. Res.* 23:7. doi: 10.13966/j.cnki.kfjyyj.2017.06.009
- Parks-Stamm, E. J., Zafonte, M., and Palenque, S. M. (2017). The effects of instructor participation and class size on student participation in an online class discussion forum. *British J. Educ. Technol.* 48, 1250–1259. doi: 10.1111/bjet.12512
- Peltokorpi, J., and Niemi, E. (2019). Effects of group size and learning on manual assembly performance: an experimental study. *Int. J. Prod. Res.* 57, 452–469. doi: 10.1080/00207543.2018.1444810
- Pfister, H. R., and Oehl, M. (2009). The impact of goal focus, task type and group size on synchronous net-based collaborative learning discourses. *J. Comput. Assist. Learn.* 25, 161–176. doi: 10.1111/j.1365-2729.2008.00287.x
- Qiu, M., and McDougall, D. (2015). Influence of group configuration on online discourse reading. *Comput. Educ.* 87, 151–165. doi: 10.1016/j.compedu.2015.04.006
- Rannastu, M., Siiman, L. A., Mäeots, M., Pedaste, M., and Leijen, Ä. (2019). "Does group size affect students' inquiry and collaboration in using computer-based asymmetric collaborative simulations?" in *Advances in Web-Based Learning—ICWL 2019*. eds. M. A. Herzog, Z. Kubincová, P. Han and M. Temperini (Cham: Springer), 143–154.
- Ryan, R. M., and Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Educ. Psychol.* 25, 54–67. doi: 10.1006/ceps.1999.1020
- Saqr, M., Nouri, J., and Jormanainen, I. (2019). "A learning analytics study of the effect of group size on social dynamics and performance in online collaborative learning" in *Transforming Learning with Meaningful Technologies*. eds. M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou and J. Schneider (Cham: Springer), 466–479.
- Segundo Marcos, R. I., López Fernández, V., Daza González, M. T., and Phillips-Silver, J. (2020). Promoting children's creative thinking through reading and writing in a cooperative learning classroom. *Think. Skills Creat.* 36:100663. doi: 10.1016/j.tsc.2020.100663
- Shaw, R.-S. (2013). The relationships among group size, participation, and performance of programming language learning supported with online forums. *Comput. Educ.* 62, 196–207. doi: 10.1016/j.compedu.2012.11.001
- Shen, C.-Y., and Wu, C.-H. (2011). An exploration of students' participation, learning process, and learning outcomes in web 2.0 computer supported collaborative learning. *Int. J. Online Pedagog. Course Des.* 1, 60–72. doi: 10.4018/ijopcd.2011040105
- Shimazoe, J., and Aldrich, H. (2010). Group work can be gratifying: understanding & overcoming resistance to cooperative learning. *Coll. Teach.* 58, 52–57. doi: 10.1080/87567550903418594
- Siampou, F., Komis, V., and Tselios, N. (2014). Online versus face-to-face collaboration in the context of a computer-supported modeling task. *Comput. Hum. Behav.* 37, 369–376. doi: 10.1016/j.chb.2014.04.032
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: a best-evidence synthesis. *Rev. Educ. Res.* 57, 293–336. doi: 10.3102/00346543057003293
- Strauß, S., and Rummel, N. (2020). Promoting interaction in online distance education: designing, implementing and supporting collaborative learning. *Inf. Learn. Sci.* 121, 251–260. doi: 10.1108/ils-04-2020-0090
- Sugai, M., Horita, T., and Wada, Y. (2018). "Identifying optimal group size for collaborative argumentation using SNS for educational purposes" in *2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI)*. Yonago: IEEE, 226–231.
- Sung, Y.-T., Yang, J.-M., and Lee, H.-Y. (2017). The effects of mobile-computer-supported collaborative learning: meta-analysis and critical synthesis. *Rev. Educ. Res.* 87, 768–805. doi: 10.3102/0034654317704307
- Talan, T. (2021). The effect of computer-supported collaborative learning on academic achievement: a meta-analysis study. *Int. J. Educ. Math. Sci. Technol.* 9, 426–448. doi: 10.46328/ijemst.1243
- van der Meijden, H., and Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Comput. Hum. Behav.* 21, 831–859. doi: 10.3390/sym9080141
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Walsh, S. L., Gregory, E., Lake, Y., and Gunawardena, C. N. (2003). Self-construal, facework, and conflict styles among cultures in online learning environments. *Educ. Technol. Res. Dev.* 51, 113–121. doi: 10.1007/bf02504548
- Wang, Y.-H., and Liao, H.-C. (2017). Learning performance enhancement using computer-assisted language learning by collaborative learning groups. *Symmetry* 9:141. doi: 10.3390/sym9080141
- Wang, M., Sun, B., Le, H., and Guo, W. (2020). Does group size influences middle school students' interaction in online forums? *J. Comput. Educ.* 7, 29–49. doi: 10.1007/s40692-019-00141-9
- Wu, J. (2018). Application of computer aided collaborative learning model in english virtual electronic teaching. *Int. J. Emerg. Technol. Learn.* 13, 137–148. doi: 10.3991/ijet.v13i07.8801
- Xie, K., and Ke, F. (2011). The role of students' motivation in peer-moderated asynchronous online discussions. *Br. J. Educ. Technol.* 42, 916–930. doi: 10.1111/j.1467-8535.2010.01140.x
- Yang, T., Luo, H., and Sun, D. (2022). Investigating the combined effects of group size and group composition in online discussion. *Act. Learn. High. Educ.* 23, 115–128. doi: 10.1177/1469787420938524
- Zhan, Z., He, G., Li, T., He, L., and Xiang, S. (2022). Effect of groups size on students' learning achievement, motivation, cognitive load, collaborative problem-solving quality, and in-class interaction in an introductory AI course. *J. Comput. Assist. Learn.* 38, 1807–1818. doi: 10.1111/jcal.12722



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Malgorzata Siekanska,
University of Physical Education in Krakow,
Poland

Mark Urtel,
Indiana University, Purdue University
Indianapolis, United States

*CORRESPONDENCE

Mustafa Levent Ince
✉ mince@metu.edu.tr

RECEIVED 16 February 2023

ACCEPTED 02 May 2023

PUBLISHED 24 May 2023

CITATION

Kılıç K and Ince ML (2023) Perceived use and
value of reproductive, problem-solving, and
athlete-initiated teaching by coaches and
athletes.

Front. Psychol. 14:1167412.

doi: 10.3389/fpsyg.2023.1167412

COPYRIGHT

© 2023 Kılıç and Ince. 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.

Perceived use and value of reproductive, problem-solving, and athlete-initiated teaching by coaches and athletes

Koray Kılıç¹ and Mustafa Levent Ince^{2*}

¹Faculty of Sport Sciences, Kırşehir Ahi Evran University, Kırşehir, Türkiye, ²Physical Education and Sports
Department, Faculty of Education, Middle East Technical University, Ankara, Türkiye

In the sports coaching environment, it is recognized that developing athletes' autonomy and problem-solving skills are crucial to support holistic development and ensure optimal performance. However, there needs to be more information on how coaches use and value different teaching methods in training and how athletes perceive and value these methods. This study aimed to examine coaches' and athletes' perceptions of the use and value of reproductive, productive problem-solving, and productive athlete-initiated teaching methods. To this end, the Coaches' Use of Teaching Methods Scale which is validated for the use of coaches and athletes, was applied to 70 coaches and their 294 athletes of youth sports teams purposefully selected from four cities in Türkiye. Data were analyzed by nonparametric methods, including Friedman's and Mann-Whitney tests ($p < 0.05$). Although there were statistically significant differences between the responses of coaches and athletes regarding the use of different teaching methods in their training and the value they gave to these methods, both groups marked the frequent use of reproductive, occasional use of productive problem-solving and rare use of productive athlete-initiated teaching methods during training. The value given to productive athlete-initiated teaching methods in terms of enjoyment, learning, and motivation by the athletes was higher than the value given to them by the coaches. The study's findings strongly indicate the coaches' professional needs in their pedagogical knowledge, specifically on their value perceptions of productive problem-solving and productive athlete-initiated teaching methods and the capacity to apply them.

KEYWORDS

coaching, teaching methods, reproductive, problem-solving, athlete-initiated

1. Introduction

The primary indicator of effective coaching is athletes' learning and its positive effect on their development in sport (Côté and Gilbert, 2009). Prominent athlete development models such as Long-Term Athlete Development (Balyi et al., 2013) and the Developmental Model of Sport Participation (the DMSP; Côté, 1999; Côté et al., 2007) emphasize biological age and maturation-dependent sport-specific competence (e.g., technique and fitness), and psychosocial skills such as athletes' confidence and character development. These developmental frameworks provide coaches with conceptual guidance for coaching aims, content identification, implementation, and measurement and evaluation of their practice outcomes to improve athletes' learning and performance at a sustainable level (e.g., Lauer and Dieffenbach, 2013).

Based on the coaching aims, setting the objectives, identifying the content, implementing training, and measuring and evaluating the outcomes should be well aligned to ensure athletes' optimum development (Bompa and Haff, 2009; Lara-Bercial et al., 2013). Studies examining the aims of contemporary coaching and the appropriate content to reach those aims stress the holistic development of athletes in all developmental areas and using relevant content during the training (sport-specific technique, fitness, tactics, psychosocial, and emotional development; Fraser-Thomas et al., 2005; Côté et al., 2010). In coaching and teaching literature, as the athletes' performance develops, a need for higher-order learning in all developmental aspects of athletes with the recognition of their individual differences is underlined (e.g., Wikeley and Bullock, 2006; Cushion and Nelson, 2013). Specifically, athletes should develop deep learning capacity by having more opportunities for application, synthesis, evaluation, and creation in the cognitive domain (Thorpe, 1997; Potrac and Cassidy, 2006; Cassidy et al., 2008), adaptation and organization in the psychomotor domain, and characterization by a value or value complex in the affective domain. Athletes' superficial learning in these developmental areas may not be sufficient to realize the current training aims and improve sports performance significantly (Wikeley and Bullock, 2006, p. 18).

Presenting the content (how to teach/instruct) is a critical dimension of coaching knowledge (Gilbert and Côté, 2013), which enables coaches' training implementation. Coaches' instructions may be beneficial or harmful to athletes' development depending on the extent that they are informed about the optimal strategies for athletes' learning and performance needs (Gearity, 2012). Previous work on instructional aspects of coaching naturally depends on the theories of learning and instruction (e.g., Kidman, 2005; Armour, 2010). Accordingly, the intended learning in sport occurs most effectively when the learning intention and instructional strategies or teaching methods are correctly matched. Coaches need to be aware of the consequences of their instructional practices for coaching effectiveness (e.g., Cassidy et al., 2008, p. 33).

For this reason, it is essential for coaches to understand the theories of learning (how athletes learn) and instruction (how coaches should teach) and to use them effectively in practice (Light, 2008; Roberts and Potrac, 2014). Whenever the coaching focuses on ensuring higher-order learning, independent of coaching context (Cassidy et al., 2008), direct teaching methods (reproductive: replication of a known model) would not be sufficient; coaches should also use productive teaching methods (production of knowledge and skills new to the athlete and/or coach), including productive problem-solving and productive athlete-initiated approaches in their training to facilitate deep learning (Wikeley and Bullock, 2006; Ertmer and Newby, 2013; Pill et al., 2021). That brings about the necessity of coaches' regular assessment of athletes' higher-order learning in cognitive, psychomotor, and affective domains due to coaching practices. The current study focuses on the perceived use of various instructional strategies and teaching methods during training by coaches and athletes. The following paragraphs in this section will synthesize the related literature to create a rationale for the study.

In the sports education literature, teaching methods are usually characterized by Mosston and Ashworth's (2008) spectrum of teaching styles. Mosston and Ashworth's Spectrum of Teaching Styles identify 11 different teaching methods in a continuum from teacher-centered to learner-centered: A: Command, B: Practice, C: Reciprocal, D:

Self-Check, E: Inclusion, F: Guided Discovery, G: Convergent Discovery, H: Divergent Discovery, I: Learner Designed Individual Program, J: Athlete-initiated, and K: Self Teaching. Teachers' dominance of decision-making processes in planning, implementation, and measurement and evaluation of an outcome in the teaching process decreases as teaching styles progress from A to K. Styles from A to E use a direct teaching approach. Therefore, they are called reproductive methods of teaching in the sports education setting. The styles from E to K are more learner-centered, and the learner's discovery of new information or skills is at their core. Due to their focus on learners' discovery by problem-solving, styles from E to H can be named productive problem-solving (henceforth problem-solving) teaching methods. On the other hand, learner autonomy in the styles from I to K is dominant in planning, implementation, and measurement and evaluation of the outcome. They can be named under the title of productive athlete-initiated (henceforth athlete-initiated) teaching methods (Kilic and Ince, 2017; Pill et al., 2021).

There is limited research on coaches' use of teaching methods in their practices, although in many sport pedagogy textbooks (e.g., Kidman, 2005; Jones, 2006; Cassidy et al., 2008; Mitchell et al., 2020) and relevant work in teaching (e.g., Light, 2008) the importance of teaching methods, especially the use of learner/athlete-centered methods, is highlighted by well-known scholars in the field. Ample research indicates the use and value given to the various teaching methods by physical education teachers in school physical education classes. Those studies in physical education settings reported heavy use of teacher-centered reproductive teaching methods (e.g., command and practice) but occasional use of learner-centered approaches (e.g., problem-solving and athlete-initiated methods; Curtner-Smith et al., 2001; Cothran et al., 2005; Ince and Hunuk, 2010). Interestingly, these studies also reported that physical education teachers highly valued reproductive methods while valuing the learner-initiated teaching methods for learners' learning, motivation, and enjoyment less.

A few pilot studies have been conducted to assess coaches' use of teaching methods so far (Ince and Kilic, 2016; Kilic and Ince, 2017). The findings of this work were similar to that of the school physical education setting. The preliminary results indicated the coaches' tendency to use reproductive methods, especially command and practice, while rarely using problem-solving and athlete-initiated teaching methods.

Studies examining instruction in physical education and coach development settings have mainly focused on teachers' practices (e.g., Ince and Hunuk, 2010) and coaches (Potrac and Jones, 2009; Trudel et al., 2010). While the research on examining coaching behaviors indicated instruction is the most preferred and motivating form of coaching practice by athletes (Gearity, 2012) coaches may lack the knowledge and ability to apply the types of instruction for different learning situations. Consequently, their instruction can be poor in meeting youth athletes' unique learning needs (Gearity, 2012). There needs to be a more profound understanding of coaches' instructional practices and tendencies and, more importantly, athletes' perceptions of coaching practices to improve coaching effectiveness. To our knowledge, how the athletes perceive the specific teaching methods coaches apply needs to be clarified in the coaching literature. Exploring the use of and value given to reproductive, problem-solving, and athlete-initiated teaching methods in coaching settings could provide critical information on the use and preference of

athlete-centered discovery strategies in sports education settings better with the view of both coaches and their athletes. Such information would help design professional development programs for the coaches' professional knowledge of improving youth athletes' higher-order learning in the defined developmental areas.

Considering the rationale mentioned above, this study explores the coaches' and their athletes' perceived use and value of reproductive, problem-solving, and athlete-initiated teaching methods. To this end, the following research questions were asked;

1. What are the perceptions of coaches and athletes about the use of reproductive, problem-solving, and athlete-initiated teaching methods during training?
 - a. Are there any significant differences between the coaches' perceived use of reproductive, problem-solving, and athlete-initiated teaching methods during the training?
 - b. Are there any significant differences in coaches' use of reproductive, problem-solving, and athlete-initiated teaching methods during the training by the athletes' perception?
2. Are there any significant differences between coaches and athletes in value given to reproductive, productive problem-solving, and productive athlete-initiated teaching methods concerning enjoyment, learning, and motivation during training?

2. Materials and methods

2.1. Participants

Coaches ($n=70$) and their athletes ($n=294$) of youth sports teams from Ankara, Istanbul, Bartın, and Kırşehir cities of Türkiye participated in the study. The coaches were between 18 and 50 years of age ($M=32.69$), representing each level of coaching certificate in a 5-level coaching system ($M=2.61$). In Türkiye, the first and second levels of coaching are equal to developmental level certification, while the higher levels represent elite coaching contexts at national and international levels (GSD, 2019). The athletes' ages were between 12 and 18, with an average of 4.90 years of sport experience ($SD=2.52$). The athletes train 4.88 days a week ($SD=1.33$). With respect to gender representation, 9 and 61 out of 70 coaches were women and men, respectively. Out of 294 athletes, 130 of them were women, and 164 of them were men. The representation of low women coaches in the sampling is related to the relatively low representation of women coaches in the study setting. According to Koca (2018, p.142), 20–25% of the coach population in Türkiye are women. Demographics of coaches and athletes based on sports type are presented in Table 1.

2.2. Data collection instruments

A multi-perspective approach was adopted to measure coaches' use and value perceptions of teaching methods. In examining the teaching methods applied during training and to what extent these methods were valued, the two versions of the "Coaches' Use of Teaching Methods Scale" (CUTEMS-Coach/Athlete; Kılıç and İnce,

TABLE 1 Demographic characteristics of participants.

Sport		Coach		Athlete
	<i>n</i>	Mean age (SD)	<i>n</i>	Mean age (SD)
Badminton	9	32.78 (8.32)	44	13.66 (0.68)
Basketball	12	33.08 (9.1)	62	15.08 (1.44)
Box	2	41.50 (10.61)	14	16.36 (1.22)
Judo	2	32.00 (0.00)	8	14.25 (1.75)
Kick box	5	32.40 (5.18)	30	16.20 (1.03)
Soccer	4	35.50 (5.80)	18	14.94 (0.73)
Swimming	4	27.75 (2.1)	25	13.32 (1.03)
Taekwondo	4	29.00 (2.16)	13	14.92 (1.26)
Tennis	7	35.29 (6.10)	32	14.47 (1.19)
Athletics	10	33.50 (6.98)	25	15.64 (1.44)
Volleyball	7	31.00 (6.83)	16	14.31 (1.45)
Wrestling	4	29.75 (5.91)	7	14.43 (1.13)
Gender				
Women	9		130	
Men	61		164	
Total	70		294	

2019, 2021a) were used. The scales for coaches and athletes share the same structure but are worded slightly differently to reflect who is rating. For both versions of the scale, the items were generated based on the adapted form of the "Use of Teaching Styles and Perceptions of Styles Questionnaire" (Kulinna and Cothran, 2003) for the Turkish physical education context by Ince and Hunuk (2010). The scale items include 11 scenarios and four questions answered for each on a 5-point Likert scale (never to always). The 11 scenarios are broken down into three approaches of teaching methods, namely reproductive (items 1–5), problem-solving (items 6–8), and athlete-initiated (items 9–11) for both of the scales (See Table 2 for sample scenarios). The first question of each scenario is to determine the level of 'coaches' use of a teaching method (I train my athletes with this method/My coach trains with this method), and the other three questions are to examine the value perceptions about the teaching method regarding "enjoyment," "learning," and "motivation" (e.g., I think this method will make training fun).

The scales were developed by first conducting exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) on two different data comprised of 275 athletes and 148 athletes from various sports, respectively (Kılıç and İnce, 2019). EFA findings revealed a 3-factor construct in line with the theoretical foundations (i.e., behaviorist, cognitivist, and constructivist approaches) the items fit in, with internal consistency values ranging from 0.72 to 0.81. CFA findings proved the construct validity of the scale subdimensions ($\chi^2/df=1.34$; GFI=0.93; CFI=0.94; TLI=0.92; RMSEA=0.05). Then, the validity and reliability of the scale sub-dimensions for coaches were tested by examining the concurrent validity and internal consistency of 'coaches' and their 'athletes' ratings (Kılıç and İnce, 2021a). Findings showed strong correlations between the ratings of coaches and athletes for each subdimension (reproductive, $r=0.83$, $p<0.01$; problem-solving, $r=0.62$, $p<0.01$; athlete-initiated, $r=0.62$,

$p < 0.01$). Internal consistency values of the examined factors ranged from 0.69 to 0.86.

2.3. Data collection procedures

Before data collection, permission was obtained from the Research Ethics Committee of Middle East Technical University (No: 28620816/154). A purposeful sampling strategy was adopted to reach a highly representative number of competitive youth sports coaches and their athletes and select the coaches and athletes who have worked together for at least 1 year. While primarily aiming at collecting data within the largest cities of Türkiye (Ankara and İstanbul), two other smaller cities of Türkiye (Bartın and Kırşehir) were also included in the data collection to ensure the representativeness of coaches' use of teaching methods in the study setting examined.

The first researcher collected the data by visiting the sport club settings. Coaches and their athletes completed the adapted versions of the scales (CUTEMS-Coach/Athlete). Athletes and coaches completed the scales separately to ensure the trustworthiness of responses. Coaches and athletes rated the scales considering the current coaching practices regarding teaching methods. Coaches and athletes completed the scales in approximately 15 min.

2.4. Data analysis

Data were initially screened for missing cases and matched the representation of coaches and their athletes in the data set. Fifty-six

TABLE 2 Sample scenarios for reproductive, problem-solving, and athlete-initiated teaching.

Scenario samples
<i>Reproductive</i>
1. The coach breaks down the skills into parts and demonstrates the right way to perform the skill. Athletes try to move when and exactly how the coach tells them. The coach provides feedback and the athletes try to emulate the coach's model.
2. The coach makes several stations in the gym where athletes work on different parts of a skill or different skills. Athletes rotate around the stations and do the tasks at their own pace. The coach moves around and helps athletes when needed.
<i>Productive problem-solving</i>
6. The coach asks athletes to discover a solution to a movement problem. The coach asks athletes a series of specific questions and the athletes try out their answers until they discover the right answer that the coach wanted them to discover.
7. Athletes try to learn a skill or concept by using logical reasoning. The coach asks a question and athletes try to reason and think about different solutions. By critically thinking about the question and trying solutions, athletes can discover the single, right answer.
<i>Productive athlete-initiated</i>
10. The athlete decides what will be learned as well as how it will be learned. The coach and athlete set some basic criteria, but the athlete is responsible for all the decisions about how and what to learn. The coach can help with information if the athlete needs it.
11. The athlete decides everything about learning something new. They even decide if they want to involve the coach or not. The coach accepts the athlete's decisions about learning.

Scenarios are adapted from Kulinna and Cothran (2003) by Kılıç and Ince (2019).

athletes from artistic gymnastics and 'women's rugby were eliminated from the data set due to a lack of their coaches' data. Then, the data were examined to meet the normality assumptions. As coaches' and athletes' data on the use of teaching methods and the value given to teaching methods needed to meet the normality assumptions, the analyses were conducted by nonparametric tests. Perceived use of teaching methods data of coaches and athletes were analyzed separately by the Friedman test for the first research question.

Coaches' and athletes' comparison of the value given to teaching methods regarding enjoyment, learning, and motivation were analyzed by the Mann–Whitney U test ($p < 0.05$).

3. Results

Rq1. What are the perceptions of coaches and athletes about the use of reproductive, problem-solving and athlete-initiated teaching methods during training?

Regarding the first research question, according to the descriptive analysis, both coaches and athletes reported the predominant use of reproductive teaching methods during the training. Data from coaches and athletes also indicated the occasional use of problem-solving teaching methods and almost no use of athlete-initiated teaching methods during training (Table 3).

(a). Are there any significant differences between the coaches' perceived use of reproductive, problem-solving, and athlete-initiated teaching methods during the training?

According to Friedman's test, there were significant differences in the coaches' use of reproductive, problem-solving, and athlete-initiated teaching methods during the training, $\chi^2(2) = 140.000$, $p = 0.001$. *Post hoc* analysis with a Bonferroni correction indicated that the mean rank of perceived use of reproductive, problem-solving, and athlete-initiated teaching methods were 3, 2, and 1, respectively. There were significant differences between the use of reproductive and problem-solving ($Z = -5.916$, $p = 0.001$), reproductive and athlete-initiated ($Z = 11.832$, $p = 0.001$), and problem-solving and athlete-initiated ($Z = 5.916$, $p = 0.001$) teaching methods by coaches. Coaches reported the dominant use of reproductive teaching methods and the rare use of athlete-initiated teaching methods during training (Figure 1).

(b). Are there any significant differences in coaches' use of reproductive, problem-solving, and athlete-initiated teaching methods during the training by the athletes' perception?

Friedman's test showed significant differences in coaches' use of reproduction, problem-solving, and athlete-initiated teaching methods during training, according to the athletes' perception,

TABLE 3 Coaches' and their athletes' reports on the use of reproductive, problem solving, and athlete-initiated teaching methods during the training.

Teaching methods	Data source	
	Coach ($n = 70$) Mean (SD)	Athlete ($n = 294$) Mean (SD)
Reproductive*	4.66 (0.30)	3.70 (0.72)
Problem-solving*	3.15 (0.38)	2.78 (0.96)
Athlete-initiated*	1.45 (0.35)	1.67 (0.70)

*Significant differences among the reported use of teaching methods during the training. In a 5 level Likert Scale (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Usually; 5 = Always).

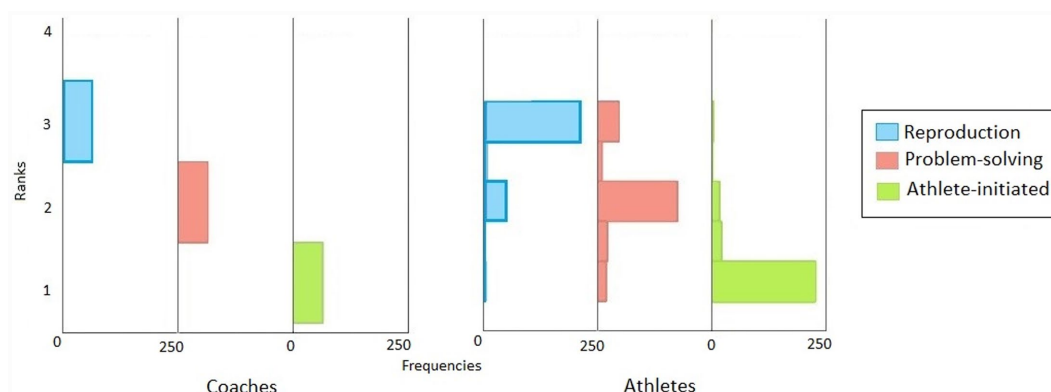


FIGURE 1 Coaches' and their athletes' perception on the use of reproductive, problem-solving and athlete-initiated teaching methods during training.

$\chi^2(2) = 410.870$, $p = 0.001$. *Post hoc* analysis with a Bonferroni correction indicated that the mean rank of perceived use of reproductive, problem-solving, and athlete-initiated teaching methods were 2.78, 2.08, and 1.14, respectively. There were significant differences between the use of reproductive and problem-solving ($Z = 8.433$, $p = 0.001$), reproductive and athlete-initiated ($Z = 19.898$, $p = 0.001$), and problem-solving and athlete-initiated ($Z = 11.465$, $p = 0.001$) teaching methods used during training. The athletes reported their coaches' dominant use of reproductive teaching methods and the rare use of athlete-initiated teaching methods during training (Figure 1).

Rq2. Are there any significant differences between the coaches and the athletes in value given to reproductive, problem-solving and athlete-initiated teaching methods concerning enjoyment, learning and motivation during training?

According to the descriptive data analysis, in terms of enjoyment, learning, and motivation, coaches' and athletes' values given to teaching methods from the highest to the least were reproductive, problem-solving, and athlete-initiated teaching methods, respectively (Table 4).

A Mann-Whitney test indicated that athletes valued reproductive teaching methods in terms of learning (Athletes Mdn = 4.2; Coaches Mdn = 4.0), $U = 7760.5$, $p = 0.001$ and motivation (Athletes Mdn = 4.0; Coaches Mdn = 3.8), $U = 8,230.5$, $p = 0.009$, problem-solving teaching methods in terms of motivation (Athletes Mdn = 4.0; Coaches Mdn = 3.7), $U = 8640.0$, $p = 0.035$, and athlete-initiated teaching methods in terms of enjoyment (Athletes Mdn = 3.3; Coaches Mdn = 2.3), $U = 5453.0$, $p = 0.001$, learning (Athletes Mdn = 3.7; Coaches Mdn = 2.3), $U = 5193.0$, $p = 0.001$ and motivation (Athletes Mdn = 3.7; Coaches Mdn = 2.3), $U = 5079.5$, $p = 0.001$, were significantly higher than those given by the coaches. There was no significant change in the value given to the reproductive teaching methods regarding enjoyment, and problem-solving teaching methods regarding enjoyment and learning between the coaches and athletes ($p > 0.05$; Table 4).

4. Discussion

This study aimed to examine the coaches' and their athletes' perceived use and value given to reproductive, problem-solving, and

TABLE 4 The coaches' and their athletes' value perception on the use of reproductive, problem solving, and athlete-initiated teaching methods regarding enjoyment, learning, and motivation during training.

Teaching methods	Value component	Data source	
		Coach (n = 70) Mean (SD)	Athlete (n = 294) Mean (SD)
Reproductive	Enjoyment	3.69 (0.60)	3.73 (0.73)
	Learning*	3.92 (0.58)	4.16 (0.62)
	Motivation*	3.85 (0.60)	4.04 (0.68)
Problem-solving	Enjoyment	3.50 (0.90)	3.63 (1.39)
	Learning	3.80 (0.88)	3.90 (0.81)
	Motivation*	3.67 (0.89)	3.89 (0.89)
Athlete-initiated	Enjoyment*	2.45 (1.10)	3.41 (1.05)
	Learning*	2.48 (1.14)	3.49 (1.05)
	Motivation*	2.50 (1.21)	3.58 (1.04)

*Significant differences between the coaches and athletes in value given the teaching methods. In a 5 level Likert Scale (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Usually, 5 = Always).

athlete-initiated teaching methods during training in the competitive youth sport context.

Initial descriptive analysis of the responses of the coaches and their athletes on coaches' use of the teaching methods indicated that the coaches widely use reproductive teaching methods while occasionally using problem-solving teaching methods. Little or no use of athlete-initiated teaching methods during training was reported. A Friedman's test with *post hoc* analysis of the coaches' and the athletes' responses revealed significant differences in the use of each teaching method. The athletes perceived the use of reproductive and problem-solving teaching methods as significantly lower than the coaches did during training. While the coaches and the athletes perceive the coaches' rare use of athlete-initiated teaching methods, the athletes scored higher than the coaches in their use.

This present study's findings give insight into the coaches' instructional knowledge and practices reflecting the perspective of their athletes in addition to the coaches' views. The findings corroborate the previous relevant work and arguments pointing out

the coaches' predominant use of reproductive teaching methods during training (Cassidy et al., 2008; Ince and Kiliç, 2016; Kiliç and Ince, 2017). Considering the sport settings examined, the findings imply the coaches' professional needs in providing instruction aligned to athletes' higher-order learning needs, specifically regarding using athlete-initiated and problem-solving teaching methods. High dependence on reproductive teaching methods during training, in which various learning situations and accompanying learning needs arise, may create misalignment between the instructional methods used and the athletes' age and competitive level contingent learning needs. The study findings on the coaches' pervasive use of reproductive teaching methods can be associated with 'poor coaching,' which youth athletes from a variety of backgrounds defined as 'not providing useful instruction, not aligning their instruction to each of athlete's unique needs, and not being knowledgeable about the effective use of teaching methods' (Gearity, 2012). A recent study's findings on youth athletes' developmental outcomes in context, which indicates a significant decrease in athletes' developmental sport outcomes as they age (Kiliç and Ince, 2021b), may also partially give hints of 'poor coaching' through which youth athletes may have been excessively exposed to reproductive teaching methods regardless of their learning needs.

Coaching effectiveness is linked to coaches' capacity to apply an array of teaching methods with the awareness of their implications on athletes' learning (Cassidy et al., 2008, p. 40). With the precondition of having an extensive teaching method repertoire, coaches also need to have a basic understanding of the theories supporting the teaching methods and their sets of assumptions about learning when applying them (Light, 2008). While bearing structural similarities with participation context (Jones and Turner, 2006), the competitive sport may become more demanding in solving complex problems such as skill acquisition (e.g., Wikeley and Bullock, 2006). Coaches' use of teaching methods appropriate to athletes' cognitive, psychological, social, and emotional learning needs (Fraser-Thomas et al., 2005) may be critical for effective athlete learning (Cassidy et al., 2008). Furthermore, applying solely reproductive teaching methods in training when teaching inexperienced athletes would also be flawed as the nature of learning before formal education (sport or schools) occurs through the learner's active engagement of discovery and problem-solving (Cassidy et al., 2008). Although important in teaching basic skills, reproductive methods are associated with behaviorist learning theory and its heavy use have been criticized for neglecting learners' freedom, choice, and self-direction (Nelson and Colquhoun, 2013).

The descriptive analysis of coaches' and their athletes' responses on the value (enjoyment, learning, and motivation) given to the examined teaching methods revealed that reproductive teaching methods were ranked the highest, followed by the problem-based and athlete-initiated teaching methods. A Mann-Whitney test indicated that the athletes' value perceptions were significantly higher than that of coaches in (1) reproductive teaching methods regarding learning and motivation; (2) problem-solving teaching methods regarding motivation; and (3) athlete-initiated teaching methods regarding enjoyment, learning, and motivation. A distinct difference between the mean scores of coaches and their athletes was observed regarding the value given to athlete-initiated teaching methods. The athletes valued athlete-initiated teaching methods significantly higher in all value components, while the coaches' scores were comparably quite low.

Excessive reliance on using reproductive teaching methods and placing a high value on them in the study context echoes the findings of the studies conducted in the physical education context. It may be attributed to the continuation of the traditional culture associated with the sport (Cassidy et al., 2008). As a result, the athletes may also become advocates of the dominant teaching methods and follow this tradition, especially when these methods work well during the early phases of skill learning (Cassidy et al., 2008). The prevailing sport culture may impose the taken-for-granted practices involving predominantly reproductive teaching methods through interaction with other coaches or other mechanisms such as mentoring. These random learning experiences may lead to the perpetuation of existing reproductive teaching methods in coaching practices by no more than passing on 'tricks of the trade' (Cushion et al., 2003). On the contrary, teaching methods (A spectrum of teaching styles) were designed to enhance practitioners' teaching approach by reflecting on their instruction (Mosston, 1972; Cassidy et al., 2008). Coaches' reflection on their teaching experiences using problem-solving and athlete-initiated teaching methods in addition to reproductive teaching methods can help coaches evolve their instruction and, as a result, improve athletes' holistic developmental sport outcomes (Potrac and Cassidy, 2006; Camiré et al., 2014).

While the athletes' perceptions were found to be generally higher than that of coaches in the value given to the teaching methods examined, importantly, among other teaching methods, the athletes perceived only the athlete-initiated teaching methods as significantly more enjoyable than their coaches did. In athlete development literature, involvement in enjoyable activities is linked to young athletes' development of intrinsic motivation through which they sustainably participate in sports and develop the resilience needed to overcome future athletic difficulties (e.g., Balyi et al., 2013, p. 52; Côté et al., 2010). According to the self-determination theory, the feeling of control over a person's actions (autonomy) is one of the primary basic psychological needs for developing intrinsic motivation. It involves the learner's interest, enjoyment, and inherent satisfaction (Ryan and Deci, 2017). Learner enjoyment increase in the social-contextual conditions that encourages learner autonomy in less structured activities and is critical for enhancing youth sport participation (e.g., Barnett et al., 1992; Fraser-Thomas et al., 2005).

4.1. Limitations

When interpreting the findings, the limitations of the study need consideration. Firstly, the study data comprised survey responses of coaches and their athletes in the competitive youth sport context. We suggest that future studies examining coaches' instruction integrate systematic field observations and use qualitative inquiry in addition to the survey responses of coaches and their athletes to obtain a more comprehensive understanding of coaches' capacity to apply teaching methods. Secondly, the data collection was limited to the two major cities and two little towns of Türkiye. Scanning a wider population from different coaching contexts can be done with the use of survey forms for athletes and coaches. Thirdly, in this study, the focus was on youth sports coaches and their athletes' use and value perception of teaching methods in training without considering their age and gender subgroups in the study population. Data was collected in the natural training setting from the available population. Further

studies should consider the effect of coaches' and athletes' age and gender on their use and value of teaching methods.

4.2. Conclusion and recommendations

One of the critical elements of coaching knowledge is pedagogical knowledge (Trudel et al., 2013, p. 385). Coaching effectiveness is directly affected by coaches' capacity to use this knowledge and provide instruction tailored to athletes' learning needs. Researchers urged a learner-centered pedagogy to athlete development to improve athletes' higher-order thinking in the physical and psychosocial aspects of sport (e.g., Fraser-Thomas et al., 2005; Potrac and Cassidy, 2006; Wikeley and Bullock, 2006; Cassidy et al., 2008). To realize this, coaches should be well-informed about a variety of teaching methods (Cassidy et al., 2008; Pill and SueSee, 2022) and, more importantly, their theoretical roots to appropriately apply these methods according to athletes' different learning needs (Light, 2008; Gearity, 2012; Roberts and Potrac, 2014). To develop coaches' instructional capacities in this vein, the teaching methods coaches currently apply during their practices in a definite sport context needs careful consideration. This study provides a detailed examination of the use of and value given to reproductive, problem-solving, and athlete-initiated teaching methods in a competitive youth sport context, drawing on the tenets of the spectrum of teaching styles (Mosston and Ashworth, 2008). The study findings enabled a detailed assessment of the coaches' pedagogical knowledge and practices from the athletes' perspectives in addition to their perceptions of their instruction.

The study's findings strongly indicate the coaches' professional needs in their pedagogical knowledge, specifically their repertoire of problem-solving and athlete-initiated teaching methods and the capacity to apply them. The athletes rating athlete-initiated teaching methods as significantly more enjoyable, motivating, and instructive, contrary to the coaches, also highlights the high demand for the athlete-centered approach to instruction in the sport context studied from athletes' part. To improve coaches' capacity to apply knowledge in the areas the present study findings addressed, community-based learning research has been suggested in the literature (Gilbert et al., 2009; Trudel et al., 2013). Such collaborative learning environments designed based on the social theory of learning (communities of

practice; Wenger, 1998) have been evidenced as effective in providing quality learning environments for coaches (e.g., Culver and Trudel, 2006; Stoszowski and Collins, 2014; Duarte et al., 2021).

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 Middle East Technical University Human Subject Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

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

Conflict of interest

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

Publisher's note

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

References

- Armour, K. (2010). "The learning coach... the learning approach: Professional development for sports coach professionals" in *Sports Coaching: Professionalisation and Practice*. Oxford, UK: Elsevier, 153–164.
- Balyi, I., Way, R., and Higgs, C. (2013). *Long-term Athlete Development*. Champaign, IL: Human Kinetics.
- Barnett, N. P., Smoll, F. L., and Smith, R. E. (1992). Effects of enhancing coach-athlete relationships on youth sport attrition. *Sport Psychol.* 6, 111–127. doi: 10.1123/tsp.6.2.111
- Bompa, T. O., and Haff, G. G. (2009). *Periodization: Theory and Methodology of Training*. 5th Edn. Champaign, IL: Human Kinetics.
- Camiré, M., Trudel, P., and Forneris, T. (2014). Examining how model youth sport coaches learn to facilitate positive youth development. *Phys. Educ. Sport Pedagog.* 19, 1–17. doi: 10.1080/17408989.2012.726975
- Cassidy, T. G., Jones, R. L., and Potrac, P. (2008). *Understanding sports coaching: The social, cultural and pedagogical foundations of coaching practice*. London: Routledge.
- Côté, J. (1999). The influence of the family in the development of talent in sport. *Sport Psychol.* 13, 395–417. doi: 10.1123/tsp.13.4.395
- Côté, J., Baker, J., and Abernethy, B. (2007). Practice and play in the development of sport expertise. *Handbook Sport Psychol.* 3, 184–202. doi: 10.1002/9781118270011.ch8
- Côté, J., Bruner, M., Erickson, K., Strachan, L., and Fraser-Thomas, J. (2010). Athlete development and coaching. In: *Sports Coaching: Professionalisation and Practice*. Oxford: Elsevier, 63, 84.
- Côté, J., and Gilbert, W. (2009). An integrative definition of coaching effectiveness and expertise. *Int. J. Sports Sci. Coach.* 4, 307–323. doi: 10.1260/174795409789623892
- Cothran, D. J., Kulinna, P. H., Banville, D., Choi, E., Amade-Escot, C., MacPhail, A., et al. (2005). A cross-cultural investigation of the use of teaching styles. *Res. Q. Exerc. Sport* 76, 193–201. doi: 10.1080/02701367.2005.10599280
- Culver, D. M., and Trudel, P. (2006). "Cultivating coaches' communities of practice: developing the potential for learning through interactions" in *The sports Coach as Educator*. ed. R. L. Jones (London: Routledge), 115–130.
- Curtner-Smith, M. D., Todorovich, J. R., McCaughy, N. A., and Lacon, S. A. (2001). Urban teachers use of productive and reproductive teaching styles within the confines of the national curriculum for physical education. *Eur. Phys. Educ. Rev.* 7, 177–190. doi: 10.1177/1356336X010072005

- Cushion, C. J., Armour, K. M., and Jones, R. L. (2003). Coach education and continuing professional development: experience and learning to coach. *Quest* 55, 215–230. doi: 10.1080/00336297.2003.10491800
- Cushion, C., and Nelson, L. (2013). “Coach education and learning: developing the field” in *Routledge Handbook of Sports Coaching*, eds. P. Potrac, W. Gilbert and J. Denison (London: Routledge), 359–374.
- Duarte, T., Culver, D. M., and Paquette, K. (2021). Assessing the value created in a social learning space intervention: four vignettes of parasport coaches. *Int. Sport Coaching J.* 8, 348–361. doi: 10.1123/iscj.2020-0006
- Ertmer, P. A., and Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: comparing critical features from an instructional design perspective. *Perform. Improv. Q.* 26, 43–71. doi: 10.1002/piq.21143
- Fraser-Thomas, J. L., Côté, J., and Deakin, J. (2005). Youth sport programs: an avenue to foster positive youth development. *Phys. Educ. Sport Pedagog.* 10, 19–40. doi: 10.1080/1740898042000334890
- Gearity, B. T. (2012). Poor teaching by the coach: a phenomenological description from athletes' experience of poor coaching. *Phys. Educ. Sport Pedagog.* 17, 79–96. doi: 10.1080/17408989.2010.548061
- Gilbert, W., and Côté, J. (2013). “A focus on coaches' knowledge” in *Routledge Handbook of Sport Coaching*, eds. P. Potrac, W. Gilbert and J. Denison (London: Routledge), 147–159.
- Gilbert, W., Gallimore, R., and Trudel, P. (2009). A learning community approach to coach development in youth sport. *J. Coaching Educ.* 2, 3–23. doi: 10.1123/jce.2.2.3
- GSD. (2019). Coach education regulations of Turkey [in Turkish]. Resmi Gazete Tarihi ve Sayısı: 30978, Ministry of Youth and Sport of Turkey, General Directory of Sport. Available at: <https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=34028&MevzuatTur=7&MevzuatTertip=5>. Accessed March 5, 2023.
- Ince, M. L., and Hunuk, D. (2010). Experienced physical education teachers' use of teaching styles and perceptions of styles during the educational reform period. *Educ. Sci.* 35:128.
- Ince, M. L., and Kilic, K. (2016). “The use of teaching styles in youth sport by the views of coaches and athletes: preliminary findings”, in: *2016 AIESEP International Conference*. (Laramie, Wyoming, US: University of Wyoming).
- Jones, R. L. (2006). *The sports coach as educator*. London: Taylor & Francis.
- Jones, R. L., and Turner, P. (2006). Teaching coaches to coach holistically: can problem-based learning (PBL) help? *Phys. Educ. Sport Pedagog.* 11, 181–202. doi: 10.1080/17408980600708429
- Kidman, L. (2005). *Athlete-centred Coaching: Developing Inspired and Inspiring People*. Christchurch: IPC Print Resources.
- Kilic, K., and Ince, M. L. (2017). “Examining youth sport coaches' and athletes' use and value perceptions of teaching styles”, In: *11th ICCE Global Coach Conference Challenging Sport Coaching Frontiers: The Role of Sports Science and Technology*, (Liverpool: International Council for Coaching Excellence), 177.
- Kılıç, K., and Ince, M. L. (2019). Development of coaches' use of teaching methods scale – athlete perception version (Cutems–athlete) [in Turkish]. *Int. J. Sport Exerc. Training Sci. (IJSETS)* 5, 119–128. doi: 10.18826/useabd.579590
- Kılıç, K., and Ince, M. L. (2021a). Concurrent validity and reliability of coaches' use of teaching methods scale: coach version (Cutems–coach) [in Turkish]. *Hacettepe J. Sport Sci.* 32, 64–74. doi: 10.17644/sbd.887547
- Kılıç, K., and Ince, M. L. (2021b). Youth athletes' developmental outcomes by age, gender, and type of sport. *J. Hum. Sport Exerc.* 16, 212–225. doi: 10.14198/jhse.2021.161.19
- Koca, C. (2018). *Gender Equality in Sports Mapping and Monitoring Study*. Ankara: CEID Publications.
- Kulinna, P. H., and Cothran, D. J. (2003). Physical education teachers' self-reported use and perceptions of various teaching styles. *Learn. Instr.* 13, 597–609. doi: 10.1016/S0959-4752(02)00044-0
- Lara-Bercial, S., Duffy, P., and Harrington, M. (2013). *International Sport Coaching Framework 1.2*. Champaign, IL: Human Kinetics.
- Lauer, L., and Dieffenbach, K. (2013). “Psychosocial training interventions to prepare youth sport coaches” in *Routledge Handbook of Sports Coaching*, eds. P. Potrac, W. Gilbert and J. Denison (London: Routledge), 469–480.
- Light, R. (2008). Complex learning theory—its epistemology and its assumptions about learning: implications for physical education. *J. Teach. Phys. Educ.* 27, 21–37. doi: 10.1123/jtpe.27.1.21
- Mitchell, S., Mitchell, S. A., Oslin, J., and Griffin, L. L. (2020). *Teaching Sport Concepts and Skills: A Tactical Games Approach*. Champaign, IL: Human Kinetics.
- Mosston, M. (1972). *Teaching: From Command to Discovery*. Michigan: Wadsworth Publishing Company.
- Mosston, M., and Ashworth, S. (2008). Teaching physical education: First online edition. © 2008 Sara Ashworth. Available at: <https://spectrumofteachingstyles.org/index.php?id=16> (Accessed February 12, 2023).
- Nelson, L., and Colquhoun, D. (2013). “Athlete (non) learning: is it time for an interdisciplinary understanding?” in *Routledge Handbook of Sports Coaching*, eds. P. Potrac, W. Gilbert and J. Denison (London: Routledge), 284–295.
- Pill, S., and SueSee, B. (2022). “Coaching pedagogy” in *Routledge Handbook of Coaching Children in Sport*, eds. M. Toms and R. Jeanes. (New York: Routledge), 16–24.
- Pill, S., SueSee, B., Rankin, J., and Hewitt, M. (2021). *The Spectrum of Sport Coaching Styles*. New York: Routledge.
- Potrac, P., and Cassidy, T. (2006). “The coach as a ‘more capable other’” in *The Sports Coach as Educator*, ed. R. L. Jones (London: Routledge), 57–68.
- Potrac, P., and Jones, R. (2009). Power, conflict, and cooperation: toward a micropolitics of coaching. *Quest* 61, 223–236. doi: 10.1080/00336297.2009.10483612
- Roberts, S., and Potrac, P. (2014). Behaviourism, constructivism and sports coaching pedagogy: a conversational narrative in the facilitation of player learning. *Int. Sport Coaching J.* 1, 180–187. doi: 10.1123/iscj.2014-0097
- Ryan, R. M., and Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.
- Stoszowski, J., and Collins, D. (2014). Communities of practice, social learning and networks: exploiting the social side of coach development. *Sport Educ. Soc.* 19, 773–788. doi: 10.1080/13573322.2012.692671
- Thorpe, R. (1997). *Game Sense: Developing Thinking Players (Video Recording)*. Belconnen, ACT: Australian Sports Commission.
- Trudel, P., Culver, D., and Werthner, P. (2013). “Looking at coach development from the coach-learner's perspective: considerations for coach development administrators,” in *Routledge Handbook of Sports Coaching*. (London: Routledge), 375–387.
- Trudel, P., Gilbert, W., and Werthner, P. (2010). “Coach education effectiveness” in *Sport Coaching: Professionalisation and Practice*. Oxford, 135–152.
- Wenger, E. (1998). Communities of practice: learning as a social system. *Systems thinker* 9, 2–3.
- Wikeley, F., and Bullock, K. (2006). “Coaching as an educational relationship” in *The Sports Coach as Educator*, ed. R. L. Jones (London: Routledge), 32–42.



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

José Luis Gómez Ramos,
University of Castilla-La Mancha, Spain
Keri L. Heitner,
Saybrook University, United States

*CORRESPONDENCE

Lingling Wang
✉ linglingwang@upc.edu.cn

†These authors share first authorship

SPECIALTY SECTION

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

RECEIVED 08 February 2023

ACCEPTED 31 March 2023

PUBLISHED 30 May 2023

CITATION

Li S, Hu Y, Wang L and Xia E (2023)
Involvement, leadership and social practice
to the development of postgraduate
attributes: evidence from extracurricular
education with Chinese characteristics.
Front. Psychol. 14:1161308.
doi: 10.3389/fpsyg.2023.1161308

COPYRIGHT

© 2023 Li, Hu, Wang and Xia. 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.

Involvement, leadership and social practice to the development of postgraduate attributes: evidence from extracurricular education with Chinese characteristics

Sen Li^{1†}, Yusheng Hu^{2†}, Lingling Wang^{3*} and Enjun Xia¹

¹School of Management and Economics, Beijing Institute of Technology, Beijing, China, ²School of Information Management, Beijing Information Science and Technology University, Beijing, China, ³School of Economics and Management, China University of Petroleum, Qingdao, China

Limited by students' time and energy, participation in extracurricular activities is not necessarily beneficial to the development of postgraduate attributes. Thus, it is necessary to explore the impact path of extracurricular activities and education outcomes on the development of postgraduate attributes. From a configuration perspective, this study identifies the asymmetric causal effects of engagement and extracurricular education on postgraduate attributes. First, this study proposes a theoretical framework for postgraduate attribute development in extracurricular education with Chinese characteristics based on the input-environment-output (IEO) theory. Second, 166 academic scholarship applications submitted by the whole third-grade postgraduates who are from a science and engineering school at a double first-class university in China are taken as the sample. Finally, utilizing data envelopment analysis (DEA) and fuzzy set qualitative comparative analysis (fsQCA), this study conducts the effect of the combination of causal conditions on the development of postgraduate attributes. Results are as follows: (1) the development efficiency of postgraduate attribute in extracurricular education with Chinese characteristics is practical but still insufficient; (2) four configurations consistently linked to high development efficiency of postgraduate attributes. Specifically, in context with outstanding academic research achievement and excellent moral character, participating in extracurricular education or not consistently linked to high development efficiency. In contrast, in a context characterized by academic achievement or moral award not outstanding enough, involvement in extracurricular activities or social practice is consistently linked to high development efficiency. In addition, no configuration links student leadership to high development efficiency, and non-scientific research ability is consistently linked to low development efficiency; (3) there is an asymmetric causal relationship between the high and

low development efficiency paths, indicating that the conditions affecting the development of postgraduate attributes have multiple concurrencies. These findings provide a new practical path and perspective for promoting the development of postgraduate attributes through extracurricular education with Chinese characteristics.

KEYWORDS

postgraduate attribute, extracurricular education, student involvement, student leadership, social practice

1. Introduction

Student development is some kind of positive change that occurs in the student, such as intellectual and ethical growth, cognitive complexity, self-awareness, and social identity (Patton et al., 2016). Rodgers (1990) defined student development as the ways that a student grows, progresses, or increases his or her developmental capabilities as a result of enrollment in an institution of higher education. Higher education serves several fundamental roles in that it imparts knowledge, develops students' professional skills, promotes other academic and professional outcomes, and shapes students' attributes and qualities (Anderson et al., 2013). Archer and Davison (2008) found that personal attributes, such as communication skills, integrity, self-confidence, and personality, were among the top 10 qualities that employers thought important when recruiting graduates. These professional skills and personal attributes must be acquired through formal and informal experiences in different learning spaces (Jackson, 2008), with academic study and extracurricular activities (ECAs) representing other spaces within the student experience (Barnett, 2011).

Numerous studies have reported that extracurricular activity participation (ECAP) during adolescence is positively associated with academic performance (Bartko and Eccles, 2003; Roeser and Peck, 2003; Peck et al., 2008; Guilmette et al., 2019), psychological health (Viau and Poulin, 2015), persistence (Lufi and Tenenbaum, 1991; Allen et al., 2015), self-regulation (Hansen et al., 2003), and social interaction (Mahoney et al., 2005). Similarly, research on the value of ECAs in higher education has shown that ECAP benefits academic achievement (Terenzini et al., 1996; Shulruf, 2010; Van der Zanden et al., 2018) and that the quality of ECAP significantly predicts adaptation and transition to university (Harvey et al., 2006; Tieu and Pancer, 2009; Tieu et al., 2010). For example, data drawn from the U.S. National Survey of Student Engagement indicate that spending more time engaging in extracurricular activities increases students' likelihood of acquiring valuable work skills and attributes such as improved critical thinking, leadership, and social skills (Kuh, 1995; Tieu et al., 2010; Thompson et al., 2013). In university settings, ECAP could also promote civic engagement, prosocial activities, and social integration (Fredricks and Eccles, 2006) as well as decreased internalization problems such as symptoms of depression and anxiety (Bartko and Eccles, 2003).

Extracurricular education with Chinese characteristics has become an essential supplement to classroom learning since 1983 (Zhu et al., 1983). Many Chinese scholars have demonstrated that

extracurricular education with Chinese characteristics has unique functions and advantages with regard to cultivating students' comprehensive quality (Song and Zeng, 2018; Chen et al., 2019; Xiao and Ren, 2021), and this field has gradually developed into a new front for the comprehensive implementation of quality-oriented education in China. For example, the Central Committee of the Communist Youth League and the Ministry of Education jointly proposed the university extracurricular activity record system in 2016 (CYD, 2018). This record system covers innovation and entrepreneurship, social practice, culture and art, sports activities, work history, skills and expertise, ideological growth (Chen et al., 2019). At the end of 2017, the Central Committee of the Communist Youth League started to pilot and promote this system on a larger scale (CYD, 2018). By May 2018, 507 universities and 291 junior colleges in China had implemented the university extracurricular activity record system, and 427 universities and 234 junior colleges had issued or were on the verge of releasing relevant documents (CYD, 2018). Generally, all colleges and universities in China have reached a preliminary consensus regarding the implementation of the university extracurricular activity record system (CYD, 2018).

It is worth noting that, although these domestic and foreign findings have highlighted the association between ECAP and educational outcomes, not all such studies have reported a positive association in this context (Lindsay and Paton-Saltzberg, 1994). Moreover, due to the limitations of research methods, more evidence supporting causal effects is needed (Shulruf, 2010). In addition, given the richness of extracurricular education with Chinese characteristics and the limited time and energy of students, increasing numbers of participants in extracurricular education entail a reduction in academic investment. The mention fact leads to the following reflections: "—Is extracurricular education better for students? —How can students achieve effective development through extracurricular education?" Student development is like a systematic project that is affected not only by input factors but also by output factors which, in turn, affect the further development of students (Zou and Tan, 2013). The impacts of various factors on student development are not independent, and the factors have different combination effects through linkage and matching (Zepke et al., 2011). Therefore, it is necessary to explore how extracurricular education affects the development of students from a configuration perspective.

Based on student development theory and graduate attribute theory, this study proposes a theoretical framework for postgraduate attribute development in extracurricular

education with Chinese characteristics. Based on a combination of DEA and fsQCA, this study aims to identify the multiple concurrent causalities (Ragin, 2009a) that might improve/hinder postgraduate attribute development. The sample consists of 166 academic scholarship applications submitted by the whole third-grade postgraduates who are from a science and engineering school at a double first-class university in China. The primary contributions of this study can be summarized as follows: (1) based on the IEO model, this study develops a collaborative linkage framework of postgraduate attribute development in extracurricular education with Chinese characteristic; (2) using the DEA method, this study measures the efficiency of the development of postgraduate attributes and verifies the effectiveness of extracurricular education with Chinese characteristics; and (3) from the configurational perspective, this study proposes four practical ways of promoting postgraduate attribute development in extracurricular education with Chinese characteristics, which are refined into an output-driven model and an input-output linkage model.

2. Theoretical background

2.1. Student development theory

According to different perspectives, student development theories can be divided into psychosocial, cognitive-structural, typological, and person-environment schools (Patton et al., 2016). Psychosocial theories focus on analyzing the content of student development and emphasize the critical issues that students face in college. Representative theories include Erikson's life cycle theory (Erikson, 1994) and Chickering's seven-vector development theory (Chickering and Reisser, 1993). Cognitive-structural theories focus on describing the process of students' epistemological and intellectual development that occurs during college (Kohlberg, 1969; Perry, 1999). Typological theories treat differences among students as relatively fixed characteristics reflected in the cognitive process. Kolb (2007) and Myers and Myers (2010) are representative authors. Person-environment theories regard student development as the result of the interaction between individuals and the environment, such as Astin's (1970) IEO model and student involvement theory (Astin, 1984), Pascarella's comprehensive causality model (Pascarella, 1985), and Tinto's student departure theory (Tinto, 2012).

Astin's IEO model is one of the most frequently used frameworks for understanding the effects of college on a range of outcomes of undergraduate education (Gayles et al., 2012), which has been used extensively in higher education settings for nearly four decades (Stein, 2007). The model includes three major components—inputs, environments, and outputs. Inputs refer to the personal traits, attitudes, and characteristics of students themselves, such as gender, race, high school grade point average, expectations, and involvement. Environments include the programs, strategies, interventions, peers' influences, and policies introduced to the student while they are in college. Outcomes are the consequences and/or results that are derived through the influence of the input and environment variables, such as academic

achievement, critical thinking disposition, and social-emotional adjustment (Van der Zanden et al., 2018).

2.2. Graduate attributes

Graduate attributes are the qualities, skills, and understandings that a university community agrees that its students should develop during their time at the institution (Bowden et al., 2000). These attributes include the disciplinary expertise or technical knowledge that has traditionally formed the core of most university courses, and the qualities that prepare graduates to become agents of social good in an unknown future (Bowden et al., 2000). Many studies have claimed that these attributes are core outcomes of higher education and have thus emphasized the importance of graduate attributes with regard to employability and social-civic responsibility. Barrie (2007) identified four distinct orientations: precursory, complementary, translation and enabling. In the most complex integrated conception, graduate attributes are understood to be interwoven aptitudes and abilities, such as academic inquiry and intellectual curiosity, the ability to accommodate diversity and alternative perspectives, the ability to create and defend ideas, and the ability to employ communication as a vehicle for learning (Barrie, 2007). While these outcomes might lie at the heart of formal academic and university curricula, the processes by which students develop these abilities may also extend far beyond the regular academic classroom. In other words, graduate attributes develop through student engagement during the learning experience that occurs due to the individual's immersion in the university's intellectual and social community, of which formal courses are only one part. Moreover, student-centered approaches to the development of graduate attributes are consistently associated with better graduate attribute outcomes. Therefore, universities' efforts to develop effective student-centered strategies serve to guarantee the quality of university teaching (Prosser and Barrie, 2003).

2.3. Research framework

Drawing on the research conducted by Bowden et al. (2000), this study defines postgraduate attributes in terms of the qualities, skills, and understandings that a university community agrees that its students should develop during their time at the institution. It is worth noting that the postgraduate attributes referenced in this study are based on the most complex integrated conception summarized by Barrie (2007). These attributes include not only disciplinary expertise or technical knowledge, which lie at the core of the university curriculum, but also communication ability, social ability, student leadership, a sense of responsibility, and the ability to serve society. Consequently, this study follows Barrie's emphasis on student participation in the broader university experience and the critical importance of designing student-centered teaching strategies.

Therefore, based on Astin's IEO model, this study regards extracurricular education with Chinese Characteristics as a dimension of the university environment, and divides the effect of extracurricular education on postgraduate attribute development

into three parts: the student involvement, student leadership role, and social practice of prosocial behavior in the extracurricular education as *input factors*; the extracurricular education with Chinese characteristics as *environment*, scientific research ability and moral character as *output factors* corresponding to academic education and moral education in China. Notably, postgraduate attribute development is a systematic project with the linkage effects among input factors, environment, and output factors (Zou and Tan, 2013). The research framework is drawn in **Figure 1**.

2.4. Research hypotheses

2.4.1. Student involvement and postgraduate attribute development

Student involvement is now considered to be an overarching meta-construct in which an ecosystem of students, educators, service staff, and institutions interact to create enriching tertiary experiences (Zepke, 2014; Kahu and Nelson, 2018). Bowden et al. (2021) defined student involvement as a student's positive social, cognitive, emotional, and behavioral investments in the context of interacting with their tertiary institution and its focal agents, such as peers and employees. These four dimensions jointly motivate student striving, persistence, and retention within academic contexts (Kuh et al., 2008).

For example, involvement has strong links with the emotional dimension of engagement, as the more involved a student is, the higher that student's propensity for the creation of intrinsic value, such as joy, elation, enjoyment, passion, pride (Kahu, 2013; Pansari and Kumar, 2017). Involvement can also motivate the social dimension of engagement, such as by enhancing feelings of connection, belonging, warmth, and relatedness between students and their peers, staff, and the institution community (O'Keeffe, 2013). In addition, positive involvement is central to academic achievement (De Carolis et al., 2020). Thus, we propose:

Hypothesis 1: Postgraduates who are more involved in extracurricular education with Chinese characteristics are more likely to get a higher development efficiency of the postgraduate attribute.

2.4.2. Social practice and postgraduate attribute development

Many scholars have called for more authentic forms of instruction based on Dewey's notion of experiential learning (Dewey, 1986) and Bandura's social learning theory (Bandura and Walters, 1977). The social practice represents a potentially powerful form of pedagogy because all types of social practice activities involve the incorporation of academic learning with practical service (Salam et al., 2019). Various scholars have noted that social practice offers a variety of benefits for students, such as a deep understanding of course contents (Dienhart et al., 2016), increased student practical experience (Meyer et al., 2016), acquiring specific career-oriented skills (Geller et al., 2016), enhancing social responsibility and civic leadership (Weiler et al., 2013). Therefore, we propose:

Hypothesis 2: Postgraduates who participate more in social practice are more likely to get a higher development efficiency of the postgraduate attribute.

2.4.3. Student leadership and postgraduate attribute development

Scholars have agreed that one of the core functions of universities is to prepare the leaders of tomorrow's society (Astin and Astin, 2000). Developing leadership in students is increasingly becoming an explicit priority of higher education institutions (Skalicky et al., 2020). This is because universities face the challenge of graduate employability, which has prompted institutions to pay attention to the development of graduate attributes, including leadership (Osmani et al., 2015). Student leadership development within the higher education context comes in a variety of forms, including those that are specifically designed to develop leadership capabilities within a program (e.g., student union, leadership award programs), and those that provide leadership experience incidentally through student-facing activities (e.g., tutoring peers, campus tour guides) or community volunteer activities (Skalicky et al., 2020). All students who involve themselves in leadership education have the potential to increase their skills and knowledge (Zimmerman-Oster and Burkhardt, 1999; Komives et al., 2006; Wren, 2013). Therefore, we propose:

Hypothesis 3: Postgraduates who take on a student leadership role are more likely to get a higher development efficiency of the postgraduate attribute.

2.4.4. Scientific research ability and postgraduate attribute development

Scientific research ability is an essential indicator of the quality of graduate education (Zajda and Rust, 2016; Van der Zanden et al., 2019). From the social perspective, postgraduates' scientific research ability directly affects national scientific and technological innovation and comprehensive competitiveness. Therefore, the British government identified the following as the primary goal of postgraduate education in 1974, which is meeting the future talent needs of the country, providing further training for qualified, appropriate, and keen students; and contributing to the advancement of knowledge (Clark and Clark, 1993).

The National Science Foundation of the United States implemented a strategic framework for postgraduate education investment from 2016 to 2020 (Huang et al., 2021). China's Ministry of Education also implemented a postgraduate education innovation plan in 2003–2007 to improve postgraduate innovation ability. Cultivating scientific research ability is a continuous and uninterrupted process, which is not only related to curriculum system design, academic environment, tutor relationships, but also to extracurricular social practice, community culture, and psychological capital (Zhou and Wang, 2018). Therefore, we propose:

Hypothesis 4: Postgraduates with a higher scientific research ability are more likely to get a higher development efficiency of the postgraduate attribute.

2.4.5. Moral character and postgraduate attribute development

Moral education practice has become a student-centered practice, and incorporated into various fields, such as music, health, and business (Julia et al., 2020). Recently, moral character has developed into an essential and popular theme in various countries (Rissanen et al., 2018; Darnell et al., 2019). For example,

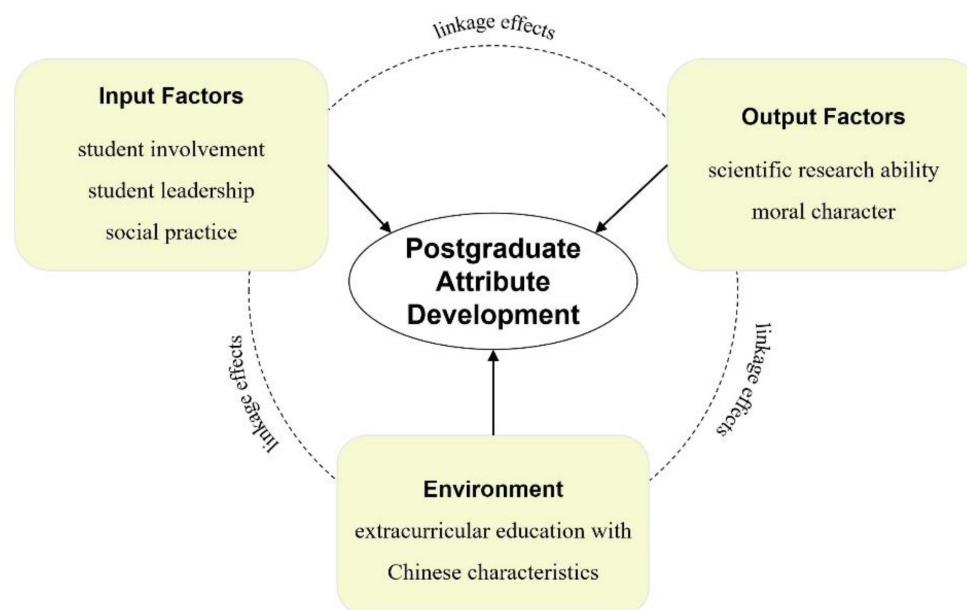


FIGURE 1

Research framework of postgraduate attribute development in extracurricular education with Chinese characteristic.

increasing numbers of schools in the U.S. have incorporated character education into their programs (Narvaez and Lapsley, 2008). A comprehensive civics and moral education program was established in Singapore, which focuses on nurturing both moral and character development and active citizenship and aims to equip students with appropriate competencies and value systems (Koh, 2012). In China, moral education belongs to citizenship education (Zhao and Tan, 2007). Integrative character education is not a package of virtues designed to control student behavior, but rather a continuation of moral development through role building, interactions among students, families, and the community, as well as student active participation, discussion and reflection (Marshall et al., 2011). The research conducted by Marshall et al. (2011) shows that comprehensive character education can improve students' prosocial and moral behaviors and their reading and test scores. Therefore, we propose:

Hypothesis 5: Postgraduate students with a higher moral character are more likely to get a higher development efficiency of the postgraduate attribute.

3. Methodology

3.1. Research method

Fuzzy set qualitative comparative analysis is a variant of qualitative comparative analysis that combines fuzzy sets with logical principles in the context of qualitative comparative analysis (Ragin, 2009b). Compared with other variance-based quantitative techniques, fsQCA aims to disentangle the relationship of these combinations of variables to a given outcome (Fiss, 2007) by expressing the relative contribution of a set of variables to a given outcome rather than individual contributions (Vis, 2012). The

essential features of this approach are as follows: (1) fsQCA is based on Boolean algebra and its associated tools, such as truth tables, consistency, coverage scores, and set coincidences (Ragin, 2014). (2) fsQCA is a method of data analysis used for small and medium sample cases that combines the case orientation of qualitative methods and the variable orientation of quantitative methods, thereby combining case analysis and cross-case comparison. However, larger samples can also be investigated (Vis, 2012). (3) fsQCA is based on combinations of variables, including equilibria, that is, different paths leading to specific outcomes (Fernández-García et al., 2020). This approach is used to explore the generative logic across different conditional configurations and outcomes and to identify the synergistic effects of multiple conditional variables based on an acknowledgment of the complexity of causality (Du and Jia, 2017).

The development of postgraduate attributes is a systematic project in which input and output factors interact with each other, and the impacts of each factor are not independent. These impacts have different combined effects through linkage matching. Therefore, to explore the multiple driving mechanisms underlying the promotion of postgraduate attribute development in extracurricular education with Chinese characteristics, this study uses fsQCA to examine the multiple conjunctural causations between student input and output in the context of a university environment featuring extracurricular education.

3.2. Sample and data collection

Referring to the study by Guilmette et al. (2019), this study selected a science and engineering school at a double first-class university in China. Given that the output of postgraduates will increase gradually over time, this study took a random sample consists of academic scholarship applications which are submitted

by the whole third-grade postgraduates voluntarily. Data collection was carried out in September 2022. Finally, 166 valid academic scholarship applications were obtained. Notably, the extensive coverage of graduate academic scholarships for as many as 90% of students guarantees the diversity of applicants and research value. The proportions of special prizes, first prizes, second prizes, and non-awarded applicants in the sample are 22.30, 21.69, 42.78, and 13.25%, respectively, thus ensuring maximum heterogeneity in the sample and avoiding excessive consistency within the sample.

3.3. Variable measurement

Conditional variables in this study include student involvement, social practice, student leadership, academic achievement, and moral development, all measured by objective indicators included in the scholarship applications, all of which are fuzzy values. The measurements of these variables are displayed in [Table 1](#).

3.3.1. Student involvement (SI)

The behavioral dimension is the most frequently measured dimension within national barometers of the student experience (Kuh et al., 2008; Zepke, 2014). According to the research by Khademi Ashkzari et al. (2018), this study selects the number of participants in extracurricular activities (e.g., science and technology activities, competitions, cultural and sports activities, and other activities) as the measurement of student involvement in extracurricular education with Chinese characteristics.

3.3.2. Social practice (SP)

Referring to the research by Scales et al. (2006), this study selects the total score of volunteer service, summer social practice, and winter social practice in the scholarship applications as the measurement of students' social practice in extracurricular education with Chinese characteristics.

3.3.3. Student leadership (SL)

According to the research by Skalicky et al. (2020), the work performance score of the student leadership role in the scholarship applications is selected as the measurement of student leadership.

3.3.4. Scientific research ability (SR)

According to the studies by Thursby (2000), Bonaccorsi et al. (2006), Abramo et al. (2008), this study selects the academic achievement score in the scholarship applications, that is, the scores of published papers and authorized patents, as the measurement of scientific research ability.

3.3.5. Moral character (MC)

A moral sense of honor is the praise and award people get after fulfilling their moral obligations. This study refers to the research of Kou (2011) and selects the honor and award scores in the scholarship applications as the measurement of students' moral character.

The outcome variable of this study is *Postgraduate Attribute Development* (PAD). Referring to the research of Amara et al. (2020), total efficiency is calculated through the input-oriented CRS model of DEA, in which student involvement, social practice,

and student leadership are input variables, and scientific research ability, moral development are output variables. The formula for calculating PAD is shown in formula (1):

$$PAD = total\ efficiency \times (SI + SP + SL + SR + MC) \quad (1)$$

4. Results

4.1. Efficiency assessment

This study developed an input-oriented CRS efficiency model featuring two outputs (scientific research ability and moral character) and three inputs (student involvement, social practice, and student leadership). DEAP-XP1 software was used to obtain the DEA efficiency scores for the development of postgraduate attributes in extracurricular education with Chinese characteristics.

As shown in [Table 2](#), in extracurricular education with Chinese characteristics, the average efficiency score of the sample is 62.15%, and approximately half of the postgraduates exhibited strong DEA efficiency (development efficiency = 1). However, more than 40% of postgraduates did not achieve 50% efficiency, and nearly 45% of postgraduates did not achieve 75% efficiency. These findings show that extracurricular education with Chinese characteristics can be substantially improved with regard to its ability to promote development efficiency with regard to postgraduate attributes.

4.2. Variable calibration

The first step in fsQCA is a process known as calibration, in which context each case is assigned a degree of membership in each set (Ragin, 2009b). Three thresholds (full membership, a crossover point of maximum membership ambiguity, and full non-membership) are required to rescale interval variables into fuzzy sets. The research conducted by Ragin (2009b), Fiss (2011) indicated that theoretical and empirical knowledge should be taken into full consideration when selecting these three thresholds. The full membership point can be set at the 10th percentile, the 20th percentile, or the upper percentile; the crossover point can be set at the average or the median; and the full non-membership point can be set at the 80th percentile or the 90th percentile (Sun and Li, 2021).

In this study, considering the distribution characteristics of the case data ([Table 3](#)), we use sample-dependent anchors to define set membership (Gupta et al., 2020). Specifically, the three thresholds of student involvement, moral character, and postgraduate attributes development are set as follows: the 90th percentile is the anchor for full membership; the 10th percentile is the anchor for non-membership; and the median is the crossover point anchor. Similarly, the three thresholds of scientific research ability, social practice, and student leadership are set as follows: the 90th percentile is the anchor for full membership; the 10th percentile is the anchor for non-membership; and the mean is the crossover point anchor. The calibration values of the outcomes and conditions are shown in [Table 4](#). This study uses these calibration values and fsQCA 3.1 software to calibrate the variable data.

TABLE 1 Variable list.

Construct	Measurement	References
Student involvement	The number of participants in extracurricular activities (e.g., science and technology activities, competitions, cultural and sports activities, and other activities)	Khademi Ashkzari et al., 2018; Skalicky et al., 2020
Social practice	The total score of volunteer service, summer social practice, and winter social practice	Scales et al., 2006
Student leadership	The work performance score of the student leadership role	Skalicky et al., 2020
Scientific research ability	The academic achievement score of published papers and authorized patents	Thursby, 2000; Bonaccorsi et al., 2006; Abramo et al., 2008
Moral character	The honor and award scores in the scholarship applications	Kou, 2011
Postgraduate attribute development	The total efficiency is calculated through the input-oriented CRS model of DEA, in which student involvement, social practice, and student leadership are input variables, and scientific research ability, moral development are output variables	Amara et al., 2020

4.3. Necessity analyses

Before conducting a sufficiency analysis, necessity analyses must be conducted to determine whether any condition is necessary for obtaining high or low PAD. A condition is necessary if it must be present for the outcome to occur (Ragin, 2009b). Consistency thresholds for necessity analyses are typically set above 0.90 (Ragin, 2009a,b; Schneider and Wagemann, 2012; Afonso et al., 2018; Grau and López, 2018), and a condition may be considered to be necessary if it has high consistency and coverage scores. The necessity analysis results concerning high PAD show that the consistency of each condition is lower than 0.9 (Table 5), indicating that no condition is necessary for the occurrence of high postgraduate attribute development. That is, the explanatory power of any single condition on high PAD is insufficient. Therefore, the effect of a combination of conditions on high PAD must be analyzed in further detail. Correspondingly, the necessity analysis for low PAD shows that the consistency of weak scientific research ability ($\sim$ SR) is higher than 0.9, thus suggesting that weak scientific research ability may be necessary to explain low postgraduate attribute development. However, the consistency of the other conditions are all lower than 0.9, thus indicating that further analysis of the causal condition combinations impacting low PAD is needed.

4.4. Sufficiency analysis

According to fsQCA truth-table approach, this study constructs and analyzes a data matrix known as the truth table using fsQCA 3.1, which includes 32 rows ($2^5 = 32$; the number of explanatory conditions is five). Each row represents a logically possible combination of conditions. Using Boolean algebra, the truth table is reduced into a simplified expression of combinations linked to the outcome. To simplify the truth table, we follow the criteria suggested by Gupta et al. (2020): (1) a frequency threshold for the minimum number of cases that must belong to a combination for that combination to be considered by the analysis and (b) a consistency benchmark to identify combinations that are reliably linked to an outcome. Therefore, this study sets the case frequency threshold at the minimum critical value of one and chooses a raw consistency threshold of 0.75 (Rihoux and Ragin, 2008), which is supplemented by a 0.7 proportional reduction in consistency (PRI) threshold (Greckhamer et al., 2018). In addition, in the section on counterfactual analysis, this study

does not presuppose the direction, and all factors are measured in terms of existence or absence. Namely, the development of postgraduate attributes is a system of interactions among input, environment, and output, which is challenging to judge in a uniform manner. This study uses the Quine-McCluskey method to obtain intermediate, parsimonious, and complex solutions (Mustafa et al., 2022). Furthermore, the intermediate solution is the main solution, which is supplemented by the parsimonious solution and used to identify the core and peripheral conditions (Fiss, 2011).

To understand the results more accurately, we follow the convention of using a configuration chart to report a combination of parsimonious and intermediate solutions, as shown in Table 6. Core conditions are contained in the intermediate and parsimonious solutions, while peripheral conditions are contained in the intermediate solution but not in the parsimonious solution (Pappas and Woodside, 2021). In Table 6, black circles represent the presence of a causal condition, whereas crossed circles indicate the absence of such a causal condition. Large circles represent core conditions, whereas small circles indicate peripheral conditions. Blank spaces indicate a state of “do not care.”

Table 6 shows that four configurations of student involvement, leadership, social practice, scientific research ability, and moral character, are consistently linked to high PAD. These configurations can be refined into the output-driven model (Configuration H₁) and the input-output linkage model (Configurations H₂, H₃, and H₄). The overall consistency and coverage are 0.936 (>0.80) and 0.517 (>0.10), respectively, which are higher than the thresholds suggested by Ragin (2009a,b), Woodside (2014), indicating that these four configurations have good explanatory power and are sufficient to ensure high development efficiency with regard to postgraduate attributes. The overall solution coverage is 0.517, which indicates that the four configurations

TABLE 2 Development efficiency of postgraduates in extracurricular education with Chinese characteristics.

Development efficiency (%)	Number	Proportions (%)
<25	44	26.5
25–49.99	23	13.9
50–74.99	7	4.2
75–99.99	8	4.8
100	84	50.6
Total	166	100

TABLE 3 Descriptive statistics and percentile analysis of the outcomes and conditions.

	SI	SP	SL	SR	MC	PAD
SD	5.962	1.996	1.321	4.052	2.044	8.392
Max	33.000	10.000	6.300	19.000	10.000	41.002
Min	0.001	0.001	0.001	0.001	0.001	0.000
M	5.389	1.120	0.445	2.129	1.039	6.819
P_5	0.001	0.001	0.001	0.001	0.001	0.002
P_10	0.001	0.001	0.001	0.001	0.001	0.004
P_20	1.000	0.001	0.001	0.001	0.001	0.204
P_25	1.000	0.001	0.001	0.001	0.001	0.539
P_30	2.000	0.001	0.001	0.001	0.001	1.004
P_50 (Md)	4.000	0.001	0.001	0.001	0.100	3.579
P_70	5.800	1.000	0.001	1.000	0.500	8.096
P_75	6.000	1.000	0.001	2.000	1.181	10.416
P_80	8.000	2.000	0.001	3.330	2.000	12.402
P_90	13.700	4.000	1.000	8.500	3.650	21.288
P_95	19.000	6.000	4.000	10.875	5.100	25.376

SD, standard deviation; Max, maximum; Min, minimum; M, mean; Md, median; P_5, 5th percentile; P_10, 10th percentile; P_20, 20th percentile; P_30, 30th percentile; P_50, 50th percentile; P_70, 70th percentile; P_75, 75th percentile; P_80, 80th percentile; P_90, 90th percentile; P_95, 95th percentile.

TABLE 4 Calibration values of the outcomes and conditions.

	Thresholds		
	Full membership point	Crossover point	Full non-membership point
SI	0.001 (10th percentile)	4.000 (Median)	13.700 (90th percentile)
SP	0.001 (10th percentile)	1.120 (Mean)	4.000 (90th percentile)
SL	0.001 (10th percentile)	0.445 (Mean)	1.000 (90th percentile)
SR	0.001 (10th percentile)	2.129 (Mean)	8.500 (90th percentile)
MC	0.001 (10th percentile)	0.100 (Median)	3.650 (90th percentile)
PAD	0.001 (10th percentile)	3.579 (Median)	21.288 (90th percentile)

TABLE 5 Necessary conditions for high and low development efficiency of postgraduate attributes.

	High PAD		Low PAD	
	Consistency	Coverage	Consistency	Coverage
SI	0.684965	0.673975	0.475727	0.596689
~SI	0.590112	0.468933	0.740067	0.749657
SP	0.498192	0.720549	0.297085	0.547725
~SP	0.687294	0.434086	0.848426	0.683064
SL	0.256331	0.575286	0.228027	0.652355
~SL	0.845100	0.462018	0.851544	0.593435
SR	0.530805	0.855176	0.230951	0.474303
~SR	0.673701	0.407311	0.929481	0.716331
MC	0.680580	0.741992	0.379751	0.527758
~MC	0.566844	0.417569	0.814349	0.764698

can explain 51.7% of the empirical cases. Moreover, the raw coverage of the four configurations ranges between 0.141 and 0.280, which is higher than 0.10 and can thus be considered

satisfactory (Eng and Woodside, 2012). Similarly, Table 6 shows four configurations linked to low PAD, the overall consistency and coverage values of which are 0.938 (>0.80) and 0.704 (>0.10), respectively. Moreover, the raw coverage of the four sufficient conditions ranges between 0.138 and 0.632, indicating that these conditions serve as a compelling explanation for low development efficiency with regard to postgraduate attributes.

4.5. Robustness test

To ensure the accuracy of the results, we conducted a stability test. First, the consistency threshold was increased from 0.75 to 0.80. Second, the PRI consistency was increased from 0.70 to 0.75. The results were all consistent with the configuration solutions shown in Table 6. Therefore, the configuration results obtained in this study are robust (Greckhamer et al., 2018; Yuan et al., 2022).

5. Discussion

5.1. Configuration path of high PAD

Configuration H₁ (SR*MC*~SP) shows that the configuration of scientific research ability, moral character, and a lack of social practice is linked to high development efficiency with regard to postgraduate attributes. The consistency is 0.989 (>0.800), and the raw coverage is 0.280, indicating that this path can explain approximately 28.0% of cases (Leppänen et al., 2019). The number of example cases representing this configuration is 19. Accordingly, compared with other conditions, strengthening scientific research ability and moral traits is particularly important with respect to improving the development efficiency of postgraduate attributes.

Configuration H₁ is consistent with studies concerning policy and research on higher education effects, which have generally used scientific research ability as an essential indicator of student success, academic achievement, or the quality of graduate education (Zajda and Rust, 2016; Van der Zanden et al., 2019). This configuration is also consistent with the emphasis on scientific research achievements in international/national accreditation (De Boer et al., 2015). Furthermore, Configuration H₁ validates the claim that postgraduate attributes are a multidisciplinary concept that includes various outcomes (Wegner, 2008; Universities UK, 2015). Higher education should not only focus on imparting knowledge (Mayhew et al., 2016) but should also include moral education and character formation. Moral character is both the starting point and a critical point for education (Wolf-Wendel and Ruel, 1999; Podger et al., 2010). Therefore, higher education should focus on cultivating students' essential qualities and behavioral standards, shaping students' moral responsibilities (Shaaban, 2005; Podger et al., 2010), and providing students with appropriate abilities and value systems (Nucci, 1997; Corrigan et al., 2007; Marshall et al., 2011). This approach is consistent with the central role of moral character in improving the development efficiency of postgraduate attributes emphasized in Configuration H₁.

Configuration H₂ (SR*SI*~SL) shows that the configuration of scientific research ability, student involvement, and a lack

TABLE 6 Sufficient conditions for the development efficiency of postgraduate attributes.

	High PAD				Low PAD			
	1	2	3	4	1	2	3	4
SI	-	●	-	●	⊗	-	⊗	-
SP	⊗	-	●	●	⊗	⊗	-	-
SL	-	⊗	⊗	⊗	-	●	●	●
SR	●	●	●	-	⊗	⊗	⊗	
MC	●	-	⊗	●	-	-	-	⊗
Consistency	0.989	0.983	1.000	0.905	0.958	0.965	0.870	0.943
Raw coverage	0.280	0.277	0.141	0.268	0.632	0.138	0.142	0.158
Unique coverage	0.100	0.066	0.013	0.123	0.517	0.009	0.008	0.023
Overall solution consistency	0.936				0.938			
Overall solution coverage	0.517				0.704			

Black circles represent the presence of a causal condition, whereas crossed circles indicate the absence of such a causal condition. Large circles represent core conditions, whereas small circles indicate peripheral conditions. Blank spaces indicate a state of “do not care.”

of student leadership is linked to high development efficiency with regard to postgraduate attributes. The consistency is 0.983 (>0.800), and the raw coverage is 0.277, indicating that this path is able to explain approximately 27.7% of cases. With the exception of conducting scientific research, Configuration H₂ highlights participation in extracurricular education such as scientific and technological activities, cultural and sports activities, and competitions. This path is consistent with some findings regarding the value of extracurricular activities. For example, extracurricular involvement helps enhance individual adaptive self-regulation (Larson and Verma, 1999; Hansen et al., 2003; Tieu and Pancer, 2009) and learning persistence (Allen et al., 2015), which are helpful for scientific research (Nonis et al., 2006). Participation in extracurricular activities can also induce positive feelings, such as a sense of connection with others, and allow the participant to obtain many valuable job skills or attributes, such as improved critical thinking, leadership, and social skills (Tieu et al., 2010). However, more participation in extracurricular education is not guaranteed to be beneficial for postgraduate attribute development due to the student's limitations with regard to time and energy. Previous studies have revealed that extracurricular involvement hinders academic research (Lindsay and Paton-Saltzberg, 1994) and that this negative effect depends on individual characteristics and circumstances. Therefore, the task of achieving a good balance between academic research and extracurricular activities is challenging (Thompson et al., 2013).

Configuration H₃ (SR*~MC*SP*~SL) shows that the configuration of scientific research ability, social practice, and a lack of student leadership are core conditions and that complementary non-moral character is linked to high development efficiency with regard to postgraduate attributes. The consistency is 1.000 (>0.800), and the raw coverage is 0.141, indicating that this path can explain approximately 14.1% of cases. Configuration H₃ emphasizes the fact that while conducting scientific research, graduate students should actively participate in social practice activities associated with prosocial behaviors, such as volunteer service and community service, the cultivation of moral responsibility, and the improvement of their development

efficiency and efficiency. Our findings are consistent with those of numerous studies that have identified service learning as an effective tool that can benefit students (Strage, 2000). For example, participation in community service and service learning effectively contribute to higher grades (Billig, 2004; Scales and Roehlkepartain, 2004) and academic advantages (Benson et al., 1999; Scales et al., 2006). Service learning also has a powerful impact on students' moral, social-cognitive, and emotional development (Batchelder and Root, 1994; Ostrow, 1995; Eyler and Giles, 1996; Eyler and Giles, 1999; Kendrick, 1996; Rhoads, 1997). Participation in service learning develops students' sense of responsibility, acceptance of diversity, and leadership skills in community work (Markus et al., 1993; Eyler and Giles, 1996; Shumer and Belbas, 1996; Brandell and Hinck, 1997; Myers-Lipton, 1998). This situation may be due to the fact that community service and service learning provide opportunities for students to experience meaningful engagement in various life settings (Hanson et al., 2003).

Configuration H₄ (MC*SI*SP*~SL) shows that the configuration of student involvement, social practice, moral character, and a lack of student leadership are core conditions that are linked to high development efficiency with regard to postgraduate attributes. The consistency is 0.905 (>0.800), and the raw coverage is 0.268, indicating that this path can explain approximately 26.8% of cases. Configuration H₄ indicates that when the scientific research achievements are insufficiently outstanding, postgraduate attributes can be improved by increasing involvement in extracurricular activities and social practices associated with prosocial behaviors. For example, some students actively participate in extracurricular activities because they do not expect to make strong academic research achievements, even though extracurricular education is not a substitute for exemplary achievement. Nevertheless, extracurricular education encourages the acquisition of specific skills and personal development, which can boost employability (Thompson et al., 2013).

These four configurations of high PAD can be further refined into two models: the output-driven model (configuration H₁) and the input-output linkage model (configurations H₂, H₃, and H₄). Comparing configurations H₂ and H₃, we find that strong student

involvement can substitute for strong social practice in the context of strong scientific research ability and a lack of leadership. Namely, there is a potential substitution relationship between configurations H_2 and H_3 . Scientific research ability is a core condition in three out of four paths. Therefore, scientific research ability remains essential for improving the efficacy and development efficiency of postgraduate attributes.

5.2. Configuration path of low PAD

Configuration L_1 ($\sim SR^* \sim SI^* \sim SP$) shows that in cases featuring a lack of scientific research ability and no participation in extracurricular activities, the development efficiency of postgraduate attributes is not high. Configuration L_1 exhibits a consistency of 0.958 and a raw coverage of 0.632, showing that this path can explain approximately 63.2% of cases. Configuration L_2 ($\sim SR^* \sim SP^* \sim SL$) shows that in cases of a lack of scientific research ability and no social practice, even if the student leadership role is fulfilled, the development efficiency of postgraduate attributes is not high. Configuration L_2 exhibits a consistency of 0.965 and a raw coverage of 0.138, indicating that this path can explain approximately 13.8% of cases. A total of nine example cases represent this configuration. Configuration L_3 ($\sim SR^* \sim SI^* \sim SL$) shows that in cases of a lack of scientific research ability and no engagement in extracurricular activities, even if the student leadership role is fulfilled, the development efficiency of the postgraduate attribute is not high. Configuration L_3 exhibits a consistency of 0.870 and a raw coverage of 0.142, indicating that this path can explain approximately 14.2% of cases. Configuration L_4 ($\sim SR^* \sim MC^* \sim SL$) shows that in cases of a lack of scientific research ability and no outstanding achievements in moral education, even if the student leadership role is fulfilled, the development efficiency of the postgraduate attribute is not high. Configuration L_4 exhibits a consistency of 0.943 and a raw coverage of 0.158, showing that this path can explain approximately 15.8% of cases.

Comparing the coverages of the four configurations for low PAD discussed above, we find that configuration L_1 exhibits the highest coverage, explaining 63.2% of cases. That is, more than half of cases emerge because path L_1 inhibits the development efficiency of postgraduate attributes. Similarly, comparing configurations L_2 and L_3 , we find that weak social practice and weak student involvement can be substituted for each other in a context featuring weak scientific research ability and strong student leadership. Namely, there is a potential substitution relationship between configurations H_2 and H_3 . In addition, it is worth noting that the lack of scientific research ability is a core condition in all four paths leading to the low development efficiency of postgraduate attributes, which is consistent with Section 4.3 Necessity Analyses.

6. Conclusion

6.1. Findings

In this paper, a theoretical framework of input, environment, and output is established to evaluate the influence of extracurricular education with Chinese characteristics on postgraduate attributes

via involvement, leadership, social practice, moral psychology character and research ability. To gauge the developmental efficiency of this education, we employ DEA- and fsQCA-based methods to examine various causal conditions that can enhance or obstruct postgraduate attribute development. The main findings of the study are summarized as follows: Firstly, extracurricular education with Chinese characteristics is somewhat effective in promoting postgraduate attribute development, but this efficacy is relatively modest (efficiency mean = 62.15%; 50.6% of sample attained strong DEA-efficiency), and requires improvement. Secondly, our results show that four configurations consistently result in high developmental efficiency of postgraduate attributes. Students participating in extracurricular education in an environment characterized by excellent research achievement and moral psychology attributes showed the highest developmental efficiency; while academic achievement or moral award that are insufficiently outstanding, active involvement in extracurricular activities or social practice with prosocial behavior, are also associated with high developmental efficiency. This highlights the fact that participating in extracurricular activities or social practice can provide students with valuable learning opportunities and benefit them in many ways, including developing leadership, responsibility, and social skills. Furthermore, strong student involvement and strong social practice can serve as substitutes for each other in a context featuring strong research ability and a lack of leadership, suggesting that students should be strategic in their investment of time and energy in academic research and extracurricular education. Lastly, our research reveals an asymmetric causal relationship between the high and low developmental efficiency paths, indicating that conditions affecting postgraduate attribute development exhibit multiple concurrencies (Douglas et al., 2020; Zhang and Long, 2022). All hypotheses are supported except Hypothesis 3.

6.2. Implications and limitations

The four configurations consistently linked to the high development efficiency of postgraduate attributes can be refined into the output-driven model (configuration H_1) and the input-output linkage model (configurations H_2 , H_3 , and H_4), thus indicating that postgraduate attribute development is output-oriented from the perspective of students. The input-output linkage model can explain 68.6% of cases, suggesting that graduate education should combine extracurricular education and academic research to promote the postgraduate attribute development. Classroom and laboratory learning should be linked with extracurricular activities and social practice. In addition, given the importance of scientific research ability and the substitutability between strong student involvement and strong social practice, education policy-makers should remind postgraduates to allocate time and energy efficiently.

However, there are several limitations to this study that must be acknowledged. Firstly, the sample size is limited, and the participants are drawn from a science and engineering school at a double first-class university in China. Thus, the generalizability of the conclusions is limited. Future research should explore

data drawn from other countries or regions and compare our findings with those of other studies. Secondly, this study is based on students' perspectives, and other factors such as the extra support and services provided by universities and schools may lead to different outcomes. Thus, future research should look at this issue from the perspective of schools or universities. Lastly, this study is based on cross-sectional research data collected from scholarship applications during postgraduates' final year of education. However, postgraduate attribute development is a dynamic process, and longitudinal tracking should be adopted in future research.

Data availability statement

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

Author contributions

SL conceptualization, methodology, and writing—original draft. YH software, data collection, and analyzing. LW supervision and writing—review and editing. EX project administration and funding acquisition. All

authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China, grant number 7217020877 and Beijing Institute of Technology, grant number 2021ZXJG013.

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

- Abramo, G., D'Angelo, C., and Pugini, F. (2008). The measurement of Italian universities' research productivity by a non parametric-bibliometric methodology. *Scientometrics* 76, 225–244. doi: 10.1007/s11192-007-1942-2
- Afonso, C., Silva, G. M., Gonçalves, H. M., and Duarte, M. (2018). The role of motivations and involvement in wine tourists' intention to return: SEM and fsQCA findings. *J. Bus. Res.* 89, 313–321. doi: 10.1016/j.jbusres.2017.11.042
- Allen, M. S., Vella, S. A., and Laborde, S. (2015). Sport participation, screen time, and personality trait development during childhood. *Br. J. Dev. Psychol.* 33, 375–390. doi: 10.1111/bjdp.12102
- Amara, N., Rhaïem, M., and Halilem, N. (2020). Assessing the research efficiency of Canadian scholars in the management field: evidence from the DEA and fsQCA. *J. Bus. Res.* 115, 296–306. doi: 10.1016/j.jbusres.2019.10.059
- Anderson, L., Ostrom, A. L., Corus, C., Fisk, R. P., Gallan, A. S., Giraldo, M., et al. (2013). Transformative service research: an agenda for the future. *J. Bus. Res.* 66, 1203–1210. doi: 10.1016/j.jbusres.2012.08.013
- Archer, W., and Davison, J. (2008). *Graduate employability: What do employers think and want?* London: Council for Industry and Higher Education (CIHE).
- Astin, A. W. (1970). The methodology of research on college impact, part one. *Sociol. Educ.* 43, 223–254. doi: 10.2307/2112065
- Astin, A. W. (1984). Student involvement: a developmental theory for higher education. *J. Coll. Student Personnel* 25, 297–308.
- Astin, A. W., and Astin, H. S. (2000). Leadership reconsidered: engaging higher education in social change in social change. Educational Resources Information Center (ERIC) [Preprint]. Available online at: <https://files.eric.ed.gov/fulltext/ED444437.pdf> (accessed May 9, 2023).
- Bandura, A., and Walters, R. H. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Englewood cliffs Prentice Hall.
- Barnett, R. (2011). "Lifewide education: a new and transformative concept for higher education," in *Learning for a Complex World: a Lifewide Concept of Learning, Education and Personal Development*, ed. N. Jackson (Bloomington: Authorhouse).
- Barrie, S. C. (2007). A conceptual framework for the teaching and learning of generic graduate attributes. *Stud. Higher Educ.* 32, 439–458. doi: 10.1080/03075070701476100
- Bartko, W. T., and Eccles, J. S. (2003). Adolescent participation in structured and unstructured activities: a person-oriented analysis. *J. Youth Adolesc.* 32, 233–241. doi: 10.1023/A:1023056425648
- Batchelder, T. H., and Root, S. (1994). Effects of an undergraduate program to integrate academic learning and service: cognitive, prosocial cognitive, and identity outcomes. *J. Adolesc.* 17, 341–355. doi: 10.1006/jado.1994.1031
- Benson, P. L., Scales, P. C., Leffert, N., and Roehlkepartain, E. C. (1999). *A Fragile Foundation: The State of Developmental Assets Among American Youth*. Washington, DC: ERIC.
- Billig, S. H. (2004). "Heads, hearts, and hands: The research on K-12 service-learning," in *Growing to Greatness* (Redwood city, CA: Stanford University Press), 12–25.
- Bonaccorsi, A., Daraio, C., and Simar, L. (2006). Advanced indicators of productivity of universities: An application of robust nonparametric methods to Italian data. *Scientometrics* 66, 389–410. doi: 10.1007/s11192-006-0028-x
- Bowden, J., Hart, G., King, B., Trigwell, K., and Watts, O. (2000). *Generic capabilities of ATN university graduates: Final report to DETYA*, Sydney, NSW: Teaching and Learning Committee, Australian Technology Network.
- Bowden, J. L.-H., Tickle, L., and Naumann, K. (2021). The four pillars of tertiary student engagement and success: a holistic measurement approach. *Stud. High. Educ.* 46, 1207–1224. doi: 10.1080/03075079.2019.1672647
- Brandell, M. E., and Hinck, S. (1997). Service learning: connecting citizenship with the classroom. *NASSP Bull.* 81, 49–56. doi: 10.1177/019263659708159109
- Chen, L., Tao, H., and Xie, M. (2019). On the role of the extracurricular activities in talents training: focusing on the linkage between classroom learning and extracurricular activities in universities. *J. Beijing Normal Univ.* 5, 13–23.
- Chickering, A. W., and Reisser, L. (1993). *Education and Identity. The Jossey-Bass Higher and Adult Education Series*. Washington, DC: ERIC.
- Clark, B. R., and Clark, B. R. (1993). *The Research Foundations of Graduate Education: Germany, Britain, France, United States, Japan*. Berkeley, CA: University of California Press. doi: 10.1525/9780520338715

- Corrigan, M. W., Chapman, P., Grove, D., Walls, R. T., and Vincent, P. F. (2007). The importance of multidimensional baseline measurements to assessment of integrated character education models. *J. Character Educ.* 5, 103.
- CYD (2018). Comrades in charge of the School Department of the Central Committee of the Communist Youth League answer reporters' questions on the implementation of the "Second Class Transcript" system of the Communist Youth League in colleges and universities [Online]. *China Youth Daily*. Available Online at: <https://baijiahao.baidu.com/s?id=1605101396432992024&wfr=spider&for=pc> (accessed July 5, 2018).
- Darnell, C., Gulliford, L., Kristjánsson, K., and Paris, P. (2019). Phronesis and the knowledge-action gap in moral psychology and moral education: a new synthesis? *Hum. Dev.* 62, 101–129. doi: 10.1159/000496136
- De Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., et al. (2015). *Performance-Based Funding and Performance Agreements in Fourteen Higher Education Systems*. Lausanne: Center for Higher Education Policy Studies.
- De Carolis, B., D'Errico, F., Paciello, M., and Palestra, G. (2020). "Cognitive emotions recognition in e-learning: exploring the role of age differences and personality traits," in *Proceedings of the Methodologies and Intelligent Systems for Technology Enhanced Learning, 9th International Conference*, (Berlin: Springer). doi: 10.1007/978-3-030-23990-9_12
- Dewey, J. (1986). Experience and education. *Educ. Forum* 50, 241–252. doi: 10.1080/00131728609335764
- Dienhart, C., Maruyama, G., Snyder, M., Furco, A., McKay, M. S., Hirt, L., et al. (2016). The impacts of mandatory service on students in service-learning classes. *J. Soc. Psychol.* 156, 305–309. doi: 10.1080/00224545.2015.1111856
- Douglas, E. J., Shepherd, D. A., and Prentice, C. (2020). Using fuzzy-set qualitative comparative analysis for a finer-grained understanding of entrepreneurship. *J. Bus. Vent.* 35:105970. doi: 10.1016/j.jbusvent.2019.105970
- Du, Y., and Jia, L. (2017). Configuration perspective and qualitative comparative analysis (QCA): a new way of management research. *Manag. World* 6, 155–167.
- Eng, S., and Woodside, A. G. (2012). Configurational analysis of the drinking man: fuzzy-set qualitative comparative analyses. *Addict. Behav.* 37, 541–543. doi: 10.1016/j.addbeh.2011.11.034
- Erikson, E. H. (1994). *Identity and the Life Cycle*. New York, NY: WW Norton & company.
- Eyler, J., and Giles, D. (1996). "The impact of service learning program characteristics on student outcomes," in *Proceedings of the National Society for Experiential Education Conference*, (Snowbird, UT).
- Eyler, J., and Giles, D. E. Jr. (1999). *Where's the Learning in Service-Learning?* Jossey-Bass Higher and Adult Education Series. Washington, DC: ERIC.
- Fernández-García, D., Giménez-Espert, M. D. C., Castellano-Rioja, E., and Prado-Gascó, V. (2020). What academic factors influence satisfaction with clinical practice in nursing students? regressions vs. fsQCA. *Front. Psychol.* 11:585826. doi: 10.3389/fpsyg.2020.585826
- Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *Acad. Manag. Rev.* 32, 1180–1198. doi: 10.5465/amr.2007.26586092
- Fiss, P. C. (2011). Building better causal theories: a fuzzy set approach to typologies in organization research. *Acad. Manag. J.* 54, 393–420. doi: 10.5465/amj.2011.60263120
- Fredricks, J. A., and Eccles, J. S. (2006). Extracurricular involvement and adolescent adjustment: impact of duration, number of activities, and breadth of participation. *Appl. Dev. Sci.* 10, 132–146. doi: 10.1207/s1532480xads1003_3
- Gayles, J. G., Rockenbach, A. B., and Davis, H. A. (2012). Civic responsibility and the student athlete: validating a new conceptual model. *J. High. Educ.* 83, 535–557. doi: 10.1353/jhe.2012.0024
- Geller, J. D., Zuckerman, N., and Seidel, A. (2016). Service-learning as a catalyst for community development: how do community partners benefit from service-learning? *Educ. Urban Soc.* 48, 151–175. doi: 10.1177/0013124513514773
- Grau, A. J. G., and López, F. R. (2018). Determinants of migratory flow in Europe: a fuzzy-set approach. *J. Bus. Res.* 89, 243–250. doi: 10.1016/j.jbusres.2018.01.051
- Greckhamer, T., Furnari, S., Fiss, P. C., and Aguilera, R. V. (2018). Studying configurations with qualitative comparative analysis: best practices in strategy and organization research. *Strateg. Organ.* 16, 482–495. doi: 10.1177/1476127018786487
- Guilmette, M., Mulvihill, K., Villemare-Krajden, R., and Barker, E. T. (2019). Past and present participation in extracurricular activities is associated with adaptive self-regulation of goals, academic success, and emotional wellbeing among university students. *Learn. Ind. Differ.* 73, 8–15. doi: 10.1016/j.lindif.2019.04.006
- Gupta, K., Crilly, D., and Greckhamer, T. (2020). Stakeholder engagement strategies, national institutions, and firm performance: A configurational perspective. *Strateg. Manag. J.* 41, 1869–1900. doi: 10.1002/smj.3204
- Hansen, D. M., Larson, R. W., and Dworkin, J. B. (2003). What adolescents learn in organized youth activities: a survey of self-reported developmental experiences. *J. Res. Adolesc.* 13, 25–55. doi: 10.1111/1532-7795.1301006
- Hanson, T. L., Austin, G., and Lee-Bayha, J. (2003). *Student Health Risks, Resilience, and Academic Performance*. Los Alamitos, CA: WestEd.
- Harvey, L., Drew, S., and Smith, M. (2006). *The First-Year Experience: a Review of Literature for the Higher Education Academy*. London: Higher Education Academy.
- Huang, B., Huang, H., Qiao, W., and Hu, D. (2021). Graduate education development index of major countries in the world: connotation, framework and measurement. *China High. Educ. Res.* 11, 21–26.
- Jackson, N. (2008). "Tackling the wicked problem of creativity in higher education," in *Paper Presented at the Conference Briefing Paper*, (Brisbane: ARC Centre for the Creative).
- Julia, J., Supriatna, E., Isrokatun, I., Aisyah, I., Nuryani, R., and Odeode, A. A. (2020). Moral education (2010–2019): a bibliometric study (part 1). *Univ. J. Educ. Res.* 8, 2554–2568. doi: 10.13189/ujer.2020.080639
- Kahu, E. R. (2013). Framing student engagement in higher education. *Stud. High. Educ.* 38, 758–773. doi: 10.1080/03075079.2011.598505
- Kahu, E. R., and Nelson, K. (2018). Student engagement in the educational interface: understanding the mechanisms of student success. *High. Educ. Res. Dev.* 37, 58–71. doi: 10.1080/07294360.2017.1344197
- Kendrick, J. R. Jr. (1996). Outcomes of service-learning in an introduction to sociology course. *Michigan J. Commun. Serv. Learn.* 3, 72–81.
- Khademi Ashkzari, M., Piryaee, S., and Kamelifar, L. (2018). Designing a causal model for fostering academic engagement and verification of its effect on educational performance. *Int. J. Psychol.* 12, 136–161. doi: 10.24200/ijpb.2018.58146
- Koh, C. (2012). Moral development and student motivation in moral education: a Singapore study. *Aust. J. Educ.* 56, 83–101. doi: 10.1177/000494411205600107
- Kohlberg, L. (1969). "Stage and sequence: the cognitive-developmental approach to socialization," in *Handbook of Socialization Theory and Research*, eds J. C. Grusec, P. D. Hastings, and P. D. Hastings (New York, NY: Guilford Publications).
- Kolb, D. A. (2007). *The Kolb Learning Style Inventory*. Boston, MA: Hay Group, Hay Resources Direct.
- Komives, S. R., Longerbeam, S. D., Owen, J. E., Mainella, F. C., and Osteen, L. (2006). A leadership identity development model: applications from a grounded theory. *J. Coll. Stud. Dev.* 47, 401–418. doi: 10.1353/csd.2006.0048
- Kou, D. (2011). The sense of moral honor and its cultivation. *Stud. Ethics* 3, 29–33.
- Kuh, G. D. (1995). The other curriculum: out-of-class experiences associated with student learning and personal development. *J. High. Educ.* 66, 123–155. doi: 10.1080/00221546.1995.11774770
- Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., and Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *J. High. Educ.* 79, 540–563. doi: 10.1080/00221546.2008.11772116
- Larson, R. W., and Verma, S. (1999). How children and adolescents spend time across the world: work, play, and developmental opportunities. *Psychol. Bull.* 125:701. doi: 10.1037/0033-2909.125.6.701
- Leppänen, P. T., McKenny, A. F., and Short, J. C. (2019). "Qualitative comparative analysis in entrepreneurship: Exploring the approach and noting opportunities for the future," in *Standing on the shoulders of giants*, eds B. Boyd, T. R. Crook, J. K. Le, and A. D. Smith (Bingley: Emerald Publishing Limited), 155–177. doi: 10.1108/S1479-83872019000011010
- Lindsay, R. O., and Paton-Saltzberg, R. (1994). *The effects of paid employment on the academic performance of full-time students in a British new university*. Oxford: Oxford Brookes University.
- Lufi, D., and Tenenbaum, G. (1991). Persistence among young male gymnasts. *Percept. Motor Skills* 72, 479–482. doi: 10.2466/pms.1991.72.2.479
- Mahoney, J. L., Lord, H., and Carryl, E. (2005). An ecological analysis of after-school program participation and the development of academic performance and motivational attributes for disadvantaged children. *Child Dev.* 76, 811–825. doi: 10.1111/j.1467-8624.2005.00879.x
- Markus, G. B., Howard, J. P., and King, D. C. (1993). Integrating community service and classroom instruction enhances learning: results from an experiment. *Educ. Eval. Policy Anal.* 15, 410–419. doi: 10.2307/1164538
- Marshall, J. C., Caldwell, S. D., and Foster, J. (2011). Moral education the character plus Way®. *J. Moral Edu.* 40, 51–72. doi: 10.1080/03057240.2011.541770
- Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., and Wolniak, G. C. (2016). *How College Affects Students: 21st Century Evidence that Higher Education Works*. New York, NY: John Wiley & Sons.
- Meyer, C. L., Harned, M., Schaad, A., Sunder, K., Palmer, J., and Tinch, C. (2016). Inmate education as a service learning opportunity for students: preparation, benefits, and lessons learned. *Teach. Psychol.* 43, 120–125. doi: 10.1177/0098628316636278
- Mustafa, S., Sohail, M. T., Alroobaee, R., Rubaiee, S., Anas, A., Othman, A. M., et al. (2022). Éclaircissement to understand consumers' decision-making psyche and gender effects, a fuzzy set qualitative comparative analysis. *Front. Psychol.* 13:920594. doi: 10.3389/fpsyg.2022.920594
- Myers, I. B., and Myers, P. B. (2010). *Gifts Differing: Understanding Personality Type*. Boston, MA: Nicholas Brealey.

- Myers-Lipton, S. J. (1998). Effect of a comprehensive service-learning program on college students' civic responsibility. *Teach. Sociol.* 26, 243–258. doi: 10.2307/1318766
- Narvaez, D., and Lapsley, D. K. (2008). Teaching moral character: two alternatives for teacher education. *Teacher Educator* 43, 156–172. doi: 10.1080/08878730701838983
- Nonis, S. A., Philhours, M. J., and Hudson, G. I. (2006). Where does the time go? a diary approach to business and marketing students' time use. *J. Mark. Educ.* 28, 121–134. doi: 10.1177/0273475306288400
- Nucci, L. (1997). Moral development and character formation. *Psychol. Educ. Pract.* 82, 127–157.
- O'Keeffe, P. (2013). A sense of belonging: improving student retention. *Coll. Stud. J.* 47, 605–613.
- Osmani, M., Weerakkody, V., Hindi, N. M., Al-Esmail, R., Eldabi, T., Kapoor, K., et al. (2015). Identifying the trends and impact of graduate attributes on employability: a literature review. *Tertiary Educ. Manag.* 21, 367–379. doi: 10.1080/13583883.2015.1114139
- Ostrow, J. M. (1995). Self-consciousness and social position: on college students changing their minds about the homeless. *Qualit. Sociol.* 18, 357–375. doi: 10.1007/BF02393347
- Pansari, A., and Kumar, V. (2017). Customer engagement: the construct, antecedents, and consequences. *J. Acad. Mark. Sci.* 45, 294–311. doi: 10.1007/s11747-016-0485-6
- Pappas, I. O., and Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): guidelines for research practice in information systems and marketing. *Int. J. Information Manag.* 58:102310. doi: 10.1016/j.ijinfomgt.2021.102310
- Pascarella, E. T. (1985). "College environmental influences on learning and cognitive development: a critical review and synthesis," in *Higher Education: Handbook of Theory and Research*, eds M. B. Paulsen and L. W. Perna (Berlin: Springer).
- Patton, L. D., Renn, K. A., Guido, F. M., and Quaye, S. J. (2016). *Student Development in College: Theory, Research, and Practice*. Hoboken, NJ: John Wiley & Sons.
- Peck, S. C., Roeser, R. W., Zarrett, N., and Eccles, J. S. (2008). Exploring the roles of extracurricular activity quantity and quality in the educational resilience of vulnerable adolescents: variable-and pattern-centered approaches. *J. Soc. Issues* 64, 135–156. doi: 10.1111/j.1540-4560.2008.00552.x
- Perry, W. G. Jr. (1999). *Forms of Intellectual and Ethical Development in the College Years: a Scheme. Jossey-Bass Higher and Adult Education Series*. Washington, DC: ERIC.
- Podger, D. M., Mustakova-Possardt, E., and Reid, A. (2010). A whole-person approach to educating for sustainability. *Int. J. Sustainabil. High. Educ.* 11, 339–352. doi: 10.1108/14676371011077568
- Prosser, M., and Barrie, S. C. (2003). "Using a student-focused learning perspective to align academic development with institutional quality assurance," in *Towards strategic staff development in higher education*, eds R. Blackwell and P. Blackmore (Milton Keynes: Open University Press), 191–221.
- Ragin, C. C. (2009a). "Qualitative comparative analysis using fuzzy sets (fsQCA)," in *Configurational Comparative Analysis*, ed. B. Rihoux and C. Ragin. Thousand Oaks, CA: Sage Publications doi: 10.4135/9781452226569.n5
- Ragin, C. C. (2009b). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226702797.001.0001
- Ragin, C. C. (2014). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press. doi: 10.1525/9780520957350
- Rhoads, R. A. (1997). *Community Service and Higher Learning: Explorations of the Caring Self*. Albany, NY: State University of New York Press.
- Rihoux, B., and Ragin, C. C. (2008). *Configurational Comparative methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, CA: Sage Publications. doi: 10.4135/9781452226569
- Rissanen, I., Kuusisto, E., Hanhimäki, E., and Tirri, K. (2018). The implications of teachers' implicit theories for moral education: a case study from Finland. *J. Moral Educ.* 47, 63–77. doi: 10.1080/03057240.2017.1374244
- Rodgers, R. F. (1990). "Recent theories and research underlying student development," in *College Student Development: Theory Andpractice for the 1990s*, ed. D. G. Creamer & Associates (Alexandria, VA: American College Personnel Association), 27–79.
- Roeser, R. W., and Peck, S. C. (2003). Patterns and pathways of educational achievement across adolescence: a holistic-developmental perspective. *New Dir. Child Adolesc. Dev.* 2003, 39–62. doi: 10.1002/cd.81
- Salam, M., Awang Iskandar, D. N., Ibrahim, D. H. A., and Farooq, M. S. (2019). Service learning in higher education: a systematic literature review. *Asia Pac. Educ. Rev.* 20, 573–593. doi: 10.1007/s12564-019-09580-6
- Scales, P., and Roehlkepartain, E. (2004). "Service to others: A gateway asset for school success and healthy development," in *Growing to greatness* (St. Paul, MN: National Youth Leadership Council), 26–32.
- Scales, P. C., Roehlkepartain, E. C., Neal, M., Kielsmeier, J. C., and Benson, P. L. (2006). Reducing academic achievement gaps: the role of community service and service-learning. *J. Exp. Educ.* 29, 38–60. doi: 10.1177/10538259060290105
- Schneider, C. Q., and Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: a Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139004244
- Shaaban, K. (2005). A proposed framework for incorporating moral education into the ESL/EFL classroom. *Lang. Cult. Curriculum* 18, 201–217. doi: 10.1080/07908310508668742
- Shulruf, B. (2010). Do extra-curricular activities in schools improve educational outcomes? a critical review and meta-analysis of the literature. *Int. Rev. Educ.* 56, 591–612. doi: 10.1007/s11159-010-9180-x
- Shumer, R., and Belbas, B. (1996). What we know about service learning. *Educ. Urban Soc.* 28, 208–223. doi: 10.1177/0013124596028002006
- Skalicky, J., Warr Pedersen, K., van der Meer, J., Fuglsang, S., Dawson, P., and Stewart, S. (2020). A framework for developing and supporting student leadership in higher education. *Stud. Higher Educ.* 45, 100–116. doi: 10.1080/03075079.2018.1522624
- Song, D., and Zeng, J. (2018). An empirical research on the relationship among the extracurricular courses, study satisfaction and college students' core competitiveness. *Univer. Educ. Sci.* 5, 21–29.
- Stein, J. L. (2007). Peer educators and close friends as predictors of male college students' willingness to prevent rape. *J. Coll. Stud. Dev.* 48, 75–89. doi: 10.1353/csd.2007.0008
- Strage, A. (2000). Service-learning as a tool for enhancing student outcomes in a college-level lecture course. *Michigan J. Commun. Serv. Learn.* 2000, 5–13.
- Sun, L., and Li, S. (2021). Research on the impact path of artificial intelligence empowerment on innovation performance: based on fuzzy qualitative comparative analysis. *Sci. Technol. Manag. Res.* 41, 11–19.
- Terenzini, P. T., Pascarella, E. T., and Blimling, G. S. (1996). Students' out-of-class experiences and their influence on learning and cognitive development: a literature review. *J. Coll. Stud. Dev.* 40, 610–623.
- Thompson, L. J., Clark, G., Walker, M., and Whyatt, J. D. (2013). 'It's just like an extra string to your bow': exploring higher education students' perceptions and experiences of extracurricular activity and employability. *Active Learn. Higher Educ.* 14, 135–147. doi: 10.1177/1469787413481129
- Thursby, J. G. (2000). What do we say about ourselves and what does it mean? Yet another look at economics department research. *J. Econ. Literature* 38, 383–404. doi: 10.1257/jel.38.2.383
- Tieu, T.-T., and Pancer, S. (2009). Cocurricular involvement and first-year students' transition to university: quality vs. quantity of involvement. *J. First Year Exp. Stud. Transition* 21, 43–63.
- Tieu, T.-T., Pancer, S. M., Pratt, M. W., Wintre, M. G., Birnie-Lefcovitch, S., Polivy, J., et al. (2010). Helping out or hanging out: the features of involvement and how it relates to university adjustment. *Higher Educ.* 60, 343–355. doi: 10.1007/s10734-009-9303-0
- Tinto, V. (2012). *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago: University of Chicago press.
- Universities, U. (2015). "Student mental wellbeing in higher education: good practice guide," in *Proceedings of the Standing Conference of Principals*, (London: Universities UK).
- Van der Zanden, P. J., Denessen, E., Cillessen, A. H., and Meijer, P. C. (2018). Domains and predictors of first-year student success: a systematic review. *Educ. Res. Rev.* 23, 57–77. doi: 10.1016/j.edurev.2018.01.001
- Van der Zanden, P. J., Denessen, E., Cillessen, A. H., and Meijer, P. C. (2019). Patterns of success: first-year student success in multiple domains. *Stud. Higher Educ.* 44, 2081–2095. doi: 10.1080/03075079.2018.1493097
- Viau, A., and Poulin, F. (2015). Youths' organized activities and adjustment in emerging adulthood: a multidimensional conception of participation. *J. Res. Adolesc.* 25, 652–667. doi: 10.1111/jora.12159
- Vis, B. (2012). The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociol. Methods Res.* 41, 168–198. doi: 10.1177/0049124112442142
- Wegner, G. R. (2008). *Partnerships for Public Purposes: Engaging Higher Education in Societal Challenges of the 21st Century*. National Center for Public Policy and Higher Education. Paris: UNESCO
- Weiler, L., Haddock, S., Zimmerman, T. S., Krafchick, J., Henry, K., and Rudisill, S. (2013). Benefits derived by college students from mentoring at-risk youth in a service-learning course. *Am. J. Commun. Psychol.* 52, 236–248. doi: 10.1007/s10464-013-9589-z
- Wolf-Wendel, L. E., and Ruel, M. (1999). Developing the whole student: the collegiate ideal. *New Dir. Higher Educ.* 1999, 35–46. doi: 10.1002/he.10503

- Woodside, A. G. (2014). Embrace perform model: complexity theory, contrarian case analysis, and multiple realities. *J. Bus. Res.* 67, 2495–2503. doi: 10.1016/j.jbusres.2014.07.006
- Wren, J. T. (2013). *The Leader's Companion: Insights on Leadership Through the Ages*. New York, NY: Simon and Schuster.
- Xiao, Y., and Ren, Y. (2021). A study on the enhancement of college students' entrepreneurial ability from the erspective of the second classroom. *Party Build. Ideol. Educ. Schools* 2, 61–63.
- Yuan, B., Xu, S., Chen, L., and Niu, M. (2022). How do psychological cognition and institutional environment affect the unsafe behavior of construction workers?—research on fsQCA method. *Front. Psychol.* 13:875348. doi: 10.3389/fpsyg.2022.875348
- Zajda, J., and Rust, V. (2016). “Current research trends in globalisation and neo-liberalism in higher education,” in *Globalisation and Higher Education Reforms*, eds J. Zajda and V. Rust (Berlin: Springer). doi: 10.1007/978-3-319-28191-9
- Zepke, N. (2014). Student engagement research in higher education: questioning an academic orthodoxy. *Teaching Higher Educ.* 19, 697–708. doi: 10.1080/13562517.2014.901956
- Zepke, N., Leach, L., and Butler, P. (2011). Non-institutional influences and student perceptions of success. *Studi. High. Educ.* 36, 227–242. doi: 10.1080/03075070903545074
- Zhang, H., and Long, S. (2022). How business environment shapes urban tourism industry development? configuration effects based on NCA and fsQCA. *Front. Psychol.* 13:947794. doi: 10.3389/fpsyg.2022.947794
- Zhao, Z., and Tan, C. (2007). “Citizenship education in mainland China at a crossroads,” in *Critical perspectives on values education in Asia*, eds C. Tan and K. Chong (Singapore: Prentice Hall, Pearson Education South Asia), 124–137.
- Zhou, Z., and Wang, H. (2018). “Study on scientific research and innovation oriented cultivation model of postgraduate,” in *Proceedings of the 2018 International Seminar on Education Research and Social Science (ISERSS 2018)*, (Amsterda: Atlantis Press), 252–255. doi: 10.2991/iserss-18.2018.65
- Zhu, J., Cai, K., and Yao, Q. (1983). *Management of Institutions of Higher Education*. Wuhan: Huazhong University of Science & Technology Press.
- Zimmerman-Oster, K., and Burkhardt, J. (1999). *Leadership in the Making: Impact and Insights from Leadership Development Programs in US Colleges and Universities*. Battle Creek, MI: WK Kellogg Foundation.
- Zou, M., and Tan, D. (2013). A literature review o input-output evaluation of student learning. *Educ. Measur. Eval.* 3, 10–14.



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

David Zamorano-García,
University of Castilla-La Mancha, Spain
Inês Peixoto Silva,
Instituto Politécnico de Viana do Castelo,
Portugal
María Martínez López,
University of Castilla-La Mancha, Spain

*CORRESPONDENCE

Ann-Christin Sollerhed
✉ ann-christin.sollerhed@hkr.se

RECEIVED 31 May 2023

ACCEPTED 07 July 2023

PUBLISHED 27 July 2023

CITATION

Sollerhed A-C (2023) Factors associated with
implemented teacher-led movement and
physical activity in early childhood education
and care.

Front. Psychol. 14:1221566.
doi: 10.3389/fpsyg.2023.1221566

COPYRIGHT

© 2023 Sollerhed. 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.

Factors associated with implemented teacher-led movement and physical activity in early childhood education and care

Ann-Christin Sollerhed*

Faculty of Teacher Education, Kristianstad University, Kristianstad, Sweden

Movement and physical activity (MoPA) is critical for children's development and health. This study aimed to explore early childhood education and care (ECEC) educators' reported frequency of implemented gross motor and physical activities (MoPA) among children in ECEC, as well as the educators' reported personal physical activity (PA) levels in leisure time. A cross-sectional survey was performed in 68 preschools in southern Sweden. Data were obtained from questionnaires completed by 359 ECEC educators. The participation rate was 61%. About two thirds offered MoPA once a week or more seldom, while one quarter offered MoPA at least every other day. Educators who reported personal PA three times or more per week, offered MoPA for the children at least every other day to a higher extent (37%) compared to colleagues who reported personal PA once or twice a week (26%) or colleagues who reported that they were never or seldom active (18%) ($p = 0.034$). The results from multiple logistic regression analysis showed that reported implemented MoPA among children in ECEC was significantly associated with the educators' perceptions that free play improved children's gross motor skills (OR 2.7), the educators' perceptions of needed curricular guidelines for MoPA (OR 2.1), the educators' own leisure PA level (OR 2.0) and the educators' perceptions that adequate gross motor skills were not learned at home (OR 0.4). Teacher-led MoPA occurs sparingly during the preschool day and the teachers believe that the children get sufficient MoPA in free play. The children are expected to develop their motor skills to a sufficient extent during the short moments of offered outdoor play. Teachers who are physically active in their leisure-time seem to offer gross motor training for the children to a higher extent than less active or inactive colleagues.

KEYWORDS

early childhood, motor skills, motor development, physical activity, teacher led activity, free play

1. Introduction

The 21st century has seen steady growth in early childhood education and care (ECEC) to support parents with young children's education and care (Cohen and Korintus, 2017), and ECEC is the first step into the formal education system (Broström et al., 2018). In 2021, 86% of children of preschool age (1–5 years) in Sweden participated in ECEC. On average, children spent 31 h a week in ECEC (Skolverket, 2021).

As most children spend a large proportion of their waking hours in ECEC, the educators are important for promoting, teaching, and modeling children's physical activity (PA) (Mavilidi et al., 2021), which in turn is important for children's motor development and to ensure they have the daily recommended PA. Early childhood development lays the foundation for a lifetime's mental and physical health, education, labor market productivity, and well-being (Shonkoff, 2014; Richter et al., 2020). PA offers significant physical health benefits in children, and prevents chronic diseases such as cardiovascular disease, diabetes, or osteoporosis later in life (Carson et al., 2017). Moreover, there seems to be a significant association between PA and cognitive and academic performance in children. PA improve physical self-perceptions and enhance self-esteem, which seem to be mechanisms that affect cognitive and mental health in children (Lubans et al., 2016). Activities with physical exertion (high intensity) seem to have the strongest effects on cognition, specifically on executive functions (Vorkapic et al., 2021).

Participation in PA among children under the age of five has declined globally in recent years, while childhood obesity is, worryingly, increasing (World Health Organization, 2019). The WHO recommends at least 180 min of PA per day at any intensity level, including at least 60 min of moderate-to-vigorous physical activity (MVPA), and reduced time sitting for children under the age of 5 years (World Health Organization, 2019). ECEC is regulated by governing documents for example in law and curriculum. The responsibility to organize pedagogical activities to promote the principles stated in the curriculum lies with the educators, yet specific pedagogical methods are not identified in the curriculum (Einarsdottir et al., 2015). MoPA was found to have a low priority, to varying degrees, in the ECEC policies enacted by Nordic countries, and the guidance provided to educators and stakeholders therein was inexplicit (Sollerhed et al., 2021), which may mean that educators do not implement MoPA for young children in ECEC.

Fundamental Movement Skills (FMS), defined as basic learned movement patterns, which include gross motor skills such as stability, locomotor skills, and object control (Barnett Lai et al., 2016), are vital for motor development in children (Eddy et al., 2019). The FMS can be regarded as building blocks for movement and are essential for learning more complex skills (Clark and Metcalf, 2002), which in turn may lead to higher PA levels (Hardy et al., 2010a; Lubans et al., 2010), especially in MVPA (Williams et al., 2008).

Various types of teachers' strategies and qualifications to handle MoPA have been identified in studies as well as MoPA-related outcomes among children (Mak et al., 2021). Teacher training in MoPA practice is essential to deliver adequate MoPA for children in ECEC (Vanderloo et al., 2014; Mavilidi et al., 2021). The ECEC educators meet children who are in a period of rapid brain development, growth, and motor development (Shonkoff, 2014), a period when also healthy behaviors are established (Ward et al., 2010). ECEC educators are agents who create conditions for children's learning in preschool practice (Dalli et al., 2012). This includes the teaching of adequate MoPA, which is essential for MVPA levels (Barnett Lai et al., 2016). Adequate teacher led programs have been shown to have the potential to improve motor skills in young children (Adams et al., 2009; Hardy et al., 2010b; Van Cauwenbergh et al., 2013). An adequate education in teaching MoPA is important. Teacher training could positively influence perceptions and attitudes to

increase the professional educator's perceived competence when implementing PE in ECEC (Soini et al., 2021).

Parents often think that children are highly active in preschool, and thus offer few opportunities for MoPA at home (Pate et al., 2008), but children were observed to be physically inactive during most of their time in ECEC, and physically active in MVPA only 2–3% of the time (Pate et al., 2008; Soini et al., 2014). Sedentary time has been shown to be high in ECEC (Reilly, 2010). Another study showed that mean MVPA in preschool children was about 16 min per day. It takes at least twice that amount to achieve the positive effects of bone mineralization (Raustorp et al., 2012).

The social and physical environment in ECEC has an important influence on children's MoPA (Brown et al., 2009a). The educators are role models, and they need adequate competence in MoPA to provide opportunities for children to learn motor skills. Children's motor skills have been shown to lead to better immediate and long-term health and education outcomes (Shonkoff, 2014). The educators' behavior and attitudes are the least studied correlates when it comes to correlates of children's MoPA and/or sedentary behavior in ECEC (Tonge et al., 2016).

Therefore, the aim of this study was to explore ECEC educators' reported frequency of implemented gross motor and physical activities (MoPA) among children in ECEC during the preschool day, their opinions about learning situations for motor skills in ECEC, as well as the educators' reported personal PA levels in leisure time. Additional aims were to explore the educators' perceived competence level to teach MoPA and how they perceived the occurrence of guidelines for MoPA in the policy documents for ECEC.

2. Methods and materials

2.1. Participants

The study took place in 68 preschools in southern Sweden. The sample represented preschools from the countryside and mid-size municipalities. The preschool units consisted of both mixed and single age groups. On average, teacher-child ratios were 1:5 with ranges of 1:3 to 1:8. Three hundred fifty-nine in-service ECEC educators completed the self-report questionnaire (response rate 61%). The participants (aged 20–65 years old) worked as educators in ECEC and were either preschool teachers or day-care attendants. About half of all ECEC employees in Sweden are preschool teachers that have 3.5 years of university training and the other half are day-care attendants who have upper-secondary qualifications (Sandberg and Ottosson, 2010).

2.2. The survey

A questionnaire was distributed to individual ECEC educators to obtain statistically useful information about their experiences and perceptions about MoPA. A literature search did not show any previous studies related to MoPA in ECEC in Sweden and therefore the questions of the current questionnaire were developed by the research group specifically for the aim of this study. The questions in the questionnaire were developed from the perspective of content in the curriculum for ECEC, completed

with questions about PA behavior that have been used in other studies (Sollerhed, 2006). The questionnaire was piloted with 47 preservice teachers in ECTE outside the study before it was used in the actual survey. The questionnaire consisted of either multiple-choice questions or Likert scale questions. The questions were distributed in three parts: (1) Participants' background information; (2) Participants' experiences and perceptions of their own PA behavior; (3) Participants' experiences and perceptions of MoPA in ECEC. Participants were asked to evaluate their experiences and perceptions of different activities and aspects in ECEC on a scale from 1 = completely disagree to 4 = completely agree. The teachers' perception of children's MoPA at home was also investigated in a similar Likert scale question. The data collection was performed at each preschool and the participants filled in the survey on paper individually in connection to their working day.

2.3. Statistical analysis

SPSS Statistics version 28 software was used to carry out both descriptive and analytical statistics. Data were first analyzed using descriptive statistics with frequencies and percentages. Secondly, Chi-square tests were used to investigate any associations between the frequency of implemented gross motor teaching in the preschools and the independent variables of educators' personal MVPA, perceived education level in MoPA, perceived competence to teach MoPA,

perceptions of children's learned motor skills at home, at preschool and in free play, perceptions of the actual curriculum, wish for guidelines for MoPA, and enjoyment of teaching MoPA.

Finally, multiple logistic regression analysis was conducted to examine any association between the described independent variables and the likelihood of reported implemented gross motor teaching among children in the preschool. The rationale for using a logistic regression analysis is that it allows us to see the effects of variables after adjusting for other variables. This helps us to see and verify if the associations seen in the bivariate analysis are due to the influence of other variables. The method "enter" was used in the logistic regression analysis, which is a procedure for variable selection in which all independent variables are entered in a single step. The assumption of absence of multicollinearity was examined among the variables by calculating variance inflation factors (Menard, 2010). The responses in dependent and independent variables were collapsed and dichotomized before the multiple logistic regression analyses. The coding and dichotomizing of response options are shown in Table 1. The significance level was set at $p < 0.05$.

2.4. Ethical consideration

The study was conducted following the ethical principles for research involving human subjects, and all procedures were in accordance with the Declaration of Helsinki and the Swedish law on research ethics (SFS:2003:460). Ethical review application was

TABLE 1 Dichotomized variables included in the logistic regression with reported frequency of implemented gross motor skills training as the dependent variable.

Item	Response options	Dichotomization
Reported personal PA "How often do you exercise in your free time for at least half an hour so that you become short of breath and sweaty?"	7 categories Never (1) → > 4 times/week (7)	PA ≥ 3 times/week (6–7) PA 2 times or less/week (1–5)
Adequate MoPA in free play "Children learn adequate gross motor skills in free play."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Adequate MoPA at home "Children learn adequate gross motor skills at home."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Adequate MoPA learning in ECEC "Children learn adequate motor skills through organized teaching in preschool."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Curriculum supportive for MoPA "The curriculum provides sufficient support for gross motor training in preschool."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Adequate competence for MoPA teaching "I have adequate competence to lead and teach gross motor skills training."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Adequate education for MoPA teaching "I have an adequate education for work with gross motor skills teaching."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Guidelines for MoPA in ECEC "I am positive towards guidelines for adequate gross motor training for children in preschool."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)
Enjoy teaching MoPA "I enjoy and find pleasure in teaching children gross motor skills."	4 categories Totally agree (1) → Totally disagree (4)	Totally agree (1) Disagree to some extent or totally disagree (2–4)

approved by the Regional Ethical Review Committee in Lund (Dnr:2017/555). The participants were informed about the study, their voluntary status, and confidentiality, and informed written consent was obtained from all participants.

3. Results

As shown in Table 2, about one quarter of the educators reported that they offered gross motor skills training for the children every other day or every day, while about two thirds of the educators reported that they offered gross motor training once a week or more seldom. About a fifth of the educators reported that they offered MVPA every other day or every day, while four fifths reported MVPA once a week or more seldom. The reported personal level of MVPA differed between the educators. About 30 percent of the educators reported MVPA very seldom or never, while about 40 percent reported regular MVPA once or twice a week and about 30 percent reported MVPA three or more times a week (Table 2).

The results of the bivariate analyses between reported implemented gross motor training in preschool and different variables are shown in Table 3. The results showed significant associations between the educators' reported personal MVPA in leisure time and the reported frequency of allocated gross motor training for the children in the preschool. The educators who reported MVPA three times or more per week, reported that they offered gross motor training for the children at least every other day to a higher extent (37%) compared to colleagues who reported personal MVPA once or twice a week (26%) or colleagues who reported they were never or seldom active (18%) ($p=0.034$). There was also a significant association between the educators' opinion that children learn sufficient gross motor skills in free play and the reported frequency of allocated gross motor training ($p<0.001$) (Table 3).

The results from the multiple logistic regression analysis showed that reported implemented gross motor training among children in the preschool was significantly associated with the educators' perception that free play improved children's gross motor skills adequately (OR 2.7; CI 1.3–5.5), the educators' perceptions of needed guidelines for MoPA in the ECEC curriculum (OR 2.1; CI 1.1–4.1), the educators' reported personal leisure PA level (OR 2.0; CI 1.9–3.8) and the educators' perception that adequate gross motor skills were not learned at home (OR 0.4; CI 0.2–0.9) (Table 4).

4. Discussion

The findings of our study among 359 ECEC educators, revealed that about two thirds of them reported that they offered organized MoPA sessions once a week or more seldom, while less than one fifth of the educators offered MoPA daily for the children. The question is why few occasions are offered in general. The answer seems to be multifactorial and complex. Other studies have shown different explanations for barriers to incorporating MoPA in ECEC. One explanation was shown to be the desire for educators to favor work for academic schooling over MoPA training (Reilly, 2010). Another explanation could be the educators' lack of competence to teach MoPA. It has been shown in other studies that it could be due to the educators' personal attitudes and self-efficacy in MoPA-related issues

(Parks et al., 2007; Copeland et al., 2012; Webster et al., 2015), as well as insufficient pedagogical content knowledge to teach MoPA, fear of injury and their own low fitness levels (Sollerhed, 2023). Educators' insufficient understanding of the value and benefits of MoPA, which contribute to overall development, including cognitive and academic achievements (Lu and Montague, 2016) also affects the implemented MoPA.

The playground and the outdoor environment, including spending time outdoors, are often pointed out as crucial for children's MoPA (Bower et al., 2008), especially if the educators are asked about prerequisites for MoPA. The effects of educators' underlying attitudes about MoPA and actions are seldom pointed out, especially not by the educators themselves. Individual educators make daily decisions about when and how MoPA should be integrated or denied for children in ECEC. The policy documents regulate ECEC activities and should guide educators in their work. MoPA was shown to be of low priority in the ECEC policies, and the governmental guidance provided to educators and stakeholders therein is inexplicit (Sollerhed et al., 2021), which can increase the impact from the educator's attitudes, perceived competence and habitus. The importance of educators' behaviors toward children's MoPA, including positive or negative prompts and modeling, has been shown to be vital (Brown et al., 2009b).

In our study, it was shown that the educators' statements that children learned motor skills sufficiently through free play in ECEC was associated with the reported level of implementation of MoPA. The educators can act either as facilitators for or barriers to children's motor development. This cannot be ignored from a public health perspective. Children spend a significant amount of their waking hours in ECEC, and educators are key players in shaping young children's active behavior, which is important not only for children's actual development but also for lifelong healthy behavior. According to the ECEC educators' answers in the questionnaire, the likelihood for reported implemented MoPA at least every other day was almost three times higher when the educators perceived that the children learned motor skills in free play. It can thus be interpreted that the implemented MoPA was usually in the form of free play and not in the form of intentional teacher-led activity. Free play might offer efficient MoPA for the children, but it is not a guarantee. If the children have access to an enriching play area, they might choose to be physically active, and the free play might give the children the possibility to develop motor skills and endurance capacity. On the other hand, children can choose to do sedentary activities instead and exclude MoPA and learning motor skills in the free play period. In short, free play does not guarantee the effects of MoPA and therefore it is important to have adequately led MoPA sessions to ensure that all children get a chance to develop FMS. Besides general motor development, in terms of health, the importance of increased PA for influencing children's (0–5 years) adiposity, bone density and cardiometabolic factors must be emphasized (Timmons et al., 2012).

Several studies have shown that children generally spend less than 50 percent of a free play period in ECEC participating in MoPA (Verstraete et al., 2006), especially when the restrictions are comprehensive to exclude any risks. Structured and intentional MoPA sessions have been shown to substantially increase the total amount of children's PA in comparison with the amount in free play (Frank et al., 2018), as well as to enhance the FMS (Barnett Lai et al., 2016). The reason for the educators' strong beliefs that children are highly active in MoPA in free play and their choice to often let the children have free play

TABLE 2 Distribution of ECEC educators' answers to questions and statements in the questionnaire.

Item	Options	Frequency <i>n</i> (%)
Planned gross motor training “How often does planned teacher led gross motor training occur in your preschool?”	Every day Every other day Once a week More seldom than once a week	57 (16.9) 37 (10.9) 185 (54.7) 59 (17.5)
Planned MVPA “How often does planned MVPA occur in your preschool?”	Every day Every other day Once a week More seldom than once a week	39 (11.7) 30 (9.0) 160 (48.2) 100 (30.1)
Spent time outdoor “How often do the children spend time outdoors?”	Every day Every other day Once a week More seldom than once a week	346 (98.9) 2 (0.6) 2 (0.6) -
Walking a little longer distance “How often does the group of children walk a longer distance?”	Every day Every other day Once a week More seldom than once a week	11 (3.2) 38 (11.1) 202 (59.2) 90 (26.5)
Categories of staff members “What is your highest education?”	Day-care attendant Preschool teacher Other	107 (30.2) 206 (58.2) 41 (11.6)
Reported personal PA “How often do you exercise in your free time for at least half an hour so that you become short of breath and sweaty?”	Never A few times a year A few times a month Regularly once a week Regularly twice a week Regularly three times a week Regularly four times or more a week	10 (2.8) 23 (6.4) 66 (18.6) 59 (16.5) 84 (23.5) 71 (19.9) 44 (12.3)
Adequate MoPA in free play “Children learn adequate gross motor skills in free play.”	Totally agree Partially agree Partially disagree Totally disagree	218 (62.3) 128 (36.6) 2 (0.6) 2 (0.6)
Adequate MoPA at home “Children learn adequate gross motor skills at home.”	Totally agree Partially agree Partially disagree Totally disagree	58 (17.7) 214 (65.2) 52 (15.9) 4 (1.2)
Adequate MoPA learning in ECEC “Children learn adequate motor skills through organized teaching in preschool.”	Totally agree Partially agree Partially disagree Totally disagree	214 (61.3) 127 (36.4) 7 (2.0) 1 (0.3)
Curriculum supportive for MoPA “The curriculum provides sufficient support for gross motor training in preschool.”	Totally agree Partially agree Partially disagree Totally disagree	99 (29.5) 176 (52.4) 51 (15.2) 10 (3.0)
Adequate competence for MoPA teaching “I have adequate competence to lead and teach gross motor skills training.”	Totally agree Partially agree Partially disagree Totally disagree	223 (64.1) 118 (33.9) 6 (1.7) 1 (0.3)
Adequate education for MoPA teaching “I have an adequate education for work with gross motor skills teaching.”	Totally agree Partially agree Partially disagree Totally disagree	90 (28.7) 135 (43.0) 72 (22.9) 17 (5.4)
Guidelines for MoPA in ECEC “I am positive towards guidelines for adequate gross motor training for children in preschool.”	Totally agree Partially agree Partially disagree Totally disagree	189 (54.8) 133 (38.6) 19 (5.5) 4 (1.2)

TABLE 3 Relationship between ECEC educators' reported frequency of gross motor training in ECEC and reported personal PA, and opinions about MoPA.

	Gross motor training every day	Gross motor training every other day	Gross motor training once a week or more seldom	P-value*
Reported personal PA (%)				
3 times or more/week	20.4	16.8	62.8	0.034
1–2 times/week	16.8	9.5	73.7	
Seldom or never	12.5	5.7	81.8	
Adequate MoPA in free play (%)				
Agree	22.3	11.0	66.7	<0.001
Disagree	5.5	11.2	80.3	
Adequate MoPA at home (%)				
Agree	14.0	7.0	78.9	0.597
Disagree	15.2	11.3	73.5	
Adequate MoPA learning in ECEC (%)				
Agree	18.0	12.2	69.8	0.241
Disagree	12.5	9.4	78.1	
Adequate competence for MoPA teaching (%)				
Agree	20.1	11.0	68.9	0.061
Disagree	10.4	9.6	80.0	
Adequate education for MoPA teaching (%)				
Agree	18.0	11.2	70.8	0.721
Disagree	15.0	9.8	75.2	
Guidelines for MoPA in ECEC (%)				
Agree	17.6	12.2	70.2	0.241
Disagree	14.2	8.8	77.0	
Enjoy teaching MoPA (%)				
Agree	18.9	12.4	68.7	0.068
Disagree	12.2	6.8	81.0	

*Chi-square test. P-value in bold when significant.

instead of various teacher-led MoPA is unclear, but is probably related to their perceived insufficient competence to lead and instruct MoPA (Sollerhed, 2023), but may also be associated with indolence. In children's free play, the educators often take the position of an inactive supervisor (Tandon et al., 2018), which might be perceived as a more convenient task compared to the position as an instructor with the responsibility to teach the children different techniques in MoPA. The educators may use free play perfunctorily and take the easiest way with the least energy-demanding effort for themselves. Adequate educator-led programs have the potential to improve motor skills in young children (Adams et al., 2009; Hardy et al., 2010b; Van Cauwenbergh et al., 2013), which assumes that the educators are more actively participating in the activities and are not just supervising them. Also the intensity of PA among the children heightens during teacher led FMS practice, especially when involving locomotor skills (Cliff et al., 2009; Kain et al., 2018).

In free play, children are given access to certain areas in the environment with or without restrictions. The educators are the gatekeepers to available play areas and maintainers of restrictions. There is a strong belief that children should be kept as safe as possible, and the increasing focus on safety affects restrictions on children's

freedom and possibility to develop motor skills in the play area. Outdoor play may offer great opportunities for children to use natural elements and may involve challenges, heights, and speed, including exploring the environment and taking risks.

As risky play is associated with injury it is highly limited with subsequent limited MoPA training (Brussoni et al., 2012, 2015; Bento and Dias, 2017). Risks of injury must be considered in ECEC, but the children's living conditions must not be de-risked to the extent that they do not learn to develop their own risk assessment, which is a normal part of childhood and development (Wyver et al., 2010; Tremblay et al., 2015). In ECEC, children are under adult supervision most of the time and the adults frequently decide what children are allowed to do and where they are allowed to play (Kytä, 2004). Thus, preschoolers are under the influence of the preferences and fears of the educators. Adult-imposed restrictions on risk-taking in free play and MoPA restrictions are often justified for safety reasons (Brussoni et al., 2015), with children being banned from a variety of experiences of MoPA and from accessing forbidden areas (Thomson, 2014). The reason behind heightened sensitivity to risk in children's play is a fundamental professional dilemma experienced by practitioners on

TABLE 4 Variables associated with the educators' reported gross motor teaching in a multiple logistic regression analysis ($n = 348-359$).

Variable	<i>B</i>	SE	χ^2	<i>P</i> *	OR	95% CI
Reported personal PA	0.709	0.318	4.958	0.026	2.03	1.089–3.794
Adequate MoPA in free play	0.990	0.368	7.226	0.007	2.693	1.308–5.544
Adequate MoPA at home	−0.918	0.425	4.660	0.031	0.399	0.174–0.919
Adequate MoPA learning in ECEC	0.396	0.342	1.339	0.247	1.486	0.760–2.905
Curriculum supportive for MoPA	−0.228	0.359	0.403	0.525	0.796	0.394–1.608
Adequate competence for MoPA teaching	0.240	0.384	0.391	0.532	1.272	0.599–2.700
Adequate education for MoPA teaching	−0.017	0.350	0.002	0.962	0.983	0.496–1.951
Guidelines for MoPA in ECEC	0.730	0.346	4.444	0.035	2.075	1.053–4.091
Enjoy teaching MoPA	0.032	0.398	0.007	0.935	1.003	0.474–2.252

**P*-value in bold when significant. $\chi^2(8) = 273.24$, $p < 0.001$, Hosmer–Lemeshow, $p = 0.414$, Nagelkerke $R^2 = 0.125$. Variance Inflation Factors (VIFs) were calculated to detect the presence of multicollinearity between independent variables. All independent variables in the regression model have VIFs of 1.01–1.06.

a day-to-day basis. There is a dilemma in dealing with children's safety; on the one hand, it is important that the children do not suffer acute injuries, but on the other hand, it is important not to prevent the children's participation and exploration in MoPA, which can affect their motor and physical development negatively in the longer term. Thus, daily practice is fraught with conflicting priorities between safety and development imperatives. The educators should be encouraged to allow children to experience risk in their play by evaluating the developmental benefits of risk-taking and by having increased awareness of the value of risk-taking (Sandseter, 2014). While the fear of injuries prevents MoPA in general, perceived lack of sufficient pedagogical content knowledge in MoPA prevents intentional MoPA sessions where FMS is trained (Sollerhed, 2022).

In our study, a high percentage of the educators were convinced that the children learned adequate motor skills in free play in ECEC, while far fewer were convinced that children learned adequate motor skills at home. As children engage in free play both in the preschool context and at home this seems to be a paradox. The study cannot explain this discrepancy.

An association between the educator's reported personal PA level and the reported frequency of implemented MoPA for the children in ECEC was shown. The educators' own interest in PA seems to be an important determinant of the possibility for children's MoPA in ECEC. About one third of the participating educators in our study reported that they were physically active in their leisure time. The likelihood that intentional MoPA would be implemented for the children in ECEC was doubled if the educator reported personal PA in their own leisure time. Educators' beliefs and experiences of their own PA could thus be seen as a possible explanation for integrated MoPA among children. It is positive that these physically active educators implement MoPA sessions for the children in ECEC, but on the other hand, it could be questioned if the occurrence of MoPA among children is due to educators' personal attitudes. Many teachers value MoPA but lack confidence to teach it. Teachers may believe in the benefits of PA, but would rather teach other subjects (McBride et al., 2002; Morgan and Hansen, 2008).

Interestingly, educators' own experiences and perceptions about PA seem to be determinants for promotion and teaching of MoPA in children in ECEC, despite existing curriculum. The educators'

actions are not always performed in accordance with the curriculum, which is the governmental guidance provided to educators. Human behavior and actions are not performed in accordance with rules, but are often based on mental states, such as beliefs, attitudes, desires, goals and intentions (Gibbs, 2001), and also in habitus and ingrained habits (Bourdieu, 2017). MoPA should be legitimized and implemented from the perspective of children's development irrespective of educators' own past experiences and habitus for PA. The educators' wish for clearer guidelines for MoPA in the curriculum for ECEC was significantly associated with the reported implemented teacher led MoPA. The educators perceived that MoPA was insufficiently and vaguely mentioned in the curriculum, and they wished for more explicit guidance through the curriculum. MoPA was shown to be of low priority in the ECEC curricula (Sollerhed et al., 2021), which strengthened the educators' reflections.

The strengths of this study are worth mentioning, but also the limits. The study had a relatively large sample size, and the response rate was acceptable. However, a skewness in the positive direction could be noted. The survey may have attracted those educators interested in MoPA while those who felt resistance toward MoPA did not participate to the same level. Despite a sense of being skewed in a positive direction, the results showed that MoPA seems to occur sparingly in ECEC and seems to depend on the teachers' benevolence and self-interest. The study was a cross-sectional survey and was unable to establish causality. As a literature search did not show any previous studies related to MoPA in ECEC in Sweden the questions in the questionnaire were developed by the research group, which can be seen as a weakness. The questions were developed on basis of the perspective of the content in the curriculum for ECEC, which could be seen as a strength. The PA was self-reported and not objectively measured, which is another limitation. The questions about reported PA have been used in several studies before and have shown to have acceptable validation. The questionnaire was piloted outside the study before it was used in the actual survey.

5. Conclusion

The teacher led MoPA seem to occur sparingly during the preschool day and the educators believe that the children get sufficient

PA and develop adequately the gross motor skills in free play. The children are expected to develop their gross motor skills adequately during the short moments of offered outdoor play. Educators who report that they are physically active in PA three times or more per week in their leisure-time seem to offer gross motor training for the children to a higher extent than less active or inactive colleagues. Thus, educators' own experiences and perceptions about PA seem to be determinants for promotion and teaching MoPA in ECEC.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the author upon request.

Ethics statement

The studies involving human participants were reviewed and approved by Regional Ethical Review Committee in Lund (Dnr:2017/555). The participants provided their written informed consent to participate in this study.

Author contributions

A-CS contributed to the design of the study. A-CS contributed to the data collection and interpretation of the data, drafted and revised the manuscript, provided final approval, and agreed to be accountable for all aspects of the work to ensure integrity and accuracy.

References

- Adams, J., Zask, A., and Dietrich, U. (2009). Tooty fruity veggie in preschools: an obesity prevention intervention in preschools targeting Children's movement skills and eating behaviours. *Health Promot. J. Austr.* 20, 112–119. doi: 10.1071/HE09112
- Barnett Lai, S. K., Veldman, S. L., Hardy, L. L., Cliff, D. P., Morgan, P. J., Zask, A., et al. (2016). Correlates of gross motor competence in children and adolescents: a systematic review and meta-analysis. *Sports Med.* 46, 1663–1688. doi: 10.1007/s40279-016-0495-z
- Bento, G., and Dias, G. (2017). The importance of outdoor play for young children's healthy development. *Porto Biomed. J.* 2, 157–160. doi: 10.1016/j.pbj.2017.03.003
- Bourdieu, P. (2017). *Habitus: a sense of place* Routledge. e-Book ISBN 9781315253701
- Bower, J. K., Hales, D. P., Tate, D. F., Rubin, D. A., Benjamin, S. E., and Ward, D. S. (2008). The childcare environment and children's physical activity. *Am. J. Prev. Med.* 34, 23–29. doi: 10.1016/j.amepre.2007.09.022
- Broström, S., Einarsdottir, J., and Samuelsson, I. P. (2018). *The Nordic perspective on early childhood education and care*. Nordrecht, Netherlands (NL): Springer.
- Brown, W. H., Googe, H. S., Mciver, K. L., and Rathel, J. M. (2009a). Effects of teacher-encouraged physical activity on preschool playgrounds. *J. Early Interv.* 31, 126–145. doi: 10.1177/1053815109331858
- Brown, W. H., Pfeiffer, K. A., Mciver, K. L., Dowda, M., Addy, C. L., and Pate, R. R. (2009b). Social and environmental factors associated with preschoolers' nonsedentary physical activity. *Child Dev.* 80, 45–58. doi: 10.1111/j.1467-8624.2008.01245.x
- Brussoni, M., Gibbons, R., Gray, C., Ishikawa, T., Sandseter, E. B. H., Bienenstock, A., et al. (2015). What is the relationship between risky outdoor play and health in children? A systematic review. *Int. J. Environ. Res. Public Health* 12, 6423–6454. doi: 10.3390/ijerph120606423
- Brussoni, M., Olsen, L. L., Pike, I., and Sleet, D. A. (2012). Risky play and Children's safety: balancing priorities for optimal child development. *Int. J. Environ. Res. Public Health* 9, 3134–3148. doi: 10.3390/ijerph9093134
- Carson, V., Lee, E.-Y., Hewitt, L., Jennings, C., Hunter, S., Kuzik, N., et al. (2017). Systematic review of the relationships between physical activity and health indicators in the early years (0–4 years). *BMC Public Health* 17, 33–63. doi: 10.1186/s12889-017-4860-0
- Clark, , and Metcalf, J. (2002). "The mountain of motor development: a metaphor" in *Motor development: Research and reviews*. eds. I. N. J. Clark and J. Humphrey, vol. 2 (Reston, US: National Association Of Sport And Physical Education), 163–190.
- Cliff, D. P., Okely, A. D., Smith, L. M., and Mckeen, K. (2009). Relationships between fundamental movement skills and objectively measured physical activity in preschool children. *Pediatr. Exerc. Sci.* 21, 436–449. doi: 10.1123/pes.21.4.436
- Cohen, B. J., and Korintus, M. (2017). Making connections: reflections on over three decades of Eu initiatives in early childhood education and care (ECEC). *Early Years* 37, 235–249. doi: 10.1080/09575146.2016.1181050
- Copeland, K. A., Kendeigh, C. A., Saelens, B. E., Kalkwarf, H. J., and Sherman, S. N. (2012). Physical activity in child-care centers: do teachers hold the key to the playground? *Health Educ. Res.* 27, 81–100. doi: 10.1093/her/cyr038
- Dalli, C., Miller, L., and Urban, M. (2012). *Early childhood grows up: towards a critical ecology of the profession* Springer.
- Eddy, L. H., Wood, M. L., Shire, K. A., Bingham, D. D., Bonnick, E., Creaser, A., et al. (2019). A systematic review of randomized and case-controlled trials investigating the effectiveness of school-based motor skill interventions in 3-to 12-year-old children. *Child Care Health Develop.* 45, 773–790. doi: 10.1111/cch.12712
- Einarsdottir, J., Puroila, A.-M., Johansson, E. M., Broström, S., and Emilson, A. (2015). Democracy, caring and competence: values perspectives in Ecec curricula in the Nordic countries. *Int. J. Early Years Educ.* 23, 97–114. doi: 10.1080/09669760.2014.970521
- Frank, M. L., Flynn, A., Farnell, G. S., and Barkley, J. E. (2018). The differences in physical activity levels in preschool children during free play recess and structured play recess. *J. Exerc. Sci. Fitness* 16, 37–42. doi: 10.1016/j.jesf.2018.03.001
- Gibbs, Jr. R. W. (2001). Intentions as emergent products of social interactions. In *Intentions and intentionality: Foundations of social cognition*. eds. B. F. Malle, L. J. Moses and D. A. Baldwin pp 105–22.
- Hardy, L. L., King, L., Farrell, L., Macniven, R., and Howlett, S. (2010a). Fundamental movement skills among Australian preschool children. *J. Sci. Med. Sport* 13, 503–508. doi: 10.1016/j.jsams.2009.05.010

Funding

This research was sponsored by the Gyllenstiernska Krapperup Foundation (KR 2021-0045) and The Research Platform for Collaboration for Learning, Faculty of Teacher Education at Kristianstad University.

Acknowledgments

The author would like to thank the participants for participating in the study, as well as Hanna Hermansson and Emelie Håkansson for their help with the collection of the data. A special thank you to Jan-Eric Ekberg for cooperation.

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.

- Hardy, L. L., King, L., Kelly, B., Farrell, L., and Howlett, S. (2010b). Munch and move: evaluation of a preschool healthy eating and movement skill program. *Int. J. Behav. Nutr. Phys. Act.* 7, 1–11. doi: 10.1186/1479-5868-7-80
- Kain, J., Leyton, B., Soto-Sánchez, J., and Concha, F. (2018). In preschool children, physical activity during school time can significantly increase by intensifying locomotor activities during physical education classes. *BMC. Res. Notes* 11, 1–5. doi: 10.1186/s13104-018-3536-x
- Kyttä, M. (2004). The extent of Children's independent mobility and the number of actualized affordances as criteria for child-friendly environments. *J. Environ. Psychol.* 24, 179–198. doi: 10.1016/S0272-4944(03)00073-2
- Lu, C., and Montague, B. (2016). Move to learn, learn to move: prioritizing physical activity in early childhood education programming. *Early Childhood Educ. J.* 44, 409–417. doi: 10.1007/s10643-015-0730-5
- Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., and Okely, A. D. (2010). Fundamental movement skills in children and adolescents. *Sports Med.* 40, 1019–1035. doi: 10.2165/11536850-000000000-00000
- Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., et al. (2016). Physical activity for cognitive and mental health in youth: a systematic review of mechanisms. *Pediatrics* 138, 1–13. doi: 10.1542/peds.2016-1642
- Mak, T. C., Chan, D. K., and Capio, C. M. (2021). Strategies for teachers to promote physical activity in early childhood education settings—a scoping review. *Int. J. Environ. Res. Public Health* 18:867. doi: 10.3390/ijerph18030867
- Mavilidi, M. F., Rigoutsos, S., and Venetsanou, F. (2021). Training early childhood educators to promote children's physical activity. *Early Childhood Educ. J.* 50, 785–94. doi: 10.1007/s10643-021-01191-4
- Mcbride, R. E., Xiang, P., Wittenburg, D., and Shen, J. (2002). An analysis of preservice teachers' dispositions toward critical thinking: a cross-cultural perspective. *Asia Pac. J. Teach. Educ.* 30, 131–140. doi: 10.1080/13598660220135649
- Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. Los Angeles, US: Sage.
- Morgan, P. J., and Hansen, V. (2008). Physical education in primary schools: classroom Teachers' perceptions of benefits and outcomes. *Health Educ. J.* 67, 196–207. doi: 10.1177/0017896908094637
- Parks, M., Solmon, M., and Lee, A. (2007). Understanding classroom Teachers' perceptions of integrating physical activity: a collective efficacy perspective. *J. Res. Child. Educ.* 21, 316–328. doi: 10.1080/02568540709594597
- Pate, R. R., Mciver, K., Dowda, M., Brown, W. H., and Addy, C. (2008). Directly observed physical activity levels in preschool children. *J. Sch. Health* 78, 438–444. doi: 10.1111/j.1746-1561.2008.00327.x
- Raustorp, A., Pagels, P., Boldemann, C., Cosco, N., Söderström, M., and Mårtensson, F. (2012). Accelerometer measured level of physical activity indoors and outdoors during preschool time in Sweden and the United States. *J. Phys. Act. Health* 9, 801–808. doi: 10.1123/jpah.9.6.801
- Reilly, J. J. (2010). Low levels of objectively measured physical activity in preschoolers in child care. *Med. Sci. Sports Exerc.* 42, 502–507. doi: 10.1249/MSS.0b013e3181cea100
- Richter, L. M., Cappa, C., Issa, G., Lu, C., Petrowski, N., and Naicker, S. N. (2020). Data for action on early childhood development. *Lancet* 396, 1784–1786. doi: 10.1016/S0140-6736(20)32482-X
- Sandberg, A., and Ottosson, L. (2010). Pre-school teachers', other professionals', and parental concerns on cooperation in pre-school—all around children in need of special support: the Swedish perspective. *Int. J. Incl. Educ.* 14, 741–754. doi: 10.1080/13603110802504606
- Sandseter, E. B. H. (2014). Early childhood education and care practitioners' perceptions of children's risky play; examining the influence of personality and gender. *Early Child Dev. Care* 184, 434–449. doi: 10.1080/03004430.2013.794797
- Shonkoff, J. P. (2014). Changing the narrative for early childhood investment. *JAMA Pediatr.* 168, 105–106. doi: 10.1001/jamapediatrics.2013.4212
- Skolverket. (2021). *Statistik Över Barn Och Personal I Förskola 2021*
- Soini, A., Villberg, J., Sääkslahti, A., Gubbels, J., Mehtälä, A., Kettunen, T., et al. (2014). Directly observed physical activity among 3-year-olds in Finnish childcare. *Int. J. Early Childhood* 46, 253–269. doi: 10.1007/s13158-014-0111-z
- Soini, A., Watt, A., and Sääkslahti, A. (2021). Finnish pre-service teachers' perceptions of perceived competence in early childhood physical education. *Int. J. Environ. Res. Public Health* 18:6454. doi: 10.3390/ijerph18126454
- Sollerhed, A.-C. (2006). *Young today-adult tomorrow!: studies on physical status, physical activity, attitudes, and self-perception in children and adolescents*. Lund, Sweden: Faculty Of Medicine, Lund University.
- Sollerhed, A.-C. (2022). Perceived insufficient pedagogical content knowledge in teaching movement and physical activity. Experiences from an action-oriented study among educators in early childhood education and care. *Front. Sports Active Living* 4. doi: 10.3389/fspor.2022.1050311
- Sollerhed, A.-C. (2023). Perceived insufficient pedagogical content knowledge in teaching movement and physical activity. Experiences from an action-oriented study among educators in early childhood education and care. *Front. Sports Active Living* 4:514. doi: 10.3389/fspor.2022.1050311
- Sollerhed, A.-C., Olesen, L. G., Froberg, K., Soini, A., Sääkslahti, A., Kristjándóttir, G., et al. (2021). Movement and physical activity in early childhood education and care policies of five Nordic countries. *Int. J. Environ. Res. Public Health* 18:13226. doi: 10.3390/ijerph182413226
- Tandon, P. S., Saelens, B. E., Zhou, C., and Christakis, D. A. (2018). A comparison of preschoolers' physical activity indoors versus outdoors at child care. *Int. J. Environ. Res. Public Health* 15:2463. doi: 10.3390/ijerph15112463
- Thomson, S. (2014). 'Adulterated play': an empirical discussion surrounding adults' involvement with children's play in the primary school playground. *J. Playwork Pract.* 1, 5–21. doi: 10.1332/205316214X13944535662152
- Timmons, B. W., Leblanc, A. G., Carson, V., Connor Gorber, S., Dillman, C., Janssen, I., et al. (2012). Systematic review of physical activity and health in the early years (aged 0–4 years). *Appl. Physiol. Nutr. Metab.* 37, 773–792. doi: 10.1139/h2012-070
- Tonge, K. L., Jones, R. A., and Okely, A. D. (2016). Correlates of Children's objectively measured physical activity and sedentary behavior in early childhood education and care services: a systematic review. *Prev. Med.* 89, 129–139. doi: 10.1016/j.ypmed.2016.05.019
- Tremblay, M. S., Gray, C., Babcock, S., Barnes, J., Costas Bradstreet, C., Carr, D., et al. (2015). Position statement on active outdoor play. *Int. J. Environ. Res. Public Health* 12, 6475–6505. doi: 10.3390/ijerph120606475
- Van Cauwenberghe, E., De Craemer, M., De Decker, E., De Bourdeaudhuij, I., and Cardon, G. (2013). The impact of a teacher-led structured physical activity session on preschoolers' sedentary and physical activity levels. *J. Sci. Med. Sport* 16, 422–426. doi: 10.1016/j.jsams.2012.11.883
- Vanderloo, L. M., Tucker, P., Johnson, A. M., Van Zandvoort, M. M., Burke, S. M., and Irwin, J. D. (2014). The influence of centre-based childcare on preschoolers' physical activity levels: a cross-sectional study. *Int. J. Environ. Res. Public Health* 11, 1794–1802. doi: 10.3390/ijerph110201794
- Verstraete, S. J., Cardon, G. M., De Clercq, D. L., and De Bourdeaudhuij, I. M. (2006). Increasing children's physical activity levels during recess periods in elementary schools: the effects of providing game equipment. *Eur. J. Pub. Health* 16, 415–419. doi: 10.1093/eurpub/ckl008
- Vorkapic, C. F., Alves, H., Araujo, L., Borba-Pinheiro, C. J., Coelho, R., Fonseca, E., et al. (2021). Does physical activity improve cognition and academic performance in children? A systematic review of randomized controlled trials. *Neuropsychobiology*, 454–82. doi: 10.1159/000514682
- Ward, D. S., Vaughn, A., McWilliams, C., and Hales, D. (2010). Interventions for increasing physical activity at child care. *Med. Sci. Sports Exerc.* 42, 526–534. doi: 10.1249/MSS.0b013e3181cea406
- Webster, C. A., Buchan, H., Perreault, M., Doan, R., Douth, P., and Weaver, R. G. (2015). An exploratory study of elementary classroom teachers' physical activity promotion from a social learning perspective. *J. Teach. Phys. Educ.* 34, 474–495. doi: 10.1123/jtpe.2014-0075
- Williams, H. G., Pfeiffer, K. A., O'Neill, J. R., Dowda, M., Mciver, K. L., Brown, W. H., et al. (2008). Motor skill performance and physical activity in preschool children. *Obesity* 16, 1421–1426. doi: 10.1038/oby.2008.214
- World Health Organization (2019). *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age: web annex: evidence profiles*. Geneva, CH: World Health Organization.
- Wyver, S., Tranter, P., Naughton, G., Little, H., Sandseter, E. B. H., and Bundy, A. (2010). Ten ways to restrict Children's freedom to play: the problem of surplus safety. *Contemp. Issues Early Child.* 11, 263–277. doi: 10.2304/ciec.2010.11.3.263



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo, Portugal

REVIEWED BY

Inês Peixoto Silva,
Escola Superior de Educação, Instituto
Politécnico de Viana do Castelo, Portugal
María Martínez López,
University of Castilla La Mancha, Spain

*CORRESPONDENCE

Guoxia Wang
✉ wanggx690@nenu.edu.cn

RECEIVED 08 May 2023

ACCEPTED 24 July 2023

PUBLISHED 04 August 2023

CITATION

Wu Q, Wang G and Li C (2023) Relationship
between quality of professional capacity
building for kindergarten teachers and
children's language development: the
mediating role of kindergarten language
education activities quality.
Front. Psychol. 14:1219330.
doi: 10.3389/fpsyg.2023.1219330

COPYRIGHT

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

Relationship between quality of professional capacity building for kindergarten teachers and children's language development: the mediating role of kindergarten language education activities quality

Qiong Wu¹, Guoxia Wang^{2*} and Chang Li¹

¹Faculty of Education, Northeast Normal University, Changchun, Jilin, China, ²School of Psychology, Northeast Normal University, Changchun, Jilin, China

Strengthening the professional capacity building for kindergarten teachers is crucial for improving the quality of kindergarten education and promoting children's development. In order to explore the impact of the quality of professional capacity building for kindergarten teachers and kindergarten language education activities quality on children's language development, a study was conducted. A total of 1,584 children from 90 kindergartens in 5 provinces in China were randomly selected as research participants. Utilizing an independent sample *t*-test and correlation analysis (with SPSS 26.0 as the tool), the study analyzed the disparities in the professional capacity building for kindergarten teachers of different types, kindergarten language education activities quality, and the children's language developmental levels, as well as the interrelationships among these factors. A multilevel structural equation model (using MPLUS 8.3 as the tool) was employed to analyze the mediating role that the kindergarten language education activities quality plays between the quality of professional capacity building for kindergarten teachers and children's language development levels. The results indicate that children in public kindergartens demonstrate significantly higher language development levels than those in private kindergartens. A positive correlation exists between the quality of professional capacity building for kindergarten teachers, kindergarten language education activities quality and children's language development levels. Furthermore, kindergarten language education activities quality shows partial mediating effects on the relationship between them. These findings highlight the need to enhance the quality of education in private kindergartens, reduce the disparities in language development levels between private and public kindergartens, and reinforce the quality of professional capacity building for kindergarten teachers to provide effective support for children's language development. Ultimately, the study suggests concentrating on the organization and implementation of kindergarten language education activities as a means of evaluating the quality of professional capacity building for kindergarten teachers.

KEYWORDS

children's language development level, quality of professional capacity building for kindergarten teachers, kindergarten language education activities quality, teacher training, preschool education reform

1. Introduction

Strengthening teacher professional development is a key step towards implementing the connotation construction and achieving high-quality development in preschool education in China. Since 2010, China has issued a series of policy documents that provide specific opinions and requirements on how to build the kindergarten teacher workforce. In 2012, the Ministry of Education issued the “Professional Standards for Kindergarten Teachers (Trial Implementation),” which positioned “ability as the priority” as one of the basic principles that kindergarten teachers must adhere to (Ministry of Education of the People’s Republic of China, 2012). In 2018, the Central Committee of the Communist Party of China and the State Council issued the “Opinions on Deepening the Reform of Teacher Development in the New Era,” which emphasized the need to effectively enhance the professional childcare and education skills of kindergarten teachers (The State Council of the CPC Central Committee, 2018). In 2022, the Ministry of Education issued a notice on the “Guidelines for Evaluating the Quality of Preschool Care and Education,” which included key indicators for evaluating the teacher workforce such as professional development, staffing, professional ethics, and incentive mechanisms, with the aim of strengthening the ethical work of kindergarten teachers and emphasizing their professional development (Ministry of Education of the People’s Republic of China, 2022). It is evident that China is currently emphasizing the professional development of kindergarten teachers, focusing on improving their professionalization level, and promoting the in-depth development of preschool education reform.

The developmental status of children is an important indicator for measuring the quality of professional capacity building of kindergarten teachers and the success of preschool education reform (Hong and Pang, 2006; Li et al., 2021). Against the backdrop of China’s emphasis on the quality of professional capacity building for kindergarten teachers, it is a highly concerning issue for policy makers and researchers to explore whether the quality of professional capacity building for kindergarten teachers can promote children’s development and to what extent. The main approaches to professional capacity building for kindergarten teachers include teacher training, teaching and research activities, incentives and evaluation, and autonomous development (Xiao and Hu, 2007; Wang, 2011; Xu and Yun, 2013; Sun and Zhang, 2019). Currently, there have been some empirical studies abroad exploring the relationship between teacher training and child development. A meta-analysis study on the impact of teacher professional development program quality on child development suggests that teacher training has a significant effect on both teacher professional growth and child development (Egert et al., 2018). Teacher training and support in building intimate teacher-child relationships can maximize children’s readiness for learning by promoting social/emotional, behavioral, cognitive, and physical health development (Smith and Glass, 2019). When innovative methods such as after-school consultation, online face-to-face guidance or supervision are used to train teachers, it can effectively improve the quality of the course implementation process and promote the early math skills of preschool children (Domitrovich et al., 2009; Kinzie et al., 2014). Compared with teachers who did not receive professional development programs, those who received project training made significant overall progress in teacher-student communication (Wasik et al., 2006) and were better able to plan and think of appropriate professional measures

to support children’s language development (Michel et al., 2014). Children who interacted with teachers participating in professional development programs were more efficient in language and used more complex language than those who interacted with teachers using traditional methods (Piasta et al., 2012), and their knowledge of grammar and letters was also expanded (Cabell et al., 2011). Increasing expenditure on kindergarten teacher training is an effective measure to improve the quality of the education process and has a positive impact on children’s development and learning (Hu et al., 2017). However, some studies have suggested that there is no relationship between teacher training and child development (Lin and Magnuson, 2018). This may be due to the fact that teacher training has not provided teachers with the necessary knowledge and skills to promote child development (Hyson et al., 2009). Another possible reason is that in low-income social environments, due to limited educational resources, teachers may not be able to apply the knowledge and skills learned from training in the classroom (Wang et al., 2020).

Teaching and research activities refer to practical and reflective research activities in which kindergarten teachers analyze and solve specific problems faced in education and teaching in a planned, process-oriented, and methodical manner under the organization of research leaders, with the purpose of promoting teacher professional development. Previous studies have analyzed the value of teacher research activities and found that they are of great value in promoting the comprehensive development of young children (Torquati et al., 2007), but there is a lack of evidence-based educational research in this area. Researchers mainly focus on the implementation, existing problems, and implementation strategies of policy requirements for teacher motivation and evaluation, as well as teacher autonomous development, but there are very few studies on their effectiveness. As a result, although various methods have been adopted to promote the professional development of kindergarten teachers in practice, the actual effectiveness, including their impact on child development, is unknown. Moreover, it is unclear which aspects of the existing content and pathways of teacher professional development need to be improved to effectively enhance teacher professional development.

Therefore, it is imperative to conduct in-depth research on the relationship between the quality of professional capacity building for kindergarten teachers and children’s development. Current research on the relationship between teacher training and child development is inconclusive, and it is necessary to conduct a thorough analysis of the relationship between the quality of professional capacity building for kindergarten teachers in other forms and children’s development in kindergartens. Child development refers to the comprehensive development of children in areas such as physical, cognitive, emotional, social adaptation, and language (UNICEF, 2017). Language is an important indicator for assessing children’s developmental level. In different countries and regions, language development is considered an important area or key dimension in guidelines for children’s learning and development and in assessments of child development (OECD, 2020).

Teacher professional capacity building aims at promoting the professional development for kindergarten teachers that can improve the quality of kindergarten education activities. Kindergarten language education activities refers to various forms of teaching process carried out in kindergartens with children as the subject and language as the object (Zhang, 2010). Studies have shown that kindergarten language education activities can systematically and efficiently help children acquire the core experience of language learning and development

(Wang et al., 2021). According to global quality measures, young children in high-quality educational activities tend to score higher on receptive vocabulary and better language skills (Abreu-Lima et al., 2013). Kindergarten education activities focusing on literacy can improve children's early literacy skills (Farver et al., 2009; Goodrich et al., 2017).

This study aims to explore the relationship between the quality of professional capacity building for kindergarten teachers and children's language development, and to analyze the mediating role of kindergarten language education activities quality. The findings of this study will provide scientific evidence for improving the quality of professional capacity building for kindergarten teachers and promoting children's language development.

This study proposes the following research hypotheses: Firstly, there are significant differences in language development levels among children in different types of kindergartens (public and private). Secondly, the quality of professional capacity building for kindergarten teachers significantly affects children's language development levels. Thirdly, kindergarten language education activities quality plays a mediating role between the quality of professional capacity building for kindergarten teachers and children's language development levels.

2. Research methods

2.1. Research participants

This study employed a stratified sampling method to select 100 public and private kindergartens in urban, suburban, town, and rural areas in five provinces (autonomous regions) of China: Guangxi, Jilin, Jiangsu, Shaanxi, and Zhejiang. One junior class (3–4 years old children), middle class (4–5 years old children) and senior class (5–6 years old children) were randomly selected from each kindergarten, and 6 children (half male and half female) were randomly sampled from each class, resulting in a total of 1,800 children. Based on the research questions to be explored, 10 kindergartens with missing data on quality of professional capacity building for kindergarten teachers were excluded, a total of 90 kindergartens with an effective rate of 90% were selected as the study sample. The sample size for child language development assessment was 1,584 from 90 kindergartens (The evaluation results of 36 young children were incorrect, so they were deleted), with an effective rate of 98%. Among them, 911 children were from public kindergartens, and 673 were from private kindergartens (see Table 1 for details).

2.2. Evaluation instruments

2.2.1. Evaluation tool for quality of professional capacity building for kindergarten teachers

The study employed the evaluation tool for assessing the quality of professional capacity building for kindergarten teachers, based on the “The Path Towards Excellence—Chinese Kindergarten Education Quality Rating Standards (short for PTE-CKEQRS in the following) (Chen et al., 2021). The overall and internal consistency of the PTE-CKEQRS as well as the consistency of each domain, reached a high level. Most items demonstrated good discriminant validity and construct validity. The PTE-CKEQRS consist of five quality domains: *Management Leadership*, *Environmental Support*, *Curriculum Promotion*, *Teacher Qualifications Guarantee* and *Home-Kindergarten-Community Cooperation*. Each domain contains several items, sub-items, and finely graded indicators of different levels. For example, *Teacher Qualifications Guarantee* comprises four items: teacher staffing, salary benefits, teacher ethics, and professional development. The professional development item includes four sub-items: teacher training, teaching and research activities, incentives and evaluations, and autonomous development, each with several finely graded evaluation indicators of different levels. In the quality evaluation process, evaluators first assess each fine indicator, and the evaluation result is either yes or no. Then, they score the sub-item using the Likert seven point scoring method, where 1 point represents not applicable, 3 points represent qualified, 5 points represent good, and 7 points represent excellent. The sub-item score follows the principle of step-by-step inference from low to high, assigning a corresponding level of score (between 1 and 7) based on the evaluation results and scoring rules of the fine indicators. The item score is the average score of multiple sub-items, ranging from 1 to 7 (with two decimal places). A higher score indicates higher quality of professional capacity building for kindergarten teachers. The internal consistency reliability of the tool is good, with a Cronbach's alpha coefficient of 0.849.

2.2.2. Evaluation tool for kindergarten language education activities quality

This study evaluated kindergarten language education activities quality using the language domain of the “Curriculum Promotion” quality area in the PTE-CKEQRS (Chen et al., 2021). The language domain includes four sub-items: listening, speaking, reading, and writing. The scoring of kindergarten language education activities quality is the same as that of the quality of professional capacity

TABLE 1 Description of sampled kindergartens and children.

Province	Kindergarten ownership		Number of kindergartens	Sample classes			Number of children
	Public	Private		Junior class of kindergarten	Middle class of kindergarten	Senior class of kindergarten	
Guangxi	11	9	20	115	117	118	350
Jilin	6	6	12	62	64	69	195
Jiangsu	13	7	20	120	118	117	355
Shaanxi	10	8	18	108	108	108	324
Zhejiang	12	8	20	118	119	123	360
Total	52	38	90	523	526	535	1,584

building for kindergarten teachers. The internal consistency reliability of this tool is good, with Cronbach's alpha coefficients for the language domain ranging from 0.754.

2.2.3. Evaluation tool for children's language development level

The tool used to assess children's language development level is the Peabody Picture Vocabulary Test-revised edition type A (PPVT) (Lu and Liu, 2005). The PPVT adopts a 0–1 scoring system, with a score of 1 for correct answers and 0 for incorrect answers. Studies have shown that this assessment tool has good psychometric properties (Zhang et al., 2021), with internal consistency reliability coefficients of 0.91.

2.3. Evaluation process

Prior to the formal evaluation, researchers underwent rigorous and standardized testing training, which included evaluation methods, interpretation of evaluation indicators, and on-site testing. For the evaluation of quality of professional capacity building for kindergarten teachers and kindergarten language education activity quality, two evaluators worked independently in each group. After the evaluation was completed, they discussed and made the final evaluation results based on the PTE-CKEQRS. For the assessment of children's language development level, child development evaluators conducted one-on-one assessments on sampled children in the kindergarten. The assessment guidelines and procedures were strictly followed according to each measurement scale's instructions. The assessment usually took 10–20 min depending on the child's age and response time.

2.4. Data analysis

This study used SPSS 26.0 for independent samples *t*-test and correlation analysis. As the independent variable (quality of professional capacity building for kindergarten teachers) and the mediating variable (kindergarten language education activities quality) in this study belong to the kindergarten level (L2), while the dependent variable (children's language development level) belongs to the individual level (L1), a multilevel data analysis method was adopted. The analysis of mediating effects used the coefficient product method, which examines whether $a \times b$ is significant. If $a \times b$ is significant (significant level of 0.05) and its confidence interval does not include 0, it indicates a significant multilevel mediating effect (Zhao et al., 2010). Mplus 8.3 was used as the analysis software.

3. Results and analysis

3.1. Means and standard deviations of all study variables

Kindergartens can be classified into public kindergartens and private kindergartens based on their ownership. An independent samples *t*-test was conducted to examine the differences in children's language development level, quality of professional capacity building for kindergarten teachers, kindergarten language education activities quality between different types of kindergartens (see Table 2 for

results). The results showed significant differences in children's language development level, quality of professional capacity building for kindergarten teachers, and kindergarten language education activities quality between different types of kindergartens. Children attending public kindergartens had significantly higher language development levels than those attending private kindergartens ($t = 3.56, p < .01$), the quality of professional capacity building for kindergarten teachers and kindergarten language education activity quality in public kindergarten had significantly higher than those in private kindergartens ($t = 17.79, p < .001$; $t = 9.714, p < .001$).

3.2. Correlation analysis of all study variables

The raw scores of children's language development levels in each age group were transformed into standard scores. Then, Pearson correlation was used to examine the correlation between quality of professional capacity building for kindergarten teachers in the four dimensions and children's language development level (results shown in Table 3). The statistical results showed that the quality of teacher training, teaching and research activities, incentives and evaluation, autonomous development were all significantly positively correlated with children's language development level and kindergarten language education activities quality.

3.3. Mediating effect test

As the quality of professional capacity building for kindergarten teachers and kindergarten language education activities quality belong to the level of kindergarten hierarchy (L2), while children's language development level belongs to the level of individual development (L1), a multilevel structural equation modeling approach is adopted for the subsequent mediation analysis. Firstly, the null model of children's language development level is tested to examine whether the data is suitable for multilevel analysis. The result shows that the within-group correlation coefficient (ICC) of children's language development level is 0.194, which is greater than 0.06,

TABLE 2 Means and standard deviations of all study variables.

Evaluation Indicator	Public		Private		<i>t</i>
	Mean	SD	Mean	SD	
Children's language development level	52.83	23.54	48.50	23.96	3.56**
Quality of professional capacity building for kindergarten teachers	4.58	1.58	3.57	1.10	17.79***
Kindergarten language education activities quality	3.61	0.75	3.27	0.56	9.714***

** $p < 0.01$; *** $p < 0.001$.

TABLE 3 Correlation analysis of all study variables.

Variables	Correlation							Mean $\pm$ SD
	1	2	3	4	5	6	7	
1. Quality of professional capacity building for kindergarten teachers	1							4.147 $\pm$ 1.223
2. Teacher training	0.786**	1						3.022 $\pm$ 2.034
3. Teaching and research activities	0.766**	0.414**	1					5.39 $\pm$ 1.547
4. Incentives and evaluation	0.673**	0.358**	0.384**	1				4.025 $\pm$ 1.212
5. Autonomous development	0.814**	0.466**	0.562**	0.479**	1			4.152 $\pm$ 1.587
6. Children's language development level	0.222**	0.157**	0.171**	0.153**	0.201**	1		0.025 $\pm$ 0.983
7. Kindergarten language education activities quality	0.396**	0.299**	0.310**	0.297**	0.309**	0.232**	1	3.465 $\pm$ 0.696

** $p < 0.01$.

indicating the necessity of conducting multilevel analysis on the explanatory mechanism of the antecedents of children's language development level. A 2-2-1 model is constructed, which includes quality of professional capacity building for kindergarten teachers (L2) – kindergarten language education activities quality (L2) – children's language development level (L1).

The following are the steps for analyzing the mediating effect of kindergarten language education activities quality between the quality of professional capacity building for kindergarten teachers and children's language development level, using a multilevel structural equation modeling approach:

Step 1: Multilevel mediation model (see Figure 1) is tested, with quality of professional capacity building for kindergarten teachers as the independent variable, kindergarten language education activities quality as the mediator, and children's language development level as the dependent variable. The model fit indices are as follows: CFI = 1.00, TLI = 1.00, RMSEA < 0.001, SRMRB = 0.000, SRMRw < 0.001, indicating good model fit.

Step 2: The results of the multilevel structural equation modeling analysis indicate that the path coefficient from quality of professional capacity building for kindergarten teachers to the kindergarten language education activities quality (a) is significant ($\gamma = 0.236$, $p < 0.001$). The path coefficient from the kindergarten language education activities quality to children's language development level (b) is significant ($\gamma = 0.244$, $p < 0.01$). The mediating effect of $a \times b$ is significant ($\gamma = 0.058$, $p < 0.001$), and the direct effect from the quality of professional capacity building for kindergarten teachers to children's language development level (c) is also significant ($\gamma = 0.128$, $p < 0.01$) (see Table 4). This indicates that kindergarten of language education activities quality partially mediates the relationship between quality of professional capacity building for kindergarten teachers and children's language development level in a 2-2-1 model.

4. Discussion

4.1. Significant differences in language development levels exist among children in kindergartens of different characteristics

This study reveals that significant differences exist in language development levels among children attending kindergartens of different characteristics in China. Children in public kindergartens exhibit considerably higher language development levels than those in private kindergartens. This is consistent with previous research findings. For instance, studies have shown that children in public kindergartens score significantly higher in areas such as free play, language, and social domains compared to those in private kindergartens (Liu et al., 2008). Other research indicates that children in private kindergartens score lower in language, mathematics, science, and social domains than their public kindergarten counterparts (Shi and Ye, 2016).

Various factors contribute to these differences in children's language development levels, with disparities in educational quality being a crucial factor. Studies have shown that high-quality kindergarten education can bring positive impacts on children's future development (Nores et al., 2005). The higher the quality of kindergarten education, the greater the added value for children's development (Vandell et al., 2010); Therefore, China should strive to improve the educational quality of private kindergartens. On one hand, the proportion of funding for inclusive private kindergartens should be clarified, providing adequate financial support for the supply of high-quality inclusive educational resources. On the other hand, the quality supervision and dynamic regulation of private kindergartens should be strengthened, with periodic diagnostic assessments and rectifications, linking quality assessment results with

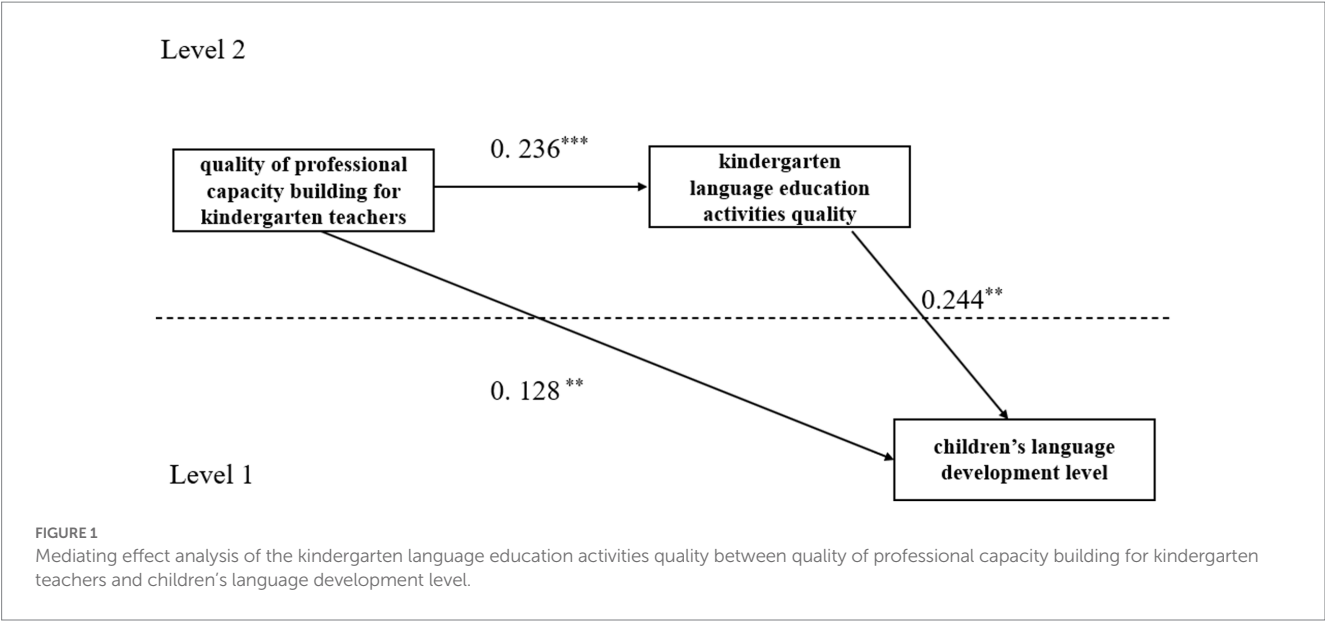


TABLE 4 Hypotheses model path coefficient.

Model pathway	Point estimate	95% CI	S.E.	Est./S.E.	p
Quality of professional capacity building for kindergarten teachers→kindergarten language education activities quality	0.236	[0.139, 0.334]	0.059	3.979	<0.001
Kindergarten language education activities quality→children's language development level	0.244	[0.121, 0.366]	0.074	3.282	<0.01
Quality of professional capacity building for kindergarten teachers→children's language development level	0.128	[0.061,0.195]	0.041	3.129	<0.01
Quality of professional capacity building for kindergarten teachers→kindergarten language education activities quality→kindergarten language education activities quality	0.058	[0.031, 0.084]	0.016	3.546	<0.001

funding subsidies to motivate private kindergartens to continuously enhance their educational quality.

4.2. The quality of professional capacity building for kindergarten teachers is significantly positively correlated with children's language development levels

This study finds that quality of professional capacity building for kindergarten teachers is significantly positively correlated with children's language development levels. These findings are consistent with previous

research perspectives. Some studies have pointed out that research support significantly affects the added value of children's development in language and early reading domains, especially for children with lower initial levels in language and early reading (Li et al., 2021). Simultaneously, the results of this study also validate the theoretical perspectives of some researchers. Some researchers believe that research activities can enhance the quality of kindergarten educational activities and improve children's development outcomes (Huang, 2019); kindergarten-based research is an essential approach to promoting kindergarten teachers' professional development, serving as an essential platform for their professional growth and a mediator for translating educational concepts into educational practices, playing a crucial role in improving kindergarten

education quality and promoting children's development (Zou and Hua, 2015). Autonomous learning of kindergarten teachers not only enhances their professional literacy and kindergarten education quality but also fosters the comprehensive and healthy development of children (Hai and Liu, 2018). Therefore, kindergartens should solidly carry out teacher training, effectively organize teaching and research activities, implement teacher incentives and evaluation, and promote teacher autonomous development. Through the effective implementation of these activities, the continuous improvement of teachers' professional capacity can be promoted, thereby providing robust support for children's language development.

4.3. The kindergarten language education activities quality plays a partial mediating role between the quality of professional capacity building for kindergarten teachers and children's language development levels

This study demonstrates that the kindergarten language education activities quality plays a partial mediating role between quality of professional capacity building for kindergarten teachers and children's language development levels. In other words, the quality of professional capacity building for kindergarten teachers not only directly affects children's language development levels but also indirectly influences them through the impact on the kindergarten language education activities quality. This suggests that teachers can apply the teaching experiences or strategies gained from professional capacity building activities to the implementation of kindergarten language education activities, thereby promoting children's language development. This validates existing research perspectives, as studies have indicated that teachers' sensitivity and effective responses to children's behavior during the organization and implementation of educational activities can effectively predict children's language and social interaction abilities (Burchinal et al., 2008). Based on this, kindergartens can focus on the implementation of language education activities, evaluate teachers' performance in designing and implementing language education activities, examine the quality of professional capacity building for kindergarten teachers, and encourage teachers to translate their acquired knowledge and skills from understanding and acceptance to practical implementation, ultimately promoting children's language development.

5. Conclusion

Firstly, the significant differences exist in the language development levels of children attending kindergartens with different types. Children in public kindergartens demonstrate notably higher language development levels compared to those in private kindergartens.

Secondly, the quality of professional capacity building for kindergarten teachers has a significant positive predictive effect on children's language development levels. Improving the quality of professional capacity building for kindergarten teachers can enhance children's language development levels.

Thirdly, kindergarten language education activities quality plays a partial mediating role between the quality of professional capacity building for kindergarten teachers and children's language development levels. That is, the quality of professional capacity building for kindergarten teachers affects children's language development levels through two pathways: directly influencing children's language development levels and indirectly affecting children's language development levels by influencing kindergarten language education activities quality.

6. Limitations of the study and future research directions

Several limitations of the present study warrant note. First, our evaluation tool for quality of professional capacity building for kindergarten teachers may not have been sensitive enough to reflect the overall performance of kindergarten teachers' professional capacity building. Nevertheless, we believe that the evaluation tool supports the validity of our findings. Because the tool covers many key aspects of professional capacity building for kindergarten teachers, the score was obtained on the basis of on-site observation and data review.

Secondly, the study only focuses on the relationship between kindergarten language education activities quality, the quality of professional capacity building for kindergarten teachers and children's language development. There may be other situations that lead to children's language development, such as other areas education activities in kindergarten, more and higher quality dialogue between teachers and children, which need to be further studied.

Thirdly, this study only made a brief analysis of language development levels exist among children in public kindergartens and private kindergartens. Future studies should deeply explore the factors affecting children's language development levels in private kindergartens and take effective measures to support these children.

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 Institutional Review Board of Northeast Normal University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

QW and GW designed the study and devised the project. QW and CL took the lead in writing the manuscript. GW verified the analytical methods and contributed to manuscript revision. All

authors contributed to the article and approved the submitted version.

Funding

This study was supported by Humanities and Social Science of Ministry of Education Planning Fund: “Research on Strategies for Improving the Quality of Teacher Child Interaction in Kindergartens” (Reference No.: 19YJA880065). This study was also partially funded by China National Society of Early Childhood Education key topics in the 13th Five-Year Plan: “Research on the Evaluation Standards of China’s High Quality Kindergartens.”

References

- Abreu-Lima, I. M. P., Leal, T. B., Cadima, J., and Gamelas, M. (2013). Predicting child outcomes from preschool quality in Portugal. *Eur. J. Psychol. Educ.* 28, 399–420. doi: 10.1007/s10212-012-0120-y
- Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. *Appl. Dev. Sci.* 12, 140–153. doi: 10.1080/10888690802199418
- Cabell, S. Q., Justice, L. M., Piasta, S. B., Cumenton, S. M., and Petscher, Y. (2011). The impact of teacher responsiveness education on preschoolers’ language and literacy skills. *Am. J. Speech Lang. Pathol.* 20, 315–330. doi: 10.1044/1058-0360(2011/10-0104)
- Chen, D. Z., Li, K. J., and Zhou, J. (2021). Psychometric property analysis on the path towards excellence—Chinese kindergarten education quality rating standards (PTE-CKEQRS) based on the data about 100 kindergartens and 1670 preschoolers in China. *Stud. Early Child. Educ.* 313, 3–16. doi: 10.13861/j.cnki.sece.2021.01.001
- Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., and Jones, D. (2009). Fostering high-quality teaching with an enriched curriculum and professional development support: the head start REDI program. *Am. Educ. Res. J.* 46, 567–597. doi: 10.3102/0002831208328089
- Egert, F., Fukkink, R. G., and Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: a meta-analysis. *Rev. Educ. Res.* 88, 401–433. doi: 10.3102/0034654317751918
- Farver, J. A., Lonigan, C. J., and Eppe, S. (2009). Effective early literacy skill development for young Spanish-speaking English language learners: an experimental study of two methods. *Child. Dev.* 80, 703–719. doi: 10.1111/j.1467-8624.2009.01292.x
- Goodrich, J. M., Lonigan, C. J., and Farver, J. A. (2017). Impacts of a literacy-focused preschool curriculum on the early literacy skills of language-minority children. *Early Child. Res. Q.* 40, 13–24. doi: 10.1016/j.ecresq.2017.02.001
- Hai, Y., and Liu, T. E. (2018). The analysis of developmental level and characteristics of preschool teachers’ self-regulated learning. *J. Educ. Stud.* 14, 89–102. doi: 10.14082/j.cnki.1673-1298.2018.03.010
- Hong, X. M., and Pang, L. J. (2006). Teachers’ perceived self-efficacy on children’s development. *Stud. Early Child. Educ.* 06, 44–47. doi: 10.3969/j.issn.1007-8169.2006.06.014
- Hu, B. Y., Zhou, Y., Chen, L., Fan, X., and Winsler, A. (2017). Preschool expenditures and Chinese children’s academic performance: the mediating effect of teacher-child interaction quality. *Early Child. Res. Q.* 41, 37–49. doi: 10.1016/j.ecresq.2017.05.002
- Huang, H. (2019). Current situations and construction paths of teachers’ community culture in kindergarten-based teaching and research. *Stud. Early Child. Educ.* 11, 35–50. doi: 10.13861/j.cnki.sece.2019.11.005
- Hyson, M., Tomlinson, H. B., and Morris, C. A. (2009). Quality improvement in early childhood teacher education: faculty perspectives and recommendations for the future. *Early Child. Res. Q.* 11, 1–3. doi: 10.1016/S0885-2006(96)90025-3
- Kinzie, M. B., Whittaker, J. V., Williford, A. P., DeCoster, J., McGuire, P., and Lee, Y. (2014). MyTeachingPartner-math/science pre-kindergarten curricula and teacher supports: associations with children’s mathematics and science learning. *Early Child. Res. Q.* 29, 586–599. doi: 10.1016/j.ecresq.2014.06.007
- Li, L., Li, Z. J., Fan, J. Q., Ren, L. X., and Liu, H. (2021). The impact of early education quality on children value-added development. *Stud. Early Child. Educ.* 04, 31–47. doi: 10.13861/j.cnki.sece.2021.04.004
- Lin, Y., and Magnuson, K. A. (2018). Classroom quality and children’s academic skills in child care centers: understanding the role of teacher qualifications. *Early Child. Res. Q.* 42, 215–227. doi: 10.1016/j.ecresq.2017.10.003
- Liu, Y., Li, Z. Y., Pan, Y. J., and Zhang, T. (2008). Comparative study on environmental quality of kindergartens of different financial input and management system. *Stud. Early Child. Educ.* 08, 7–11.
- Lu, L., and Liu, H. X. (2005). *Peabody picture vocabulary test-revised edition type a[M]*. Taipei: Psychology press.
- Michel, M., Ofner, D., and Thoma, D. (2014). Early childhood educators’ competences for supporting children’s academic language skills in Germany. *Lang. Aware.* 23, 138–156. doi: 10.1080/09658416.2013.863896
- Ministry of Education of the People’s Republic of China. (2012). Professional standards for kindergarten teachers (trial implementation). Available at: http://www.moe.gov.cn/srcsite/A10/s6991/201209/t20120913_145603.html.
- Ministry of Education of the People’s Republic of China. (2022). Guidelines for evaluating the quality of preschool care and education. Available at: http://www.moe.gov.cn/srcsite/A06/s3327/202202/t20220214_599198.html.
- Nores, M., Belfield, C. R., Barnett, W. S., and Schweinhart, L. (2005). Updating the economic impacts of the high/scope Perry preschool program. *Educ. Eval. Policy Anal.* 27, 245–261. doi: 10.3102/01623737027003245
- OECD. (2020). Early learning and child well-being: a study of five-year-olds in England, Estonia, and the United States. Available at: <http://www.oecd.org/education/school/early-learning-and-child-well-being-study/early-learning-and-child-well-being-3990407f-en.html>.
- Piasta, S. B., Justice, L. M., Cabell, S. Q., Wiggins, A. K., Turnbull, K. P., and Cumenton, S. M. (2012). Impact of professional development on preschool teachers’ conversational responsivity and children’s linguistic productivity and complexity. *Early Child. Res. Q.* 27, 387–400. doi: 10.1016/j.ecresq.2012.01.001
- Shi, J., and Ye, P. Z. (2016). Research on the relationship between classroom environment quality of kindergarten and children’s school readiness. *Stud. Early Child. Educ.* 08, 41–50. doi: 10.13861/j.cnki.sece.2016.08.005
- Smith, N., and Glass, W. (2019). Ready or not? Teachers’ perceptions of young children’s school readiness. *J. Early Child. Res.* 17, 329–346. doi: 10.1177/1476718X19875760
- Sun, L. H., and Zhang, L. Y. (2019). The development of kindergarten teacher’s professional capacity—based on the empirical investigation of kindergarten teachers’ national training. *J. Chin. Soc. Educ.* 02, 18–21.
- The State Council of the CPC Central Committee. (2018). Opinions on deepening the reform of teacher development in the new era. Available at: http://www.moe.gov.cn/jyb_xwfb/moe_1946/fj_2018/201801/t20180131_326148.html.
- Torquati, J. C., Raikes, H., and Huddleston-Casas, C. A. (2007). Teacher education, motivation, compensation, workplace support, and links to quality of center-based child care and teachers’ intention to stay in the early childhood profession. *Early Child. Res. Q.* 22, 261–275. doi: 10.1016/j.ecresq.2007.03.004
- UNICEF. (2017). Integrated approaches to early childhood development: 0–3 years. Available at: <https://www.unicef.cn/en/reports/integrated-approaches-early-childhood-development-0-3-years>.
- Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., and Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child. Dev.* 81, 737–756. doi: 10.1111/j.1467-8624.2010.01431.x
- Wang, M. S. (2011). From being developed to teacher independent development: the challenge and countermeasures of Chinese teacher professional development. *Teach. Educ. Res.* 04, 1–6. doi: 10.13445/j.cnki.t.e.r.2011.04.002

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.

Publishers 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.

- Wang, L., Dang, R., Bai, Y., Zhang, S., Liu, B., and Zheng, L. (2020). Teacher qualifications and development outcomes of preschool children in rural China. *Early Child. Res. Q.* 53, 355–369. doi: 10.1016/j.ecresq.2020.05.015
- Wang, D., Wu, L., Lu, H., and Li, K. J. (2021). The relationship between kindergarten curriculum quality and early childhood language development: based on linear and nonlinear regression model analysis. *Early Child. Educ.* 12, 9–13.
- Wasik, B. A., Bond, M. A., and Hindman, A. (2006). The effects of a language and literacy intervention on head start children and teachers. *J. Educ. Psychol.* 98, 63–74. doi: 10.1037/0022-0663.98.1.63
- Xiao, C., and Hu, L. L. (2007). On school-based teaching research and teacher professional growth. *Teach. Educ. Res.* 01, 17–21. doi: 10.3969/j.issn.1672-5905.2007.01.004
- Xu, Q., and Yun, J. (2013). Incentive mechanism and teacher's professional development--on the implementation of incentive mechanism for primary and secondary school teachers in Jiangsu province. *J. Jiangsu Inst. Educ.* 06, 42–46.
- Zhang, M. H. (2010). *Early childhood language education and activity design[M]*. Beijing: higher education Press.
- Zhang, D., Zhao, W., Li, L., and Gao, G. L. (2021). Measurement of word awareness and its predictability to early reading development. *Stud. Psychol. Behav.* 19, 771–778.
- Zhao, X., Lynch, J. G., and Chen, Q. (2010). Reconsidering baron and Kenny: myths and truths about mediation analysis. *J. Consum. Res.* 37, 197–206. doi: 10.1086/651257
- Zou, H. R., and Hua, A. H. (2015). A measurement analysis of the literature research on kindergarten teachers since the 21st century. *J. Sch. Stud.* 12, 104–112. doi: 10.3969/j.issn.1005-2232.2015.02.013



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Muhammad Younas,
Soochow University, China
Xiaoming Tian,
Xi'an University of Posts and
Telecommunications, China
Keri L. Heitner,
Saybrook University, United States

*CORRESPONDENCE

Lawrence Jun Zhang
✉ lj.zhang@auckland.ac.nz

RECEIVED 27 February 2023

ACCEPTED 24 July 2023

PUBLISHED 15 August 2023

CITATION

Deng L, Zhang LJ and Mohamed N (2023)
Exploring native and non-native English
speaker teachers' perceptions of English
teacher qualities and their students' responses.
Front. Psychol. 14:1175379.
doi: 10.3389/fpsyg.2023.1175379

COPYRIGHT

© 2023 Deng, Zhang and Mohamed. 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 native and non-native English speaker teachers' perceptions of English teacher qualities and their students' responses

Liwei Deng¹, Lawrence Jun Zhang²* and Naashia Mohamed³

¹Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

Due to globalization, English has gradually become a *lingua franca*, leading to a rising demand for proficient English teachers all over the globe. In China, more EFL teachers are being recruited, particularly at the tertiary level, with a greater preference for so-called "native English speaking teachers (NESTs)" over "non-native English-speaking teachers (NNESTs)" due to the impacts of native-speakerism. Research has shown NESTs, NNESTs, and students are often misaligned in terms of beliefs about language learning and teaching which affect teaching effectiveness as well as student achievement. Recognizing this issue, this study investigated NESTs', NNESTs', and Chinese English-major students' perceptions of characteristics of effective EFL teachers at four mid-tier universities across China. Findings from semi-structured interviews with 16 students suggest that NNESTs and Chinese English-major students had similar views on language learning and teaching. Both groups valued prerequisite qualities such as having expert knowledge, language skills, teaching skills, and professionalism. NESTs, however, valued qualities such as caring, patience, flexibility, engagement, and awareness of students' learning needs. These differences are likely the result of these two groups of teachers' linguistic, cultural, and educational background differences. The highly uniform views of the two groups of teachers suggest that they tended to emphasize certain qualities while disregarding others. These findings suggest the need to raise teachers' and students' awareness of the benefits of different types of teacher qualities so that curriculum design and lesson planning can be implemented for better instructional alignment to ultimately improve teaching effectiveness.

KEYWORDS

EFL teachers, teacher quality, misalignment, NESTs, NNESTs, perceptions

Introduction

English has gradually become a *lingua franca* and is used among speakers of different first languages worldwide (Fang, 2016; Llurda and Calvet-Terré, 2022). It plays vital roles in various fields such as scientific publications, international trade, and mass international telecommunications (Yang and Zhang, 2015; Rao, 2019). As a result, an increasing number of countries are making English proficiency a national policy goal (Taguchi, 2014; Hennebry-Leung and Gao, 2022; Shan and Xu, 2022). This potentially has contributed to the drastic increase in the number of English learners worldwide, with an estimation of over 1.5 billion

(Szymigiera, 2021). For instance, China has made English a compulsory subject from Primary Three (age 9 or 10 years) onwards, and it is estimated that there are more than 400 million English learners in China (Galloway, 2017).

This expansion of the global English education market has created a growing demand for effective English as a second language (ESL) and English as a foreign language (EFL) teachers. In the field of English language teaching (ELT), the majority of teachers are multilinguals who come from language backgrounds other than English. These so-called “non-native English speaking teachers” (NNESTs) constitute about 80% of the teaching force globally (Moussu, 2018). Nevertheless, in many ESL/EFL contexts, there is a tremendous demand for English monolingual teachers—often referred to as “native English-speaking teachers” (NESTs), based on the unfounded myth that they are more proficient and skilled, making them superior English language teachers (Llurda and Calvet-Terré, 2022). Holliday (2018) used the term “native-speakerism” to describe this pervasive ideology and pointed out that it has been mainstream in the ELT field for decades. Historically, native-speakerism has been extensively accepted by the public, particularly by foreign language students and their teachers as well as education program administrators. It shapes a discriminatory and dichotomic perspective on the linguistic and cultural differences between NESTs and NNESTs. Even though this ideology is being contested, it remains a persistent issue in many ELT contexts, including China, legitimizing the discriminatory practices against NNESTs (Moussu, 2018; Zhang, 2021; Zhang and Zhang, 2021; Liu, 2021a).

Chinese universities have recently recruited more English teachers (especially NESTs) because of the pursuit of internationalization of higher education, the increasing demand for English as an important skill, and the expansion of English majors. Consequently, English majors became the largest undergraduate program in Chinese universities by 2013 (Dai and Wang, 2014; He, 2020; Zhang and Zhang, 2021). However, research has shown that NESTs’ views of effective teaching are often viewed as misaligned with local Chinese EFL teachers and students. For example, NESTs’ teaching styles conflicted with Chinese EFL students’ learning styles, and their preferences for writing feedback differed from local Chinese teachers (Rao, 2010; Cheng and Zhang, 2022). These findings indicate that NESTs in China have introduced teaching methods that might have succeeded in other settings but have not adapted their teaching to the students’ cultural and linguistic contexts.

Even though previous studies have investigated the different perceived positive and negative characteristics of NESTs and NNESTs (Zhang, 2004a,b; Javid, 2016; Cruz and Auman, 2018; Nomura and Mochizuki, 2018; Kiczowski, 2019; Wang and Fang, 2020; Zhang and Zhang, 2021; Qiu and Fang, 2022), few studies have investigated their perceptions of effective teacher qualities. In addition, most studies in China have concentrated on the response of either students or English teachers (Rao, 2010; Ma, 2016; Huang, 2017; Wang and Fang, 2020; Cheng and Zhang, 2022). The student participants often consist of non-English majors instead of English majors. Thus, this study intends to fill these gaps by investigating the perceptions of NESTs, NNESTs, and Chinese English majors of effective English teaching in the Chinese context. It aims to understand the differences between NESTs, NNESTs, and Chinese English majors’ perspectives on effective teachers’ characteristics. It provides implications for teachers to modify their teaching to

improve instructional alignment and ultimately improve teaching effectiveness.

Literature review

Definition of NESTs and NNESTs

Traditionally, a native speaker of a language is defined as someone who has been speaking the language as their first language since birth and raised in that language (Chomsky, 1965). The native speaker was assumed to possess a perfect mastery of the language. In comparison, a non-native speaker is described as someone who learns the target language as a second or foreign language. This classification of native and non-native speakers has received heavy criticism in recent decades because of its underlying raciolinguistic ideologies (Cheng et al., 2021). For instance, Braine (2010, p. 9) observed that “the term ‘native speaker’ undoubtedly has positive connotations: it denotes a birthright, fluency, cultural affinity, and sociolinguistic competence,” and “the term ‘non-native speaker’ carries the burden of the minority, of marginalization and stigmatization.” However, some researchers have viewed NNESTs labeling as a double-edged blade (Selvi, 2014). Even though using the NNEST label could have certain disadvantages such as debasing, othering NNESTs, and referring to a false standard (i.e., native-speakerism), it is also seen by some as having some benefits. This includes valuing/admitting the periphery, making it easy to organize against discrimination, using the identification already existing in the research field, and benefitting the profession (valuing education and expertise; Selvi, 2014). Therefore, the NNESTs/NESTs classification is used to emphasize the issues in China (see also Zhang and Zhan, 2014; Zhang and Zhang, 2021; Fang et al., 2022; Sun et al., 2022). Most of the Chinese EFL institutions have adopted Kachru et al.’s (2006) classification and viewed English speakers who were born and raised in “inner-circle countries” such as the UK, the USA, Anglophone Canada, Australia, and New Zealand, and specific “outer-circle countries,” or former British colonies, such as South Africa and Singapore as native English speakers. Expat English teachers who belong to this group are regarded as NESTs. Although many expats who are regarded as “NNESTs” are eligible EFL teachers, Chinese institutions are prone to hire NESTs and local Chinese English speaker teachers who are born and raised in China. While we recognize the raciolinguistic nature of this dichotomy, we use these terms in this paper as the study we present compares NESTs’ and NNESTs’ perceptions of English teacher qualities and their students’ responses. Such a comparison is significant for colleagues in the English teaching profession and administrators when working with teachers of different backgrounds in Chinese universities.

Misalignments between NESTs, NNESTs, and students on effective teaching

Cohen (1987) proposed the term “instructional alignment” and postulated that it is required across three systems (instructional processes, intended outcomes, and assessment) for effective learning. Specifically, teachers need to align how they teach, what they want students to learn, and what is assessed to teach effectively. In this study, the term “alignment” indicates teachers’ instructional processes,

intended outcomes, and assessments in accordance with students' demands and interests. In contrast, misalignments between teachers and students can cause problems such as students feeling frustrated that teachers fail to meet their needs and affect students' learning outcomes and their evaluations of teachers (Rao, 2010; Yang, 2018; Cheng and Zhang, 2022).

Today, most EFL teachers in China are still indigenous Chinese EFL teachers (Pawan, 2017; Zhang, 2021; Zhang and Cheng, 2021). Nevertheless, with the growth of the number of NESTs in China, the constitution of the faculty has become more varied. This transformation has led to new challenges and increased the existing ones for effective ELT in China. Despite that NESTs are deemed as more effective teachers (Wang and Fang, 2020), research has indicated NESTs' perceptions of effective teaching are often perceived as misaligned with local Chinese EFL teachers and students. For example, Rao (2010) interviewed 20 Chinese university students and found that NESTs' teaching styles were misaligned with Chinese EFL students' learning styles. The students were unaccustomed to NESTs' hands-on style of performing role-plays and games which made it difficult for students to focus on the curriculum requirements. They were also dissatisfied with NESTs' "open" style, as they would provide several correct answers for a given question. Furthermore, the students were unused to NESTs' "intuitive-random" style as their manners differed from the traditional Chinese figure that teachers should behave like authorities. Moreover, the students criticized NESTs for favoring a "global" teaching style and used holistic strategies such as guessing and inferencing for searching the main idea of reading and listening materials and for concentrating less on the analysis of linguistic details. The fact is that the students preferred a more analytical style for studying complex linguistic styles, structures, and rhetoric. Lastly, NESTs were viewed to be insensitive to students' linguistic difficulties and less capable of offering students with local examples. These findings suggest that NESTs in China have imported teaching approaches that may have worked in other contexts, but have not adapted their teaching in response to the students' language and cultural backgrounds.

Cheng and Zhang (2022) observed that NESTs and NNESTs in the Chinese EFL university setting held different beliefs in various aspects of writing feedback scope (i.e., the extent of the feedback teachers should provide their students). For instance, most NESTs favored focused feedback and indicated that teachers should address only a few types of errors and overlook the others; NNESTs, however, preferred comprehensive feedback and argued that EFL learners usually could not recognize global issues autonomously and lacked sufficient English writing rhetorical understanding.

The misalignment also occurs between NNESTs and students in their attitudes and beliefs toward effective teaching (Yang, 2018; Saleem et al., 2022). Yang (2018) also found divergences between NNESTs' and Chinese EFL students' beliefs regarding instruction. For example, one of the teacher participants preferred a bottom-up approach to reading instruction, believing that sentence structure and vocabulary were fundamental, while five out of the six student participants leaned toward the top-down approach since it provided context and direction for reading. The teacher participant also favored a more teacher-centered model while some students required more autonomy in learning. Moreover, misalignments can arise between teachers' beliefs and practices (Zhang and Said, 2014; Bai and Yuan, 2019; Zhao and Zhang, 2022; Zhang and Sun, 2022).

The misalignments between teachers' and students' perceptions concerning effective teaching often harm students' learning results and their assessments of teachers. For example, teachers could meet difficulties understanding students' overall needs (Lo, 2015; Meissel et al., 2017; Ranjit, 2022), which makes students underperform, feel dissatisfied with their abilities, or be disappointed with certain parts of the class (Rao, 2010; Roothoof and Breeze, 2016).

Effective teachers' characteristics

There are disagreements in the beliefs of students and teachers about the characteristics of effective teachers, as it is a complex concept that has yet to reach consensus in both general education and more specialized settings such as English education (Farrell, 2015; Stronge, 2018). However, research has often suggested that certain qualities such as caring are essential for successful English language teaching (Huang, 2017; Tatipang et al., 2022). Characteristics of effective teachers are commonly categorized into general and subject-specific characteristics. The nature of education and the commonalities across various disciplines determine that some of these effective characteristics can be broadly applicable and play an important role. These qualities are regarded as prerequisites or key characteristics. For example, Stronge (2018, p. 18) outlined some important characteristics from research literature: verbal ability, knowledge of teaching and learning, certification status, content knowledge, and teaching experience. In addition to characteristics related to teaching and management procedures, affective aspects can have a positive psychological impact on students and are viewed by some researchers as equally important in terms of student success (Stronge, 2018; Zhang et al., 2019).

Affective characteristics such as fairness, sensitivity to students' needs, and caring are often associated with effective English teachers (Huang, 2017). For example, Tatipang et al. (2022) investigated the qualities of effective English teachers recognized by EFL Students in Indonesia. A total of 120 students from grade 11 in science, social, and Language class participated in the study. The most desirable characteristics include the use of easy English language to help students understand, assess what students have learned rationally, control themselves and not get angry, and know and master English vocabulary well. Research has often considered the perspectives of students and teachers as both are valuable sources of information that can provide us with a better understanding of the issue. Students have the advantage of being exposed to teachers with different characteristics and teaching practices and are consequently able to evaluate which characteristics are effective. Teachers, on the other hand, have the benefit of being former students and have gone through training, undertaken teaching activities, and learned from other teachers (Lavin et al., 2012). This has given them a comprehensive understanding of what makes a good teacher. A few studies compared the perspectives of NESTs and NNESTs on effective teaching in China (Rao, 2010; Ma, 2016; Huang, 2017; Wang and Fang, 2020), and mostly covered the views of NNESTs and non-English majors, yet the opinions of NESTs and Chinese English majors have not been sufficiently explored. Research has clearly shown NESTs have a different understanding from local NNESTs and students, and English majors also differ from non-English majors in terms of their understanding of English teachers (Rao, 2010; Cheng and Zhang,

2022). Therefore, it is necessary to compare and analyze the opinions of the participants in the study (NESTs, NNESTs, and English majors) on teaching effectiveness to offer us valuable information and a more comprehensive view of the issues. Specifically, this study addressed the following research questions:

1. What are the native and non-native English speaker teachers' and their Chinese students' perceptions of effective English teacher qualities?
2. How do their perceptions differ?

Conceptual framework

Based on the literature review and to address the research questions stated above, this study adopted a conceptual framework that consisted of three main concepts: native-speakerism (Holliday, 2018), instructional (mis)alignment (Cohen, 1987), and effective teachers' characteristics (Stronge, 2018). These concepts helped to explain and explore the perceptions of effective teaching among students, NESTs, and NNESTs in EFL contexts. They would also be useful frames of reference for our interpretation of the potential misalignment between perceptions and actual pedagogical practices, especially when we intended to examine the implications of these perceptions for improving language teaching. Ultimately, these findings are expected to help us deepen our understanding of what it means when teaching effectiveness becomes a key area of focus.

Methods

Qualitative research design

We adopted a generic qualitative design, which gave us more flexibility, creativity, and responsiveness to deal with the research context and research questions and enabled us to capture the complexity and diversity of the phenomenon and related variables. However, a generic qualitative design also has some limitations, such as the possible lack of rigor, clarity, credibility, and generalizability of the study if not conducted properly (Queirós et al., 2017).

Participants

The research was conducted at 4 mid-tier universities across China to better represent the perceptions of most NESTs, NNESTs, and English majors. A total of 16 students, 4 NNESTs, and 3 NESTs participated in the study. Although it was initially designed to interview 1 NEST and 1 NNEST from each university, making a total of 4 NESTs and 4 NNESTs, which would result in a more balanced number of teacher participants from each university. Most of the NESTs were abroad at the time because of the Covid-19 pandemic, and few of them were available for online interviews. As a result, three NESTs from the United States of America with English as their first language participated in the study. Two had a non-English or education-related Bachelor's Degree and a Certificate in Teaching English to Speakers of Other Languages (TESOL), and the third one

had a PhD in English Literature. In comparison, three NNESTs had PhDs in English or education-related majors and one was a PhD student at the time when the data were collected. Table 1 provides the demographic information of the student participants of the quantitative study including students' gender, age, year of study, high school, and university.

Tables 2, 3 display the demographic information on the NESTs' and NNESTs' gender, nationality, first language, year of teaching, and the universities at which they were teaching at the time of data collection for this study.

The study employed convenience sampling as a result of the restrictions imposed by the Covid-19 pandemic, as our access to potential participants was limited and the participants had limited availability. The universities involved in this study adopted the national English syllabus for English majors, the *Syllabus for English Majors in Institutions of Higher Education* (Ministry of Education of China, 2000) for the provision of English major education, and the students were enrolled in four-year degree programs (He, 2020). The age of participants ranged from 19 to 24, with more than 10 years of English learning experience from the third grade of primary school to

TABLE 1 General information of student participants in the qualitative study.

		Frequency	Percent
Gender	Male	5	31.3
	Female	11	68.8
Age	19–21	12	75
	22–24	4	25
Year of study	Third-year	14	87.5
	Fourth-year	2	12.5
High school	Township	10	62.5
	City	6	37.5

TABLE 2 NEST participants for the qualitative study.

Questions	Responses from NEST participants		
Pseudonyms	Henry	Lucy	Claire
Gender	Male	Female	Female
Nationality	The USA	USA	USA
First language	English	English	English
Year of teaching	3–4 years for all NEST participants		
University	A	B	C

TABLE 3 NNEST participants for the qualitative study.

Questions	Responses from NNESTs participants			
Pseudonyms	John	David	Nora	Lily
Gender	Male	Male	Female	Female
Nationality	China	China	China	China
First language	Chinese	Chinese	Chinese	Chinese
Year of teaching	7–27 years for all NNEST participants			
University	A	B	C	D

high school. I selected juniors and seniors as the participants of this study since they had over 3 years of English learning experience and had undertaken courses taught by both NESTs and NNESTs at the university. The teacher participants were chosen based on their availability and teaching experiences, all of them had taught English for more than 2 years. Experienced teachers differed from novice teachers in time, teaching experiences, and classroom practice. They were more knowledgeable about the education system, students, and teaching material (Stronge, 2018), thus making them ideal for interviewing to gain insight into teaching effectiveness.

Instruments and data collection

To better understand students' and teachers' perceptions of teaching effectiveness, three semi-structured interview guides were designed. Students, NESTs, and NNESTs were asked to explain their perceptions of teaching effectiveness. For Chinese teachers and students, the interviews were translated into their native language (Chinese). The semi-structured interview guides contained main questions and follow-up sub-questions. The main questions encompassed the main content of the research questions and contained questions for all participants. The main questions acted as warm-up questions to familiarize participants with the subject and create a relaxing atmosphere. The follow-up sub-questions were intended to enhance the flow of the interview process and to gain more precise and detailed answers from the participants. They were selectively asked based on different participants' responses. The students were asked about their views of the effective qualities of English teachers, while the teachers were asked about their background, teaching experience, and opinions on the qualities of effective teachers. The interviews for students took 7 h and 24 min, with most of them lasting between 10 and 25 min. The interviews for teachers took 3 h and 39 min, with a majority of them being between 18 and 58 min.

To minimize the limitations of using qualitative data collection tools, we took measures to ensure the quality and rigor of the research process. We recruited a sufficient number of participants from different groups (students, NESTs, and NNESTs) to achieve representativeness and drew up our interview questions from the reviewed literature (Rao, 2010; Huang, 2017; Stronge, 2018) and the conceptual framework (Cohen, 1987; Holliday, 2018; Stronge, 2018) to ensure that they were relevant and valid for the research questions and objectives. For instance, Huang (2017) asked non-English majors' perceptions of NESTs, NNESTs, and their English accents, and some of their questions provided inspirations for the current study. Also, we conducted a pilot study with two students, two NESTs and two NNESTs to validate the interview questions.

Semi-structured interview guides for teachers

The interview guides for both native and non-native English-speaking teachers consist of the same questions. The teachers were asked about their linguistic, educational, and cultural background, which includes their nationality, first language, and experience of learning other languages. They were also asked about their English teaching experience, especially in China. The teachers were asked to explain their understanding of the qualities of an effective teacher, their teaching philosophies and approaches, and the formation of these philosophies and approaches. These questions aimed to

understand the identities of the participants and how teachers' beliefs of effective teachers' qualities have been formed and whether their teaching practices correspond to their beliefs, providing valuable information for further interpreting their perceptions.

Semi-structured interview guide for students

The interview guide for students was adjusted from those for teachers. The students were asked to explain their perceptions of the characteristics of an effective teacher, describe the best English teacher they had, and their best qualities. These questions aimed to understand students' beliefs of effective teachers' qualities and the formation of these beliefs. In addition, the students were asked about their perceptions of whether "nativeness" is a precondition for being an effective teacher, and the characteristics of native and non-native English-speaking teachers to understand the impact of "native-speakerism" on the students.

Data analysis

The semi-structured interviews were audio-recorded and transcribed by the researchers. The transcripts were checked for accuracy and anonymized, with the names of the participants being presented with pseudonyms. The qualitative data from the semi-structured interviews were analyzed using thematic analysis following Braun and Clarke's (2012) six-step process: (1) familiarization: the researchers immersed themselves in the data by listening to audio recordings, noting words of interest, thoroughly reading the data, and critically considering the meaning of the data; (2) generating initial codes: the transcripts were imported into NVivo 11 for Windows, the transcripts were coded line by line, the researchers assigned labels to segments of text to represent meaningful units of information. The codes were generated and modified as codes to capture the variety and patterns found within the data. For instance, an NNEST said "In terms of teaching, we are more compact, and also being able to know what the students want..." and it was coded as "perceived positive characteristics of NNESTs"; (3) searching for themes: the researchers examined the coded data to identify similarities and overlapping areas between codes, and group codes that were similar to create themes and subthemes. For example, the generated codes such as "(perceived) positive characteristics of NNESTs," and "(perceived) positive characteristics of NESTs" were found in three groups and they formed the theme "(perceived) differences between NESTs and NNETs"; (4) reviewing potential themes: the researchers recursively reviewed the generated themes to further refine and clarify them, and ensured that these themes related to the research question and objectives, the literature review, and the conceptual framework. (5) defining and naming themes: the researchers summarized the essence and clearly described the unique qualities of each theme (see Appendices A,B) and (6) producing the report: the researchers analyzed and synthesized the results, compared and contrasted the perspectives of different groups of participants, discussed the implications and limitations of the study, and presented them in the following sections.

Findings and discussion

The main findings of this study are based on the thematic analysis of 23 semi-structured interviews with teachers and students. The

generated themes from the thematic analysis of this section are presented in [Table 4](#).

In general, the students and NNESTs had more resemblances than the NESTs about effective teaching. Both emphasize teachers' teaching abilities, language abilities, and professionalism, and viewed knowledge as one of the most essential qualities. However, there were slight differences. The students paid more attention to personal qualities such as being confident, passionate, inclusive, humorous, strict, and having good pronunciation. On the other hand, the NNESTs were more concerned with understanding students' psychological development and learning processes and using theoretical knowledge to help students. The following conversation extracts from Ryan and John illustrate the typical students' and NNESTs' thoughts on effective teachers' qualities. The source from the male student Ryan is coded as (Ryan-MS). For Ryan, the most crucial qualities for teachers are having good knowledge, teaching skills, and being able to guide students and give feedback.

Ryan: Most importantly, an effective English teacher should be knowledgeable in many areas. English is a language, and language involves many areas, and the teacher should have a basic knowledge of each area. I think that having good speaking skills is an aspect that can attract the interest of students. Apart from these two requirements, a qualified teacher, he/she should have very solid teaching skills and be able to guide you on the academic side and give you constructive advice. (Ryan-MS)

For John, the most important requirements for teachers are also having good knowledge and teaching skills. He also stated that teachers must have a strong understanding of students' psychology and learning process.

John: I think the first and foremost thing is strong professional knowledge. For example, when you teach a class, your level of professionalism for teaching this class is definitely the most important. The second is your ability to teach. You have to be able to teach and know how to teach as a teacher... In addition, you need to understand their psychological development process. (John-MNNEST)

In comparison, the NESTs rarely brought up the qualities that Chinese students and NNESTs highly valued such as knowledge, teaching skills, and professionalism. They put more emphasis on affective qualities and teacher-student relationships such as caring,

patience, flexibility, engagement, and being aware of students' learning needs. Both Lucy and Clare placed great importance on affective qualities.

Lucy: I think you have to be patient. You have to be caring for your students. You have to be able to be flexible. I think you should be, I'm not sure, I guess that's it. Be flexible. Be caring. (Lucy-FNEST)

Clare: You need to be engaging. Chinese students have a hard time participating in class, so you need to be able to get them to volunteer and try because Chinese students have a hard time getting an answer wrong or speaking if they're not perfect. But you can't practice if you're afraid of never being imperfect. So you have to be very engaging with Chinese students. You have to have good pronunciation, and you have to be patient and have a good grasp of English in general, and know a lot of the grammatical rules of English. (Clare-FNEST)

The NESTs and the NNESTs also significantly differed in their teaching philosophies and approaches. The NESTs had a more humanistic view and emphasized a positive teacher-student relationship and mutual engagement in class. For example, one of the NESTs, Henry, compared teaching to having a journey with students, and they were all involved in this labor of love. On the other hand, the NNESTs were more practical. They preferred to talk about establishing solid knowledge and abilities for students, giving out homework, and preparing for students' future careers.

Overall, Chinese teachers and students hold more similar views on effective teachers' characteristics than NESTs. Both NNESTs and Chinese learners emphasize pedagogical aspects instead of affective characteristics, while NESTs valued affective qualities more than educational ones. These findings correspond to prior investigations conducted in the Chinese setting ([Rao, 2010](#); [Ma, 2016](#); [Huang, 2017](#); [Sarwal and Lamb, 2018](#)), which revealed that both students and NNESTs give great importance to essential characteristics such as knowledge, language abilities, instruction abilities, and professionalism. Research in general education also suggested that teachers, students, and university directors strongly prioritize subject knowledge ([Stronge, 2018](#)).

This is likely because the vast majority of EFL teachers in China are local NNESTs ([Pawan, 2017](#)). Most students only receive instruction from NNESTs from elementary school to high school. Both students and instructors are heavily influenced by the traditional

TABLE 4 Themes, codes, descriptions, and examples for effective teachers' qualities.

Theme	Code	Description and example
Effective teaching	Effective teachers' qualities and behaviors	Referring to teachers' qualities and behaviors that are regarded by learners to have a positive effect on learning outcomes.
		e.g., <i>Firstly, he should have a high level of professionalism. Secondly, he should be responsible. (Jennifer-FS)</i>
	Expectation of teaching	Learners' idealized teaching form and environment.
		e.g., <i>The expectations were that teachers would be able to show us more about the Western world, different from the Chinese model (Robert-MS)</i>
	Suggestions on improving teaching effectiveness	Learners' suggestions on improving the teaching effectiveness of English classes.
		e.g., <i>I think this is something that needs to be improved. I think I would like to learn more about culture from NESTs (Isabella-FS)</i>

teacher-centered educational model in China. This model accentuates teachers' authority and professional abilities and focuses on knowledge transmission (Yang and Zhang, 2015; Gao, 2017; Liu, 2021b). Teachers serve an important role as the source of information and keep track of students' learning progress, while students are seen as passive recipients of knowledge. It is common for teachers to give lectures for the majority of the class and students to attentively listen and take notes (Emaliana, 2017). In addition, NNESTs generally have a higher academic degree and more teaching experience. They are familiar with a variety of second language learning theories and students' psychology. Therefore, they often observe students' progress and potential to help students in their learning process and advancement (Xu, 2014). In contrast, NESTs rarely talk about these essential qualities. They give greater importance to affective characteristics such as being patient, caring, flexible, and sensitive to students. This could be explained by the fact that NESTs are usually recruited from inner-circle countries and have received Western education. For instance, the NESTs participants received their education in the USA. Western countries such as the USA have slowly moved from teacher-centered education to student-centered education (Bernard et al., 2019). NESTs are likely to have benefited from student-centered education and applied it when teaching Chinese EFL students. This educational approach focuses on being flexible in delivering content, encouraging student participation, and meeting learners' needs (Wright, 2011). Subsequently, NESTs have come to value personal and affective qualities (e.g., fairness and caring) that fit in with the student-centered educational model.

Research has demonstrated that both prerequisite qualities (e.g., knowledge and professionalism) and affective qualities are both critical to successful teaching (Al-Seghayer, 2017; Stronge, 2018; Li et al., 2022). As a result, rather than attempting to decide who has a better understanding of the qualities of an effective teacher, it is crucial that teachers and students gain a systematic understanding of the functions of the different qualities and strive to achieve a better alignment. Nevertheless, the largely uninformed opinions of the different groups imply that participants tend to place too much importance on certain qualities while disregarding others. For instance, NNESTs and students tend to focus on prerequisite qualities, while overlooking affective qualities. On the other hand, NESTs prioritize affective qualities over prerequisite qualities. These discrepancies in the perceptions of effective teaching qualities among the participants may have been caused by their diverse linguistic, educational, and cultural backgrounds, which additionally contribute to stereotypical positive and negative characteristics of NESTs and NNESTs.

Conclusion

This investigation examined the perspectives of students, NESTs, and NNESTs on teaching effectiveness for a comprehensive understanding of their similarities and disparities. Semi-structured interviews were conducted for teachers and students and thematic analysis was employed for the data analysis. Overall, this study illustrates the misalignments between students', NESTs', and NNESTs' perceptions of effective teachers' characteristics in the Chinese EFL context and discussed the underlying cultural, educational, and individual factors that contribute to the misalignments. The results demonstrated that Chinese students and NNESTs have more resemblance in their opinion of effective teacher characteristics than

NESTs. Both prioritize knowledge, teaching skills, and professionalism. In comparison, NESTs place more emphasis on individual qualities and student engagement such as being caring and flexible. These discrepancies could have been caused by differences in the participants' beliefs of educational models, given that NNESTs and students are familiar with the teacher-centered model while NESTs are more accustomed to the student-centered model. In conclusion, this research showed the misalignments between students, NESTs, and NNESTs in terms of their views on effective teacher characteristics in the Chinese EFL contexts, along with the cultural, educational, and individual aspects contributing to these differences.

The results of this study might have important implications for educators, learners, and university administrators. These implications could help enhance the effectiveness of teachers, and improve the collaboration between NESTs and NNESTs. The study has found that NEST, NNESTs, and students have different prioritizations on effective teachers' characteristics, which are likely to result from their different cultural, linguistic, and educational backgrounds. Research has shown that certain negative or neglected qualities of NESTs and NNESTs can be developed through training (Huang, 2017). Therefore, university administrators could offer training and education to NESTs and NNESTs to understand the complementary nature of various effective teacher characteristics in teaching, allowing them to bolster weaker areas and help improve instruction. For example, NNESTs could cultivate their affective qualities and be more caring, respectful, and fair to students, and learn about student-centered teaching methodologies (Wang and Fang, 2020). NESTs, in the meantime, could put more effort into understanding the Chinese culture and education system, adjusting their teaching materials and activities as needed, and making sure they are thoroughly prepared for class. NESTs and NNESTs could collaborate and teach together, as research has shown that co-teaching can result in the exchange of knowledge between teachers, and providing greater attention to learners' needs (Gbenakpon, 2018; Rao and Yu, 2021; Sanders-Smith and Dávila, 2021). It also supports the professional growth of teachers, as they can improve their teaching practices by observing and learning from colleagues, as well as through exchanging practical teaching knowledge and mutual emotional support through a conducive teaching and learning environment (Zhang et al., 2023).

The study has several limitations. This research was conducted in a Chinese EFL setting, making the findings possibly limited in their generalizability to other ELT contexts. The sampling may also lack generalizability to non-English majors with various proficiency levels since the participants were all English majors from universities in China. Additionally, due to the impact of Covid-19, the scale of the study had to be reduced and all the NESTs participants were from the USA. Qualitative research methods such as class observation were also excluded due to limited access to the university campus. Future studies are essential to reflect the multifaceted and dynamic identity of NESTs and NNESTs by recruiting participants with diverse linguistic and cultural backgrounds. Lastly, this study utilizes a geopolitical contrast between the two groups and regards English teachers from Inner Circle countries (Kachru et al., 2006) as NESTs, which may not be applicable in other contexts.

Another possible limitation is concerned with the trustworthiness of the study. Trustworthiness refers to the degree of confidence in methods, data, and interpretation used to conduct high-quality research.

“Trustworthiness or truth value of qualitative research and transparency of the conduct of the study is crucial to the usefulness and integrity of the findings” (Connelly, 2016). Clarke and Braun et al. (2017) divided the methods of providing the trustworthiness of thematic analysis (TA) into two broad types: ‘Small q’ and ‘Big Q’. The ‘Big Q’ is the utilization of techniques and instruments in the qualitative paradigm. The Big Q approach to TA works within a qualitative paradigm. It is theoretically independent and flexible, and the processes of creating codes and themes are organic. In the “Big Q” TA (e.g., the reflective thematic analysis), the analysis is more akin to a creative process than a technical process, as the researcher uses his own involvement, abilities, and analytical experiences in the analysis. Therefore, the reflective thematic analysis is more concerned with quality-assurance strategies such as a deeper engagement, and a thorough and systematic approach, than the accuracy of coding. These quality-assurance strategies were implemented to safeguard the trustworthiness of the current study. Although the student and teacher participants in four mid-tier universities may not reflect the diversity of all students and teachers. It is hoped that the findings can have implications in other similar settings. Further research is needed to address the limitations of this study. Further studies could explore how NESTs, NNESTs, and students can be better aligned in teaching practices so that they can improve their teaching toward high levels of effectiveness.

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 Ethics Committee of The University of

Auckland, New Zealand. The patients/participants provided their written informed consent to participate in this study.

Author contributions

LD and LZ conceived and designed the study. LD collected and analyzed the data and drafted the manuscript. LZ finalized it for submission as the corresponding author. All 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.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1175379/full#supplementary-material>

References

- Al-Seghayer, K. (2017). The central characteristics of successful ESL/EFL teachers. *J. Lang. Teach. Res.* 8, 881–890. doi: 10.17507/jltr.0805.06
- Bai, B., and Yuan, R. (2019). EFL teachers’ beliefs and practices about pronunciation teaching. *ELT J.* 73, 134–143. doi: 10.1093/elt/ccy040
- Bernard, R. M., Borokhovskiy, E., Schmid, R. F., Waddington, D. I., and Pickup, D. I. (2019). Twenty-first century adaptive teaching and individualized learning operationalized as specific blends of student-centered instructional events: a systematic review and meta-analysis. *Campbell Syst. Rev.* 15, 1–35. doi: 10.1002/CL2.1017
- Braine, G. (2010). *Nonnative speaker English teachers: Research, pedagogy, and professional growth*. New York, USA: Routledge.
- Braun, V., and Clarke, V. (2012). “Thematic analysis” in *APA handbook of research methods in psychology*. eds. H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf and K. J. Sher, vol. 2 (Washington, DC, USA: American Psychological Association), 57–71.
- Cheng, L., Burgess, D., Vernooij, N., Solis-Barroso, C., McDermott, A., and Nambodiripad, S. (2021). The problematic concept of native speaker in psycholinguistics: replacing vague and harmful terminology with inclusive and accurate measures. *Front. Psychol.* 12, 1–22. doi: 10.3389/fpsyg.2021.715843
- Cheng, X., and Zhang, L. J. (2022). Teachers helping EFL students improve their writing through written feedback: the case of native and non-native English-speaking teachers’ beliefs. *Front. Psychol.* 13, 1–15. doi: 10.3389/fpsyg.2022.804313
- Chomsky, N. (1965). *Aspects of the theory of syntax*, vol. 16, Massachusetts, USA: MIT Press, 393.
- Clarke, V., and Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, 12, 297–298. doi: 10.1080/17439760.2016.1262613
- Cohen, S. (1987). Instructional alignment: searching for a magic bullet. *Educ. Res.* 16, 16–20. doi: 10.3102/0013189X016008016
- Connelly, L. M. (2016). Trustworthiness in qualitative research. *Medsurg Nurs.* 25, 435–436.
- Cruz, S., and Auman, S. T. (2018). Perceptions of Korean students in Manila on non-native speaking teachers. *Oman J. ELT* 3:53. doi: 10.13140/RG.2.2.18534.04164
- Dai, W., and Wang, X. (2014). Woguo gaodeng jiaoyu neihanshi fazhan Beijing xia yingyu zhuanke de jianshe silu [on the construction of English major under the background of connotative development of higher education in China]. *Waiyujie [Foreign Language World]* 3, 2–11.
- Emaliana, I. (2017). Teacher-centered or student-centered learning approach to promote learning? *J. Soc. Human* 10, 59–70. doi: 10.12962/j24433527.v10i2.2161
- Fang, F. (2016). World Englishes or English as a lingua franca: where does English in China stand? *Engl. Today* 33, 19–24. doi: 10.1017/S0266078415000668
- Fang, F. G., Zhang, L. J., and Sah, P. (2022). Translanguaging in language teaching and learning: current practices and future directions. *RELC J.* 53, 305–312. doi: 10.1177/00336882221114478
- Farrell, T. S. C. (2015). It’s not who you are! it’s how you teach! Critical competencies associated with effective teaching. *RELC J.* 46, 79–88. doi: 10.1177/0033688214568096
- Galloway, N. (2017). *Global Englishes and change in English language teaching*. London, UK: Routledge.
- Gao, Y. (2017). Faithful imitator, legitimate speaker, playful creator and dialogical communicator: shift in English learners’ identity prototypes. In S. B. Said and L. J. Zhang (Eds.), *Intercultural dialogue: Questions of research, theory and practice*. Routledge, 59–75. doi: 10.4324/9780203795156

- Gbenakpon, S. A. (2018). An exploration of the concept and practice of co-teaching in public secondary school EFL classes in Benin. *J Lang Teach Res* 9, 765–776. doi: 10.17507/jltr.0904.13
- He, Z. (2020). Establishing disciplinary English in the curriculum for English majors in Chinese tertiary education. *J Lang Teach Res* 11, 575–582. doi: 10.17507/jltr.1104.07
- Hennebry-Leung, M., and Gao X. (Andy). (2022). Language learning motivation in a multilingual Chinese context. Routledge. doi: 10.4324/9781003025122
- Holliday, A. (2018). “Native-speakerism” in *The TESOL Encyclopedia of English language teaching*. eds. J. I. Lontos, M. Delli Carpini, G. Kessler and S. J. Nero (Hoboken, New Jersey, USA: John Wiley & Sons), 1–7.
- Huang, Z. (2017). *Native and non-native English speaking teachers in China: Perceptions and practices*. Singapore: Springer.
- Javid, C. Z. (2016). Teaching effectiveness of native and non-native EFL teachers as perceived by preparatory year students in Saudi context. *Lang India* 46, 340–349. doi: 10.1093/elt/46.4.340
- Kachru, B. B., Kachru, Y., and Nelson, C. L. (2006). *The handbook of world Englishes*. 1st Edn Blackwell.
- Kiczowski, M. (2019). Students’ teachers’ and recruiters’ perception of teaching effectiveness and the importance of nativeness in ELT. *J Sec Lang Teach Res* 7, 1–25.
- Lavin, A., Korte, L., and Davies, T. (2012). Student gender and perceptions of teaching effectiveness. *Res. High. Educ. J.* 18, 1–16.
- Li, X., Bergin, C., and Olsen, A. A. (2022). Positive teacher-student relationships may lead to better teaching. *Learning and Instruction*, 80, 101581. doi: 10.1016/j.learninstruc.2022.101581
- Liu, J. (2021a). Revisiting native speakerism in ELT: viewpoints of Chinese EFL program administrators on the recruitment and workplace situations of foreign English teachers. *Engl. Lang. Teach.* 14, 24–31. doi: 10.5539/elt.v14n9p24
- Liu, J. (2021b). “The influence of modern information technology on college English teaching under the background of educational information” in *Advances in human factors, business management and leadership*. eds. J. I. Kantola, S. Nazir and V. Salminen (Cham, Switzerland: Springer), 350–355.
- Llurda, E., and Calvet-Terré, J. (2022). Native-speakerism and non-native second language teachers: a research agenda. *Lang. Teach.*, 1–17. doi: 10.1017/s0261444822000271
- Lo, Y. Y. (2015). A glimpse into the effectiveness of L2-content cross-curricular collaboration in content-based instruction programmes. *Int. J. Biling. Educ. Biling.* 18, 443–462. doi: 10.1080/13670050.2014.916654
- Ma, L. P. F. (2016). Examining teaching behaviour of NNESTs and NESTs in Hong Kong through classroom observations. *Asian J Appl Linguist* 3, 199–215.
- Meissel, K., Meyer, F., Yao, E. S., and Rubie-Davies, C. M. (2017). Subjectivity of teacher judgments: exploring student characteristics that influence teacher judgments of student ability. *Teach. Teach. Educ.* 65, 48–60. doi: 10.1016/j.tate.2017.02.021
- Ministry of Education of China. (2000). *《高等学校英语专业英语教学大纲》[the syllabus for English majors in institutions of higher education]*. Beijing: Ministry of Education of China.
- Moussu, L. (2018). “Shortcomings of NESTs and NNESTs” in *The TESOL Encyclopedia of English language teaching*. eds. J. I. Lontos, M. Delli Carpini, G. Kessler and S. J. Nero (Hoboken, New Jersey, USA: John Wiley & Sons), 1–6.
- Nomura, K., and Mochizuki, T. (2018). “Native-speakerism perceived by “non-native-speaking” teachers of Japanese in Hong Kong” in *Towards post-native-speakerism: Dynamics and shifts*. eds. S. A. Houghton and K. Hashimoto (Singapore: Springer), 79–95.
- Pawan, F. (2017). “Situating Chinese English language teacher training and professional development in research and policy” in *Teacher training and professional development of Chinese English language teachers: Changing from fish to dragon*. eds. E. Hinkel, F. Pawan, W. Fan, M. Pei, G. Wang and W. Jinet al. (New York: Routledge), 1–9.
- Qiu, X., and Fang, C. (2022). Creating an effective English-medium instruction (EMI) classroom: Chinese undergraduate students’ perceptions of native and non-native English-speaking content teachers and their experiences. *Int. J. Biling. Educ. Biling.* 25, 641–655. doi: 10.1080/13670050.2019.1707769
- Queirós, A., Faria, D., and Almeida, F. (2017). Strengths and limitations of qualitative and quantitative research methods. *Eur J Educ Stud* 3, 369–387. doi: 10.5281/zenodo.887089
- Ranjit, R. (2022). Classroom challenges of secondary level English teachers at remote government schools in Nepal. *J Educ Lang Innov Appl Linguist* 1, 10–25. doi: 10.37058/jelita.v1i1.4443
- Rao, Z. (2010). Chinese students’ perceptions of native English-speaking teachers in EFL teaching. *J. Multiling. Multicult. Dev.* 31, 55–68. doi: 10.1080/01434630903301941
- Rao, P. S. (2019). The role of English as a global language. *Res J Engl* 4, 65–79.
- Rao, Z., and Yu, H. (2021). Enhancing students’ English proficiency by co-teaching between native and non-native teachers in an EFL context. *Lang. Teach. Res.* 25, 778–797. doi: 10.1177/1362168819873937
- Roothoof, H., and Breeze, R. (2016). A comparison of EFL teachers’ and students’ attitudes to oral corrective feedback. *Lang. Aware.* 25, 318–335. doi: 10.1080/09658416.2016.1235580
- Saleem, M., Larik, A. R., and Bukhari, S. G. A. S. (2022). Loopholes in developing English communicative competence: a comparative case study of Pakistan and China. *Res. Comp. Int. Educ.* 1–22:174549992211284. doi: 10.1177/17454999221128474
- Sanders-Smith, S. C., and Dávila, L. T. (2021). ‘It has to be in a natural way’: a critical exploration of co-teaching relationships in trilingual preschool classrooms in Hong Kong. *J. Multiling. Multicult. Dev.* 1–15, 1–15. doi: 10.1080/01434632.2021.1957902
- Sarwal, A., and Lamb, M. (2018). “Learner perceptions of inspiring English teachers in northern India” in *Explorations: Teaching and Learning English in India*, eds. T. Brian and K. Andy (Kolkata, India: British Council), vol. 11, 20–30.
- Selvi, A. F. (2014). Myths and misconceptions about nonnative English speakers in the TESOL (NNEST) movement. *TESOL J.* 5, 573–611. doi: 10.1002/tesj.158
- Shan, Z., and Xu, H. (2022). Exploring multilingual awareness development in learners of multiple foreign languages: a social cognitive perspective. *J Multiling Multicult Dev.* 1–14. doi: 10.1080/01434632.2022.2092119
- Stronge, J. H. (2018). *Qualities of effective teachers*, vol. 1 Alexandria, Virginia, USA: ASCD, 82–87.
- Sun, P. P., Zhang, L. J., and Sah, P. (2022). Effects of translanguaging in online peer feedback on Chinese university English-as-a-foreign-language students’ writing performance. *RELJ* 53, 325–341. doi: 10.1177/00336882221089051
- Szimigiera, M. (2021). The most spoken languages in 2021 worldwide. Available at: <https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/>
- Taguchi, N. (2014). English-medium education in the global society. *Int Rev Appl Linguist Lang Teach* 52, 89–98. doi: 10.1515/iral-2014-0004
- Tatipang, D. P., Manuas, M. J., Wuntu, C. N., Rorintulus, O. A., and Lengkoan, F. (2022). EFL students’ perceptions of the effective English teacher characteristics. *J Pendidikan Bahasa Inggris Undiksha* 10, 23–30. doi: 10.23887/jpb.v10i1.45709
- Wang, L., and Fang, F. (2020). Native-speakerism policy in English language teaching revisited: Chinese university teachers’ and students’ attitudes towards native and non-native English-speaking teachers. *Cogent Educ* 7, 1–22. doi: 10.1080/2331186X.2020.1778374
- Wright, G. B. (2011). Student-centered learning in higher education. *Int J Teach Learn High Educ* 23, 92–97.
- Xu, Y. (2014). Becoming researchers: a narrative study of Chinese university EFL teachers’ research practice and their professional identity construction. *Lang. Teach. Res.* 18, 242–259. doi: 10.1177/1362168813505943
- Yang, H. (2018). An analysis of the relationship between Chinese EFL teachers’ agency and beliefs from an activity theory perspective. *Asian EFL J.* 20, 31–53.
- Yang, C., and Zhang, L. J. (2015). China English in trouble: evidence from teacher dyadic talk. *System* 51, 39–50. doi: 10.1016/j.system.2015.03.008
- Zhang, F., and Zhan, J. (2014). The knowledge base of non-native English-speaking teachers: perspectives of teachers and administrators. *Lang. educ.* 28, 568–582. doi: 10.1080/09500782.2014.921193
- Zhang, L. J. (2004a). Awareness-raising in the TEFL phonology classroom: student voices and sociocultural and psychological considerations. *ITL Int J Appl Linguist* 145–146, 219–268. doi: 10.2143/ITL.145.0.562915
- Zhang, L. J. (2004b). Reforming a teacher education program for PRC EFL teachers in Singapore: sociocultural considerations and curriculum evolution. *Int J Educ Reform* 13, 223–252. doi: 10.1177/1056787904013003
- Zhang, L. J. (2021). Curriculum innovation in language teacher education: reflections on the PGDELT Program’s contribution to EFL teachers’ continuing professional development. *Chin J Appl Linguist* 44, 435–450. doi: 10.1515/CJAL-2021-0028
- Zhang, L. J., and Cheng, X. L. (2021). Examining the effects of comprehensive written corrective feedback on L2 EAP students’ performance: a mixed-methods study. *J. Engl. Acad. Purp.* 54, 101043–101015. doi: 10.1016/j.jeap.2021.101043
- Zhang, L. J., Fathi, J., and Mohammaddokht, F. (2023). Predicting teaching enjoyment from teachers’ perceived school climate, self-efficacy, and psychological wellbeing at work: a case of EFL teachers. *Percept. Mot. Skills* 1–18:2269. doi: 10.1177/00315125231182269
- Zhang, L. J., and Said, S. B. (Eds.). (2014). *Language teachers and teaching: Global perspectives, local initiatives*. Routledge. doi: 10.4324/9780203795156
- Zhang, L. J., and Sun, Q. (2022). Developing and validating the English teachers’ cognitions about grammar teaching questionnaire (TCAGTQ) to uncover teacher thinking. *Front. Psychol.* 13, 1–10. doi: 10.3389/fpsyg.2022.880408
- Zhang, T., Wang, Z., Liu, G., and Shao, J. (2019). Teachers’ caring behavior and problem behaviors in adolescents: the mediating roles of cognitive reappraisal and expressive suppression. *Personal. Individ. Differ.* 142, 270–275. doi: 10.1016/j.paid.2018.10.005
- Zhang, J., and Zhang, L. J. (2021). Learners’ satisfaction with native and non-native English-speaking teachers’ teaching competence and their learning motivation: a path-analytic approach. *Asia Pac J Educ* 41, 558–573. doi: 10.1080/02188791.2020.1833834
- Zhao, H., and Zhang, L. J. (2022). *Teaching writing in English as a foreign language: Teachers’ cognition formation and reformation*. Cam, Switzerland: Springer Nature International.



OPEN ACCESS

EDITED BY

Pedro Gil-Madróna,
University of Castilla-La Mancha, Spain

REVIEWED BY

Keri L. Heitner,
Saybrook University, United States
Antonio Morcillo Martínez,
University of Castilla-La Mancha, Spain

*CORRESPONDENCE

Hsinmei Liao
✉ hliao@mail.tcu.edu.tw

RECEIVED 17 June 2023

ACCEPTED 09 August 2023

PUBLISHED 23 August 2023

CITATION

Liao H (2023) Mapping principles and worked examples for structural learning: effects of content complexity.
Front. Psychol. 14:1241873.
doi: 10.3389/fpsyg.2023.1241873

COPYRIGHT

© 2023 Liao. 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.

Mapping principles and worked examples for structural learning: effects of content complexity

Hsinmei Liao*

Department of Human Development and Psychology, Tzu Chi University, Hualien City, Taiwan

Drawing connections between principles and worked examples is an approach to learning and instruction, but it is poorly understood. This study investigated the effects of principle and example complexity on learners' ability to map principles and worked examples. The complexity of a principle or example was determined based on the number of concepts and relationships involved. 138 college students were randomly assigned to one of the mapping conditions: principle–simple example, principle–complex example, simple example–simple example, simple example–complex example, and complex example–complex example. The participants studied related materials and completed a free-mapping and a guided-mapping task for a simple and a complex probability principle. The effects of the mapping activities were measured in terms of gains in structural and conceptual knowledge. For the simple principle, principle–example mapping led to fewer nonrelational comparisons (standalone concepts) than did example–example mapping and an equal number of relational comparisons (interconnected concepts). For the complex principle, principle–example mapping led to fewer nonrelational but more relational comparisons than example–example mapping did. Principle–example mapping of corresponding content was more difficult than example–example mapping was. However, principle–example mapping of noncorresponding content was as easy as or easier than example–example mapping. The two forms of mapping resulted in equivalent gains in structural and conceptual knowledge. The findings of this study expand the understanding of analogical reasoning and learning through mapping and comparison of abstract and concrete content. The findings indicate that principle–example mapping enables learners to overcome the obstacles of comprehending abstract or general information and to identify the interrelationships of the individual concepts in formal structures.

KEYWORDS

principle learning, worked example, mathematics learning, content complexity, analogical comparison and mapping, relational learning, conceptual knowledge, structural knowledge

1. Introduction

For school subjects such as mathematics and physics, learning involves acquiring concepts and their definitions and, occasionally, symbols representing the concepts. A crucial goal of learning is the formation of relations or structures among concepts because relations may illustrate or define some principles, such as rules, laws, and theorems, or more complex

phenomena. To illustrate, for any two numbers a and b , we can calculate their sum raised to the power of n , which can be represented by $(a + b)^n$. For all the possible n values, the results can be collectively represented through the binomial theorem $(a + b)^n = \sum_{x=0}^n \binom{n}{x} a^x b^{n-x}$. Within the theorem, the concepts of a binomial, an exponent, a sum, and a binomial expansion are linked by relations. Principles are frequently delineated in general or abstract statements to express the relations among the concepts underlying such statements. Such principles are referred to as *principled information* in the present study.

Directly tackling to-be-learned principles can be difficult because principles are abstract by nature (e.g., Chi et al., 1989; Koedinger et al., 2008). Concrete examples are often used to present situations in which a target principle can be applied or how the principle can be used to solve problems. The examples presented in the present study serve both of these functions and are considered *worked examples*. The following illustrates when the binomial theorem can be applied. Assume a random incident results in one of two outcomes A or B and the probability of A occurring is 0.2 and that of B occurring is 0.8. We can determine the probability of A happening 3 times and B happening 7 times, regardless the order of occurrence, by this part of the theorem $\binom{n}{x} a^x b^{n-x}$.

Understanding the associations between principled information and worked examples is also a crucial learning goal because worked examples serve as actualizations of target principles. Other forms of information, such as tables or figures, can supplement or even serve as targets of learning. Connecting various forms of information is the ultimate goal of learning because understanding involves developing a web of interconnected knowledge (e.g., Lopez et al., 2014). Not only is understanding valuable in itself, it is also valuable for later problem-solving. Developing a network of knowledge, however, can be difficult (e.g., Renkl, 1997; Dixon, 2005; Schwartz et al., 2011; Belenky and Schalk, 2014).

The focus of the present study was the connections learners form between domain principles and worked examples. Examples may be the most frequently used instructional aids, and the structural information they convey should reflect target principles. Numerous studies have investigated how worked examples can be used to teach target principles (e.g., Gerjets et al., 2006; Nokes-Malach et al., 2012), for example, by pairing the worked examples with instructional explanations. However, these studies have typically focused on how worked examples should be implemented rather than on domain principles. The current study investigated the characteristics of directly connecting principled information and related worked examples through mapping. This work would contribute to deepen the understanding of how principled information and worked examples become associated and to formulate effective instructional and learning strategies.

2. Connecting as analogical mapping

In analogical thinking or problem-solving, an individual uses the solution to a source problem to solve a target problem. If the source and target problems have a common conceptual structure despite belonging to different content domains, the target problem may be solved. Cognitive psychologists have elucidated the

processes underlying analogical reasoning (e.g., Gick and Holyoak, 1980; Gentner, 1989). The individual first retrieves an appropriate analogue or solution to a source problem. Then, they align or map the entities and relations of the source and target problems. In next steps, the individual makes inferences regarding the target problem on the basis of their knowledge of the source problem and generalizes or transfers the solution of the source problem to the target problem.

An analogy generally involves two targets of concrete problems or situations, such as determining whether two worked examples have the same conceptual structure of a binomial distribution. The literature on analogical learning and comparison has offered a wealth of evidence regarding the effects of such learning and its incorporation into instructional strategies (e.g., Gentner et al., 2003). In the current study, the definition of analogy was expanded to include instances involving one concrete and one abstract situation, that is, involving a worked example and its underlying principle. In a learning context in which a source and a target problem are provided to a learner, mapping is likely the first step the learner takes because through mapping, the learner can determine similarities in the conceptual structures of the two problems and thereby delineate the conceptual structure shared by the problems. In light of this, a new strategy of learning by mapping or comparing an example and its underlying principle was proposed, and its potential applicability was explored.

This hybrid approach was developed with consideration of the following. First, for many learners, principles are too abstract to understand and too general to apply. By contrast, an example may illustrate only certain aspects of its underlying principle and be too specific for generalizations to be made (e.g., Reed and Bolstad, 1991). Therefore, learning by connecting the two forms of material may be more effective than learning either alone (e.g., Chi et al., 1989; Fyfe et al., 2014) because they compensate for each other's disadvantages. Second, the conceptual structure of a principle is complete and can serve as a model for comparison. In cases in which examples or problems do not illustrate all aspects of a conceptual structure, the learner is likely to be prompted to further investigate and acquire the remaining aspects of the structure (Schwartz et al., 2011; Lachner and Nückles, 2015). Third, principle-example mapping may enable learners to focus on structural similarities between a source and a target because principles involve general entities and relations rather than distracting contextual information. These three considerations highlight the potential advantages of the proposed hybrid approach over traditional mapping approaches. Because no study on such a hybrid approach was identified in the literature, the following presents a review of empirical studies on analogical reasoning and learning.

3. Connecting worked examples

Analogizing concrete situations is crucial in both instruction and learning and has been researched and applied in various school subjects (Alfieri et al., 2013; Dumas et al., 2013; Richland and Simms, 2015; Goldwater and Schalk, 2016). However, simply providing a learner with a source and target problem may not lead to the desired learning outcomes. Gentner et al. (2003) discovered that comparing two examples and participating in analogical encoding (i.e., a guided comparison of two examples) both resulted

in more favorable learning and transfer than did studying the two examples in order and the baseline condition, that is, the participants simply solving the problems. Without guidance, many learners do not engage in active learning processes, such as analogical comparison (Chi et al., 1989; Renkl, 1997). Therefore, learning content must be deliberately connected.

Furthermore, learners may struggle to align or map the elements of different content (e.g., Bassok, 1990; Ross and Kilbane, 1997). To ensure more favorable learning outcomes, learners must be guided through analogical mapping with instruction. Such external support can lead them to focus the comparisons on structural components and thus increase the likelihood of the learner comprehending the underlying principle. If the learner does not receive guidance, they may focus on superficial features of the problems provided that are not relevant to the learning goals (e.g., Schwartz et al., 2011; Day et al., 2015). Several studies (Gentner et al., 2003; Richland and McDonough, 2010) have discovered that comparisons conducted with guidance are generally more beneficial than comparisons conducted without and have reported the effects of improvements in solving novel problems, or procedural knowledge, and in conceptual understanding, or schematic knowledge (e.g., Cummins, 1992; Gentner et al., 2003; Rittle-Johnson and Star, 2007; Richland and McDonough, 2010; Nokes-Malach et al., 2012).

With the exception of asking learners to map or compare two isomorphic problems with the same solution (e.g., Ross and Kennedy, 1990; Novick and Holyoak, 1991; Cummins, 1992), asking learners to map or compare problems that can be solved using different solution procedures but involving the same underlying principle can empower them to more effectively learn the conceptual structure of the principle (e.g., Catrambone and Holyoak, 1990; Reed et al., 2012). Such nonisomorphic problems often have different subgoals and therefore illustrate specific aspects of the target principle. These differences cause the problems to appear different, and different steps to be required to obtain the necessary values for obtaining a solution. Research indicates that comparing such problems leads to greater success in solving isomorphic problems and awareness of the conceptual aspects reflected in the problems' subgoals (e.g., Catrambone and Holyoak, 1990).

Some studies have obtained differential findings regarding the learning effects of comparing these types of problems (Reed, 1989; Reed et al., 2012). For example, in Reed et al. (2012), for some subtypes of work problems (e.g., $\text{distance} = \text{rate1} \times \text{time1} + \text{rate2} \times \text{time2}$), comparing problems with arithmetic solutions resulted in an improved ability to solve similar problems, whereas comparing problems with algebraic solutions did not. In addition, comparing arithmetic problems did not lead to a learning transfer to algebraic problems or vice versa. Comparing problems of these two kinds did not lead to improvement on learning, either. These findings imply that when problems are structured differently, even though the elements involved and the solution methods are the same, their difficulty or complexity is different.

As the studies on analogical learning have demonstrated, learning through mapping or comparing example problems can be improved through deliberate guidance that highlights conceptual structures. In addition, coupling example problems of different complexities may facilitate structural and conceptual learning. However, the beneficial effects of such coupling are not well understood because they are not consistently identified.

4. The present study

Because the literature on how learners perform principle–example mapping is lacking, this study conducted an investigation of learners' behavior under the condition of no formal instruction being provided. Three research goals were established. The primary goal was to examine the characteristics of principle–example mapping, hereinafter referred to as *PE mapping*. To enable a comparison, example–example mapping, hereinafter referred to as *EE mapping*, was also examined. Although principles are deemed important for relational learning (e.g., Goldstone and Son, 2005), empirical evidence supporting the benefits of PE mapping in this context is lacking. Therefore, the study initially investigated learners' tendencies regarding structural mapping during learning tasks to accomplish this goal. Additionally, understanding the relative difficulty of PE and EE mapping can be valuable for educators interested in adopting these mapping approaches for instructional material planning. Thus, the second issue addressed in this study was the systematic investigation of the difficulty associated with PE and EE mapping.

The secondary goal aimed to explore the factors that may affect PE and EE mapping. Content complexity might play a role in relational learning, however, little research has delved into this aspect. Therefore, this study explored two sources of complexity, namely principles and examples, to determine whether and how they affect learning through mapping. *Complexity* was defined as the number of entities and relations within a form of content, focusing on structural learning (e.g., Halford, 1992). Simple principles had fewer elements and thus lower structural complexity, whereas complex principles had the opposite characteristics. Similarly, simple examples provided all required components for direct application of a principle, whereas complex examples included some given components but required finding others on the basis of available information regarding entities and relationships.

To examine the conception of conceptual structures, the participants' responses to the unconstrained, free-mapping task were analyzed in terms of nonrelational and relational responses according to the structure-mapping theory (Gentner, 1983). *Nonrelational* responses were mappings that did not involve relational elements, such as entities, properties, or property relationships (e.g., a continuous space). *Relational* responses were mappings that involved relational elements, such as first-order (e.g., events occurring in a space) or higher-order relationships (e.g., events occurring in a space independent of events in another space). The connections between principle and example complexity and relational learning are discussed in greater detail in the “Predictions” section.

To summarize, the research questions for the above goals were as follows:

1. How do the connections established in PE mapping and EE mapping differ?
2. Is PE mapping more difficult than EE mapping?
3. Does the complexity of an example affect the results for Questions 1 and 2?
4. Does the complexity of a principle affect the results for Questions 1 and 2?

For instructors and learners to feel motivated to adopt PE mapping as a strategy, the effects of such mapping must be understood.

Therefore, the final goal of the study was to explore the effects of PE mapping on learning. As previously indicated, the PE mapping approach was expected to facilitate structural learning; thus, the learning effect to be evaluated was concerned with the learners' structural knowledge.

Researchers, however, do not have a consensus on how the outcomes of structural learning should be measured. Some have measured whether a learner could articulate connections among the entities involved in word problems. That is, they have measured the results or schemas of problem-solving (e.g., Bassok, 1990; Novick and Holyoak, 1991). While the acquisition of abstracted knowledge necessitates understanding individual entities and their interrelations, it also entails comprehending the meaning, properties, and constraints associated with these elements. In this study, these forms of knowledge were referred to as *structural* and *conceptual knowledge*, respectively. Research has revealed that acquiring one form of knowledge may not guarantee acquisition of the other (e.g., Rittle-Johnson and Alibali, 1999). Therefore, in assessing the effects of PE mapping in structural learning, both structural and conceptual knowledge should be considered.

For the final goal, the research questions investigated were as follows:

5. Is the conceptual knowledge acquired through PE mapping different from that acquired through EE mapping?
6. Is the structural knowledge acquired through PE mapping different from that acquired through EE mapping?
7. Does the complexity of an example modulate the differences of Questions 5 and 6?
8. Does the complexity of a principle modulate the differences of Questions 5 and 6?

4.1. Design

A mixed-design experiment was employed for the study. The participants were tasked with studying foundational materials related to the target content. They then completed a free-mapping and a guided-mapping task based on the assigned condition, followed by tests assessing their conceptual and structural knowledge.

The independent variables included mapping type (PE or EE), example complexity (simple or complex), and principle complexity (simple or complex). Mapping type and example complexity served as between-subjects variables and were combined to create five experimental conditions: principle-simple example (PEs), principle-complex example (PEc), simple example-simple example (EsEs), simple example-complex example (EsEc), and complex example-complex example (EcEc). Principle complexity was a within-subject variable; participants assigned to each experimental group completed tasks at each level of principle complexity.

Based on the study's definition of complexity, Geometric and Poisson distributions were, respectively, chosen as the simple and complex principles. Examples were designed so that the simple examples required one step to obtain a solution and the complex examples required three steps. Demonstrations of the principles and worked examples at two complexity levels can be found in [Appendices A2, A3](#).

The dependent variables were scores derived from the tasks the participants completed. For each principle, the following scores were assigned for every participant: mapping difficulty, nonrelational responses, relational responses, and the sum score of nonrelational and relational responses, conceptual knowledge, and structural knowledge. Further details are provided in the "Methods" section.

4.2. Predictions

The present study hypothesized that, in general, PE mapping would facilitate relational comparison whereas EE mapping would facilitate nonrelational comparison. Researchers have agreed that teaching principles offers the advantage of presenting the elements of a conceptual structure without irrelevant, distracting features (Goldstone and Son, 2005; Sloutsky et al., 2005; Koedinger et al., 2008). Although the abstractness of principles causes PE mapping to be more difficult than EE mapping, it can also effectively highlight the roles of the entities and relations among the entities involved in an example (Goldstone and Son, 2005). Consequently, PE mapping is more likely than EE mapping to lead to relational comparison and is less likely to lead to nonrelational comparison. These differences between PE and EE mapping, however, may not be substantial if a principle is simple.

Differing levels of example complexity may also affect mapping patterns. As example complexity increases, the likelihood of a learner employing relational comparison might also increase because the examples demonstrate the process of transformation to highlight the relationships among the entities, which can facilitate relational processing. When the complexity is higher in a PE pair, PE mapping is likely to yield even more relational comparisons and fewer nonrelational comparisons than EE mapping is. When the complexity is higher in an EE pair, the differences between the results of PE and EE mapping are likely less important because more relational comparisons are likely to be made in EE mapping.

This study also hypothesized that PE mapping would generally be more difficult than EE mapping. According to the structure-mapping theory (Gentner, 1983), regardless of whether they are applying PE or EE mapping, learners attempt to identify one element of a target that corresponds to a single element of the second target. When a learner knows nothing about a domain, they may consider any element of the first target to correspond to more than one element of the second. In EE mapping, a learner may more easily identify matches (e.g., Novick, 1988; Bassok and Olseth, 1995; Ross and Kilbane, 1997) when they consider the role of an element, that is, the predicate of an argument (Gentner, 1983), because the attributes or role of an element within a specific context can be more easily extracted. In PE mapping, however, a lack of context cues and the generality of the terms used to describe the principle can increase the difficulty of determining the attributes or roles of elements, which increases the difficulty of matching elements. Therefore, when the complexity of a principle is higher, the differences in difficulty between PE and EE mapping may become more significant.

The difficulty of PE and EE mapping may also change with the complexity of an example. A pair of items containing at least one complex example for PE or EE mapping is likely to be more difficult to map than a pair containing two simple and similar items. To comprehend and map such pairs, the learner must make additional

inferences and transformations to render the two items comparable. Although undertaking such a process can be an opportunity for learning, it may also be perceived to be an additional challenge to the learner, particularly during the initial learning period. When the example complexity is higher in a PE pair, the difference in the difficulty of PE and EE mapping may become more significant (because PE mapping is generally more difficult than EE mapping). However, when example complexity is higher in an EE pair, the difference may become less significant (because the difficulty of EE mapping is higher).

The effects of mapping on structural and conceptual learning can be viewed as the cumulative outcome of two mapping tasks evaluated under the condition of no intervention. The aforementioned discussions indicate that as the principle complexity or example complexity within a PE pair increases, the benefits of relational processing due to the principle may also increase in PE mapping. However, at the same time, the disadvantage arising from the abstract nature of the principle for PE mapping may also intensify. It remains unclear how the positive and negative effects of principles interact with each other to influence the overall learning outcomes. As a result, no specific predictions were made regarding the differences between PE and EE mapping.

5. Methods

5.1. Participants

Undergraduate students were recruited through advertisements posted around the campuses of a small university. Students were considered eligible if they had not taken classes on the selected topics. The sample size was determined to detect a standardized effect size of at least 0.35, resulting in a target size 145. Details regarding this estimation are provided in the [Supplementary materials](#). In the end, 149 participants took part in the study, but 11 participants were excluded from the analysis due to a high number of missing responses or failure to correctly follow the task instructions. The study included 103 female and 35 male participants; the numbers of female and male participants were similar across the five conditions.

5.2. Materials

Three sets of printed materials were used in this study. The contents were adapted from standard textbooks that were suitable for initial learning of the topics. The *prerequisite-study booklet* contained relevant probability concepts (e.g., random experiments, random variables, probability function and distribution, and expected values) that the participants could familiarize themselves with. The contents of the booklet are provided in [Appendix A1](#).

The *work booklet* contained tasks used to measure the participants' behavior of mapping. A version of the booklet was developed for each of the five experimental conditions. Each version comprised six pages: half of the pages presented tasks for the simple principle, and half presented tasks for the complex principle. The order of the tasks illustrating the two types of principles was counterbalanced within each condition. For each half of a booklet, the first page presented a principle and worked example or two worked examples. The PE and

EE pairs were prepared for the simple and complex principles using the following criteria: One example of the PE and EE pairs was presented in the same context, and the second example of the EE pair was presented in a different context. Representative PE pairs are presented in [Appendices A2, A3](#).

The second page of each half of a booklet contained a *free-mapping* task, and the remaining pages contained a *guided-mapping* task (examples provided in [Appendices A4, A5](#)). The free-mapping task was used to measure the participants' behavior when minimal guidance was provided. The participants' performances would indicate their natural tendencies or propensities with respect to mapping. In the task, the participants were asked to identify the commonalities and the differences between two targets (e.g., [Markman and Gentner, 1996](#)).

The guided-mapping task was used to measure the participants' behavior when they received guidance. The participants were more likely to demonstrate their abilities of mapping when provided with guidance (e.g., [Catrambone and Holyoak, 1990](#)). Thus, this task was used to determine the difficulty of mapping. A list of 12 minor components of the target principle or example was provided for each principle. For each set of responses, the participants were asked to construct a stem for mapping by using the provided components and to identify a corresponding part in another target example. The numbers of components required to construct the stems were two, three, four, or five or more. For each size, up to three sets of responses could be given. This task was presented over two pages.

A third booklet contained two tasks for measuring the effect of mapping. This *test booklet* was the same for all experimental conditions. However, within each condition, the task order was counterbalanced with the mapping order. One of the tasks was the *multiple-choice task*, which was used to assess the participants' conceptual understanding of the aspects of the target principles. The task comprised two parts, each of which contained five questions on one of the target principles. The order in which the two parts were presented was the same as that in which the target principles were presented in the work booklet. The questions were described in general terms, and only one option among four was correct.

The other task was the *word-problem task*, which was used to assess the participants' knowledge of the overall conceptual structures of the two target principles. The task comprised two problems related to each of the two principles and two additional problems regarding other forms of distribution that were used as foils. One of the two problems regarding each of the two principles was a simple problem, and the other was a complex problem. These six word problems were presented in the same semirandom order for each experimental condition. The instructions specified that some of the problems could be solved by using one of the two target principles and that if a problem could be solved using one of these principles, only the equation associated with the principle needed to be provided.

5.3. Procedure

The study commenced after approval was granted by a research ethics committee. Groups of eight participants or fewer completed the experiment in a laboratory on a campus or a designated room in the library of another. The participants were seated with sufficient space between them to work independently and receive individual help from

the experimenter if needed. The participants were randomly assigned to an experimental condition upon their arrival through the block randomization procedure and were not informed of which condition they had been assigned to. Both the experimenters and the participants were blinded to the research questions and hypotheses. The experimenters asked the participants to sign a consent form and comply with the requirements of the experiment. Afterwards, the participants studied the prerequisite-study booklet, completed the work booklet for their assigned condition, and then completed the test booklet within a single session. The booklets were provided all at once but the participants were instructed to work on one booklet at a time. The participants were encouraged to attempt to comprehend and memorize the learning materials but were allowed to view the pages containing the PE or EE pair as they completed the mapping tasks in the work booklet. Before the participants started working on the test booklet, the other two booklets were retrieved. No feedback was provided to the participants on any of the tasks. The experimenters emphasized to the participants that they should complete the tasks in order and attempt to answer every question even if the questions were difficult. For each principle, the time allocated for the prerequisite learning, target learning, free-mapping task, and guided-mapping task was 5, 5, 10, and 15 min, respectively. This allocation was to ensure that the participants would spend a sufficient amount of time on each part of the experiment. In addition, 20 min were allocated to complete the test booklet. At the end of the experiment, each participant received a monetary reward of \$5 for their participation.

5.4. Analysis

The PROC MIXED of SAS was the primary procedure used for the analyses. Because the order variable was not a variable of interest for this study and did not affect the results, it was excluded from the analyses. The model for the free-mapping task included the following variables as fixed effects: principle, mapping, and relation, with principle and relation used as within-subject variables. For the other tasks, the relation variable was not included. The Kenward–Roger method was used to correct the estimates (Littell et al., 2006), and an unstructured option was chosen for the covariance structure. Separate tests were conducted for each principle based on the models. To test the effect of mapping type, the two PE conditions were compared with the three EE conditions. To determine the effect of example complexity, three tests were conducted. In the first, the two PE conditions were compared because the EE conditions were held constant (i.e., the complexity was higher in the PE pair). In the other two tests, the EsEs versus EsEc and the EsEs versus EcEc conditions were compared because the PE conditions were held constant (i.e., the complexity was higher in the EE pair). Type I error rates were controlled at 0.05 for each principle at the family level, and therefore, the alpha values were also 0.05 for the variable mapping type and example complexity. When analyses were conducted on the subcategories of responses (e.g., relational and nonrelational responses), the error rates were divided by the number of categories. Finally, the effect of principle complexity was analyzed by using the overall patterns of the results of the analyses.

To calculate effect sizes, two approaches were used. For the free-mapping task, because the numbers of responses provided by the participants were different (i.e., no fixed scale), the effect sizes and

their confidence intervals were estimated using the *bootES* package developed by Kirby and Gerlanc (2013). The bias-corrected-and-accelerated bootstrap method was used to estimate Hedges's *g* and its 95% confidence interval on the basis of 2,000 resamples. The estimates obtainable through these methods are similar to Cohen's *d* but have a higher accuracy (Kelley, 2005). For the guided-mapping task and the two tests, raw estimated differences (presented as raw *d*) between or among groups and their 95% confidence intervals were reported. These estimations were based on the analytical models and accounted for covariance among the variables. If a confidence interval contained zero, it is not reported for the sake of space.

6. Results and discussion

6.1. Free-mapping task

The responses to the questions were segmented and coded based on the ideas that were expressed regarding each principle. The score for each principle was calculated as the number of coded responses for the participants. A random sample of 25% of the participants was drawn ($n = 35$), and the responses of the sample were independently coded by two coders. For the simple principle, 161 pairs of codes were identified and simple Kappa = 0.87, ASE = 0.04, two-tailed $p < 0.001$. For the complex principle, 154 pairs of codes were identified and simple Kappa = 0.76, ASE = 0.04, two-tailed $p < 0.001$. For all data, differences in the codes assigned by the two coders were resolved through discussion or the decision of the researcher.

The coded responses were classified into the following categories: entity, property, property relation, first-order relation, higher-order relation, and other. The final category was a combination of several codes, namely no similarities or differences; do not know; blank; and responses that were uninformative, incomprehensible, or conceptually irrelevant. To analyze the mapping tendency, the first three categories were combined to form a nonrelational category, and the first-order and higher-order relation categories were combined to form a relational category. For each principle, two scores were calculated as the number of summated codes for these new categories.

The data as presented in Table 1 revealed that the category sums were similar between the two principles and the PE conditions generally had lower score sums than the EE conditions did on both principles. Additionally, more relational than nonrelational responses were provided under the PE conditions, whereas the opposite pattern was shown under the EE conditions, particularly when the complex principle was presented. For the simple principle, the planned test of the nonrelational responses revealed a mapping effect in which fewer responses were provided under the PE conditions than under the EE conditions, $t(133) = -3.34$, $SE = 0.22$, two-tailed $p = 0.001$, Hedges's $g = -0.58$, CI $[-0.85, -0.21]$. The test of the relational responses did not reveal a mapping effect, $t(133) = 0.65$, $SE = 0.21$, two-tailed $p > 0.10$, Hedges's $g = 0.11$, CI $[-0.21, 0.47]$. Furthermore, the tests of the example effect showed no significant differences (two-tailed $ps > 0.10$).

For the complex principle, the planned test of the mapping effect indicated that fewer nonrelational responses were provided under the PE conditions than under the EE conditions, $t(133) = -5.12$, $SE = 0.27$, two-tailed $p < 0.001$, Hedges's $g = -0.89$, CI $[-1.25, -0.51]$, and more relational responses were provided under the PE conditions than under the EE conditions, $t(133) = 2.98$, $SE = 0.21$, two-tailed $p < 0.01$,

TABLE 1 Mean number and standard deviation of responses by principle and experimental condition on the free-mapping task.

		Nonrelational		Relational		Sum		Total	
Mapping	n	M	SD	M	SD	M	SD	M	SD
Simple									
PEs	27	1.41	1.01	2.07	0.83	3.48	1.42	3.96	1.26
PEc	28	1.14	1.08	2.00	1.36	3.14	1.58	3.61	1.29
EsEs	28	2.14	1.72	2.14	1.38	4.29	1.90	5.43	1.97
EsEc	28	1.89	1.10	1.96	1.40	3.86	1.60	4.50	1.64
EcEc	27	1.96	1.19	1.59	0.97	3.56	1.37	4.48	1.53
Average	138	1.71	1.29	1.96	1.21	3.67	1.61	4.40	1.66
Complex									
PEs	27	1.19	1.55	2.11	1.28	3.30	1.32	3.52	1.37
PEc	28	1.39	1.50	1.82	1.49	3.21	1.32	3.50	1.32
EsEs	28	3.18	2.04	1.11	0.88	4.29	2.03	5.04	2.12
EsEc	28	2.64	1.50	1.29	1.15	3.93	1.18	4.57	1.35
EcEc	27	2.26	1.16	1.63	1.15	3.89	1.37	4.11	1.25
Average	138	2.14	1.73	1.59	1.24	3.72	1.51	4.15	1.61

PEs = principle-simple example; PEc = principle-complex example; EsEs = simple example-simple example; EsEc = simple example-complex example; EcEc = complex example-complex example. Sum: combination of nonrelational and relational responses. Total: including "others" responses.

Hedges's $g = 0.52$, CI [0.16, 0.89]. The other tests of the example effect did not produce any significant results (two-tailed $ps > 0.10$).

To summarize, the mapping propensities elicited by PE mapping and EE mapping differed. The learners' mapping behaviors were affected by the interaction of mapping type and principle complexity (Question 4). When a simple principle was being learned, PE mapping led to fewer nonrelational comparisons than EE mapping but the same number of relational comparisons. When a complex principle was being learned, PE mapping led to fewer nonrelational comparisons than EE mapping but more relational comparisons (Question 1). These findings verify the predictions regarding mapping type and principle complexity.

The results regarding example complexity showed that for both types of principles, variations in the example complexity had little or no influence on the participants' mapping tendencies (Question 3). These findings support the prediction regarding simple principles but not that regarding complex principles. This finding of a nonsignificant effect of example complexity on relational mapping during the free-mapping task might be explained by the multifarious behaviors of the participants. The participants may have avoided mapping content they were not confident about; that is, the participants assigned to the five conditions may have mapped content that did not require transformation and thus had similar performances. In addition, when the participants mapped transformational content, some merely indicated the conceptual categories the content belonged to. For example, a calculation was required to determine the frequency of an event in one worked example; however, some participants simply expressed that the two examples being compared mentioned "event frequency" or had different event frequencies. These responses were coded as nonrelational. Therefore, the effect of example complexity might have been partly diminished by how the participants articulated the mappings.

Learners' mapping behaviors are likely affected most by mapping type and principle complexity. Mapping two examples tends to lead to

more nonrelational comparisons regardless of whether the underlying principle of the examples is simple or complex. Mapping a principle onto an example of the principle tends to lead to more relational comparisons when the principle is more complex. When the principle is simple, the effects of PE and EE mapping on relational comparison are similar.

6.2. Guided-mapping task

The scores for the guided-mapping task were determined for each principle as follows. Each pair of responses constructed for PE mapping or EE mapping was determined to have no correct, a partially correct, or a completely correct correspondence. If a pair had a partially correct or completely correct correspondence, the correspondence was considered correct. Notably, the scoring was focused on the correct rather than incorrect mapping the participants completed. If a correct correspondence was similar to an example provided in the instructions, that response was not counted. In addition, if duplicate correct correspondences were identified because partially correct responses had been provided, only one correspondence was counted. For each principle, the maximum possible score per participant was 12. The alpha coefficient, calculated using the R package *psych* (Revelle, 2019), on the tetrachoric correlations among the 12 items was 0.80 for the simple principle and 0.83 for the complex principle.

As the mean total scores in Table 2 demonstrate, the two types of principles had similar mean scores, and the PE conditions on average had similar mean scores to the EE conditions for both principles. However, the PEc and EsEc conditions had slightly lower mean scores than the other conditions did for both principles. The analyses based on the total scores yielded the following results. Regarding the simple principle, the planned test of the mapping effect indicated similarity in performance for PE and EE mapping, $t(131) = -0.64$, $SE = 0.30$,

TABLE 2 Mean number and standard deviation of correct elements by principle and experimental condition on the guided-mapping task.

		Corresponding		Noncorresponding		Total	
Mapping	n	M	SD	M	SD	M	SD
Simple							
PEs	27	2.96	1.97	1.37	1.24	4.33	1.90
PEc	28	2.29	1.56	1.04	1.04	3.32	1.59
EsEs	28	3.50	1.75	0.89	1.10	4.39	1.87
EsEc	28	2.62	1.30	0.46	0.65	3.08 ^b	1.55
EcEc	27	3.85	1.75	0.70	0.91	4.56	1.60
Average	138	3.04	1.75	0.90	1.04	3.94	1.79
Complex							
PEs	27	3.33	1.82	0.44	0.85	3.78	2.04
PEc	28	2.56	1.67	1.07	1.30	3.63 ^a	2.17
EsEs	28	4.00	1.44	0.79	1.13	4.79	1.52
EsEc	28	2.44	1.12	0.33	0.55	2.78 ^a	1.25
EcEc	27	3.96	1.43	0.56	0.97	4.52	1.76
Average	138	3.26	1.63	0.64	1.02	3.90	1.89

PEs = principle-simple example; PEc = principle-complex example; EsEs = simple example-simple example; EsEc = simple example-complex example; EcEc = complex example-complex example.

^aOne participant with missing data was excluded.

^bTwo participants with missing data were excluded.

two-tailed $p > 0.10$, raw $d = -0.19$, CI $[-0.78, 0.40]$. The analyses of the example effect indicated that only the EsEs condition led to a better performance than the EsEc condition did, $t(131) = 2.76$, $SE = 0.47$, two-tailed $p < 0.01$, raw $d = 1.29$, CI $[0.37, 2.21]$.

With respect to the complex principle, the planned test did not reveal a mapping effect, $t(131) = -1.04$, $SE = 0.31$, two-tailed $p > 0.10$, raw $d = -0.33$, CI $[-0.94, 0.29]$. However, the tests revealed an example effect for the EsEs versus the EsEc condition, $t(131) = 4.19$, $SE = 0.48$, two-tailed $p < 0.001$, raw $d = 2.01$, CI $[1.06, 2.96]$.

Even though no mapping effect was discovered from the overall scores, the participants in the PE conditions likely had an advantage over those in the EE conditions for the noncorresponding content because the latter did not have direct access to it. For the complex principle, for example (see [Appendix A3](#)), the first condition presented in the PE pairs was not mentioned in the EE pairs. To investigate the possible advantage, the total score was divided into a score for corresponding content and a score for noncorresponding content. The score for the noncorresponding content was calculated as follows. Of the 16 components provided for each principle in the task, three for the simple principle and five for the complex principle did not have correspondences in the examples. If a correct response contained one or more of these components, it was counted as a case of noncorresponding mapping; other responses were counted as cases of corresponding mapping.

As presented in [Table 2](#), the responses were predominantly corresponding responses. The participants in the PE conditions tended to have lower performance than those in the EE conditions did on the corresponding content for the simple principle, $t(131) = -2.40$, $SE = 0.29$, two-tailed $p = 0.02$, raw $d = -0.71$, CI $[-1.29, -0.12]$, and for the complex principle, $t(131) = -1.97$, $SE = 0.27$, two-tailed $p = 0.051$, raw $d = -0.52$, CI $[-1.05, 0.001]$. However, on the noncorresponding content, the participants in the PE conditions had higher performance than the participants in the EE conditions did for the simple principle, $t(132) = 2.94$, $SE = 0.18$, two-tailed $p < 0.01$, raw $d = 0.52$, CI $[0.17, 0.87]$.

These findings indicated that PE mapping had an advantage over EE mapping regarding noncorresponding content even though the test was not significant for the complex principle. This advantage compensated for the lower performance on the corresponding content and thus the overall performance on the task was similar between PE and EE mapping.

The results for the overall task show that the difficulty of PE mapping and EE mapping did not differ (Question 2). Furthermore, this similarity was not affected by principle complexity (Questions 4). Additional analysis, however, revealed a differential effect in the difficulty of PE and EE mapping in relation to corresponding and noncorresponding content. The performances on the two kinds of content also uncovered a principle effect. The finding that, on average, the PE conditions had higher mean than the EE conditions on the noncorresponding content with the simple principle supported the conjecture that the principle in PE mapping serves as a model for learning. This modeling effect, nevertheless, was somehow weakened as the principle complexity increased. Because the content variable was not manipulated in this study, it is difficult to determine whether this observation was due to the increased abstractness effect of the principle. Further evidence regarding the modeling effect is presented in the subsequent discussion on the example effect.

The finding that the PE conditions had a lower mean than the EE conditions on the corresponding content may be explained by (1) PE mapping leading to more noncorresponding comparisons, (2) PE mapping being more difficult than EE mapping for this kind of content, or (3) both. Nevertheless, the results for the complexity principle, in which no mapping effect was found in the overall task and the noncorresponding content, suggested that PE mapping was more difficult than EE mapping in relation to the corresponding content. This evidence supports the prediction on a mapping effect under certain conditions. To gain a complete understanding of how the content variable influences the mapping effect, future research should systematically manipulate the variable.

TABLE 3 Mean number and standard deviation of correct responses by principle and experimental condition on the multiple-choice task.

Mapping	<i>n</i>	Simple		Complex	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
PEs	27	1.96	0.98	3.11	1.34
PEc	28	1.86	1.15	2.56 ^a	1.05
EsEs	28	1.46	0.88	2.89	0.74
EsEc	28	1.65 ^b	1.06	2.70 ^a	0.99
EcEc	27	1.89	1.05	2.48	0.98
Average	138	1.76	1.03	2.75	1.05

PEs = principle-simple example; PEc = principle-complex example; EsEs = simple example-simple example; EsEc = simple example-complex example; EcEc = complex example-complex example.

^aOne participant with missing data was excluded.

^bTwo participants with missing data were excluded.

TABLE 4 Mean number and standard deviation of correct responses with odd-response patterns removed by principle and experimental condition on the multiple-choice task.

Mapping	Simple			Complex		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
PEs	18	2.17	1.10	19	3.11	1.49
PEc	21	1.86	1.24	19	2.63	1.01
EsEs	11	1.64	1.21	17	2.65	0.93
EsEc	14	1.43	1.16	16	2.44	1.03
EcEc	14	1.93	1.21	16	3.00	0.73
Average	78	1.83	1.18	87	2.77	1.09

PEs = principle-simple example; PEc = principle-complex example; EsEs = simple example-simple example; EsEc = simple example-complex example; EcEc = complex example-complex example.

With respect to example complexity, for the overall task, EsEs mapping was easier than EsEc mapping was with both principles. In other words, PE mapping was less difficult than EE mapping when an EsEc pair was used in the latter (Question 3). These findings partially support the prediction that when a pair contain items of different complexities (i.e., PEc or EsEc), the difficulty of mapping would be higher. Notably, the effect of example complexity (PEs vs. PEc) was weak when the learners were required to employ PE mapping. This weaker effect may have occurred because of the modeling effect due to the presence of the principle. Although the example effect when EE mapping was used is in line with that reported in the literature (e.g., Reed and Bolstad, 1991), the effect when PE mapping was used requires further investigation to be characterized.

To recapitulate, PE mapping may not be more difficult than EE mapping. Factors such as content characteristics, example complexity, and principle complexity may influence the difficulty of mapping. Although PE mapping of corresponding elements is more difficult than EE mapping, such mapping may be as easy as or easier than EE mapping for noncorresponding elements. Example complexity may increase or decrease the difference in difficulty between these forms of mapping, depending on whether complexity is varied within the PE or EE pair. Finally, for more complex principles, the difference in the difficulty between these two forms of mapping may be smaller.

6.3. Multiple-choice task

For each principle, the participant score was calculated as the number of correctly answered questions. The mean scores are presented in Table 3. The tetrachoric correlations between the test items for each principle were used to calculate reliability. The resultant alpha coefficient was nearly zero for both principles. Analysis of variance components in which the participants and items were considered random factors and the principle was considered a fixed factor was performed. The results revealed that the Participant $\times$ Item interaction accounted for a large proportion of the variance in the data compared with the portion accounted for by the participant and item effects (0.215 vs. 0.003 and 0.033). This indicates that some participants had unconventional response patterns for the task that were not associated with the item difficulty. Rather, they indicate that some participants had learned some aspects of the principle more effectively than other aspects. Therefore, the correlation between any two items was low, which resulted in a low alpha coefficient.

To determine the effect of these unconventional responses, another analysis was conducted for the “reliable” participants. Information regarding how these participants were identified is provided in the Supplementary materials. The alpha coefficient for the participants with normal response patterns was 0.73 ($n=78$) for the simple principle and 0.65 ($n=87$) for the complex principle. The results indicate that the normal response patterns could play a key role in determining the reliability of task performance.

Although the unconventional participants led to the measures having lower reliability, their learning progresses were not invalid. For this reason, an analysis that included the scores of all participants was conducted. The data as presented in Table 3 indicated that the simple principle had a lower mean score than the complex principle did. This was likely caused by the participants misunderstanding the concepts of experiment and trial, with the detrimental effect of this misunderstanding being larger for the simple principle. The effects of mapping type and example complexity were negligible. Because the participants assigned to the five conditions had differing performance on the guided-mapping task, a proportion-correct score for the overall task was used as a covariate in the analysis. The planned test of the mapping effect uncovered similar performance for the participants in the PE and EE conditions for the simple principle, $t(131)=1.37$, $SE=0.18$, two-tailed $p>0.10$, raw $d=0.25$, CI $[-0.11, 0.60]$, as well as for the complex principle, $t(131)=0.82$, $SE=0.18$, two-tailed $p>0.10$, raw $d=0.15$, CI $[-0.21, 0.51]$. No example effect was identified for either principle (two-tailed $ps>0.05$).

The same set of analyses was next performed for the conventional samples to enable a comparison. See Table 4 for the means of these samples. Because some removed cases differed with the principles, separate analyses were performed for each principle. The results were similar to those obtained in the previous set of analyses; therefore, they are not presented here. Overall, the two sets of analyses demonstrated that although the participants handled the content differently, as indicated by their test item scores, their overall learning outcomes were not affected differently by the factors investigated in this study.

These analyses reveal that PE mapping and EE mapping resulted in similar increases in conceptual knowledge (Question 5). This similarity was not disrupted by example complexity (Question 7) or principle complexity (Question 8). A further discussion of these

findings is presented with that on structural knowledge in the next section.

6.4. Word-problem task

An overall task score and two principle scores were calculated for each participant based on the number of problems they correctly solved. Because the problems of the task involved three sets of principles and only two items were included for each set, an alpha coefficient was calculated for the overall task with the tetrachoric correlations. The reliability coefficient was 0.80 for the overall task.

As shown in Table 5, the overall mean scores were low, and they were close to the sum scores of the two principles, which indicates the participants rarely solved the filler problems. The data also reveal that the simple principle had a higher mean score than the complex principle did, and an example effect may have been present. For the analysis, the two principle scores and the covariate proportion-correct scores from the guided-mapping task were included. The planned tests did not reveal any effects of mapping type or example complexity (two-tailed p s > 0.10), with the exception of the PEs condition leading to a better performance than the PEc condition did for the complex principle, $t(131) = 2.95$, $SE = 0.16$, two-tailed $p < 0.01$, raw $d = 0.47$, $CI [0.16, 0.79]$.

Overall, the results indicate that the amount of structural knowledge the participants acquired through PE and EE mapping was similar (Question 6). This similarity, however, was affected by the interaction between example complexity (Question 7) principle complexity (Question 8). For learning the simple principle, example complexity did not affect the similarity. For learning the complex principle, PE mapping led to less gain than EE mapping when example complexity was increased in the PE pair.

The impact of mapping on learning could be attributed to the cumulative effects of the two mapping tasks. Since guided mapping was employed as the second task completed by the participants, it is likely to have influenced the test results. As demonstrated by the findings, participants in both the PE and EE conditions exhibited comparable performance on the guided-mapping task, regardless of the type of principles involved. This suggests that both forms of mapping resulted in similar learning outcomes when the assigned

mapping tasks were completed. Consequently, this observation may elucidate why there were analogous gains in conceptual and structural knowledge for both forms of mapping, irrespective of the principle types.

An effect of example complexity was observed for both types of principles when the participants in the EE mapping conditions completed the guided-mapping task. However, this effect was not observed for the two tests used to measure conceptual and structural knowledge. On the other hand, an example effect on structural knowledge gains was identified when PE mapping was used to learn the complex principle. These results showed that example complexity and principle complexity had limited effect on the learning of conceptual and structural knowledge. The reason might be the following. During the guided-mapping task, the participants were asked to construct the stems for mapping and identify the corresponding components from the target example. Such a task is likely more challenging than the free-mapping task. However, because the guided-mapping task did not include formal instruction, the learning gains from the task were limited. The participants' low to moderate performance on the two tests provides some evidence to support this speculation. Little example effect was discovered through the tests, which was likely because the participants assigned to the different conditions had not yet learned the content well enough that differential effects could be identified.

Although PE mapping and EE mapping may elicit different mapping behaviors with principles of varying complexity and differ in difficulty with content of different features, PE mapping can result in similar, if not higher, levels of structural and conceptual knowledge about the principle compared to EE mapping. In addition, these findings remain robust even when the principle complexity increases.

7. General discussion

The current study investigated the characteristics of connecting to-be-learned information (i.e., principles and examples) by mapping their entities, relationships, and conceptual structures. The findings contribute to the knowledge of analogical reasoning and learning by mapping one abstract principle onto one concrete example. Through this line of research, learning scientists can identify additional means of enabling learners to overcome the obstacles of comprehending abstract or general information and elucidate the interrelations among the individual concepts in formal structures. Although this line of research remains at an early stage, the current findings may provide some theoretical and instructional insight.

7.1. Principle–example mapping may be as effective as example–example mapping

The study findings seem to support that principle–example mapping is as effective as example–example mapping. In the guided-mapping task, the participants in the PE conditions compared fewer corresponding components but generally compared more noncorresponding components than did those in the EE conditions. Their higher number of comparisons of noncorresponding components compensated for the lower number of comparisons of corresponding components, which led the participants in the PE and

TABLE 5 Mean number and standard deviation of correct responses by principle and experimental condition on the word-problem task.

Mapping	<i>n</i>	Simple		Complex		Whole task	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
PEs	27	1.07	0.96	0.81	0.56	1.93	1.24
PEc	28	0.93	0.86	0.33 ^a	0.55	1.26 ^a	1.02
EsEs	28	1.29	0.81	0.68	0.55	2.07	1.30
EsEc	28	1.04 ^b	0.92	0.41 ^a	0.64	1.44 ^c	1.36
EcEc	27	1.19	0.88	0.48	0.70	1.67	1.33
Average	138	1.10	0.88	0.54	0.62	1.68	1.27

PEs = principle–simple example; PEc = principle–complex example; EsEs = simple example–simple example; EsEc = simple example–complex example; EcEc = complex example–complex example.

^aOne participant with missing data was excluded.

^bTwo participants with missing data were excluded.

^cThree participants with missing data were excluded.

EE conditions to have a similar overall performance. In addition, this similarity in their performance was not affected by principle complexity. These findings indicate that the two forms of mapping are generally equally effective for learning simple materials, such as those employed in this study. More importantly, the findings support the proposal that principle–example mapping can overcome the learning difficulty resulting from the abstractness of a principle.

Notably, in the guided-mapping task, the participants were able to freely determine what they should map. They are likely to have avoided mapping content that was difficult for them. If they had been directed to map all specified components, the findings may have been different. Therefore, future studies should use different guided-mapping tasks to test whether the conclusion regarding an equal effect is generalizable.

7.2. Principle–example mapping has a modeling effect

Several findings suggest that principles have a modeling effect in principle–example mapping. On the guided-mapping task, the participants in the EsEc condition had worse performance than those in the EsEs condition did. However, the participants in the PEc condition did not have worse performance than did those in the PEs condition. These findings were robust because they did not interact with principle complexity. Other findings revealed that the participants in the PE conditions had better performance than did those in the EE conditions with respect to mapping noncorresponding content. These findings imply that, when a learner attempts PE mapping, they can compare an example with the principle it illustrates and comprehend how the differences between the principle and example can be explained by the principle because the principle provides complete, nondistracting information.

With respect to how these findings can be applied in instruction, when a variety of worked examples or word problems are used in a lesson, they should be compared with the principle they illustrate. Merely mapping examples or problems of different levels of complexity may not result in the learner comprehensively understanding the underlying principle. This is particularly true when instruction time is limited or students are required to review examples or problems outside of class (e.g., Chi et al., 1989).

7.3. Principle–example mapping facilitates structural learning

Much of the evidence of this study support that principle–example mapping facilitates structural learning. In the free-mapping task, PE mapping of the complex principle led to the participants making more relational comparisons than EE mapping did, whereas EE mapping led to the participants making more nonrelational comparisons regardless the complexity of the principle. Although the relational advantage of PE mapping was not obtained with the simple principle, the participants in the hybrid mapping conditions were more likely to complete more relational than nonrelational comparisons, whereas those in the traditional mapping conditions were more likely to complete more nonrelational than relational comparisons, as indicated by the data presented in Table 1. These findings align with the

argument that principled information facilitates relational or structural learning (e.g., Goldstone and Son, 2005).

Other evidence supporting the facilitating effect of principle–example mapping come from the results of the conceptual and structural tests. In the literature, abstract information has been reported to be difficult to learn (e.g., Reed and Bolstad, 1991). However, in the current study, the two forms of mapping yielded comparable learning outcomes in conceptual and structural knowledge. This suggests that when a learner employs PE mapping, the difficulties associated with processing abstract information are compensated for by the utility afforded by a principle for processing relational information. Although EE mapping does not have this disadvantage, the resulting learning outcomes are not superior to those of PE mapping. This implies that the level of relational learning achieved may not be as high as that in PE mapping.

7.4. Instructional implications

The learning gains resulting from the participants completing the mapping tasks were somewhat limited. Therefore, some instructional intervention is required for the hybrid mapping approach to be sufficiently effective. In particular, rich and specific instructional support for novice learners are especially important (e.g., Bokosmaty et al., 2015). Research has provided suggestions regarding how mapping tasks can be effectively utilized. Firstly, transitioning from example–example mapping to principle–example mapping can be beneficial for understanding the principle. Some theorists have proposed (Halford, 1992; Rattermann and Gentner, 1998) that initial learning most likely begins with understanding individual entities and subsequently their relations with other entities or the relations between sets of local conceptual structures. Learning a principle by comparing two concrete examples and then comparing the principle with a concrete example may mimic this process. This approach can address the potential difficulty of principle–example mapping under certain conditions by guiding the learner to identify specific examples to which the principle may be applied. In addition, comparing such examples may enable the learner to identify key entities and their roles and properties (Alfieri et al., 2013). Through example–example mapping, the relationships among key entities may be noted, and through additional principle–example mapping, the properties and roles of the entities and the relationships among them can be explicitly understood. Therefore, even if a learner does not or cannot identify the general principle underlying a set of examples (e.g., Mayer, 2004), they can still understand the principle through principle–example mapping.

Secondly, transitioning from free mapping to guided mapping can be beneficial for motivation for learning and learning for understanding. In the present study, when the participants completed a free-mapping task, their responses were occasionally general. When their responses regarding the transformational content were general, determining whether they understood the exact relations among the transformational elements became difficult. Even when constraints are imposed, as they were in the present study's guided-mapping task, participants may avoid mapping difficult content, such that requiring inferences and transformations. Such avoidance could prevent learners from learning critical elements of a domain. Therefore, when designing instructional and learning tasks, the amount of guidance provided should be considered. For tasks aimed at enabling learners to explore the

structure of a domain, a free-mapping task would be more likely to encourage such exploration (e.g., Lammina and Chase, 2019). On the other hand, when the goal is to help learners understand the elements of a conceptual structure, a guided-mapping task with specific requirements may lead to more effective and efficient learning. In addition, instruction involving free mapping followed by guided mapping may encourage learners to engage in more active learning.

7.5. Limitations and future research

In this study, two factors regarding principles and examples were investigated, and therefore, the findings likely elucidate only some aspects of principle–example mapping. Future research can expand on the present study's findings by, for example, including principles from nonscience domains or examples that do not involve problem-solving. In addition, this study analyzed the characteristics and effects of principle–example mapping when no instructional intervention was provided. An analysis of the effects of mapping when an intervention is implemented would deepen the understanding of the effectiveness of the hybrid approach.

Future research should promptly address several issues of significance. One such issue is how principle–example mapping can be adapted for instructional purposes and whether the effects of such implementations can be generalized to the target populations of learners. Another is how the effect of principle–example mapping implemented as an instructional strategy differs from those of other forms of instruction, such as providing principled explanations or intermixing instructional explanations with worked examples. Research should also consider other issues that are more theoretical. For example, researchers should answer the following questions: Is there a difference between the cognitive processes evoked in principle–example mapping and in example–example mapping? How do these cognitive processes differ, and what do they imply? These research questions, albeit incomprehensive, should be investigated to achieve a complete understanding of the hybrid approach proposed in this study.

Data availability statement

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

References

- Alfieri, L., Nokes-Malach, T. J., and Schunn, C. D. (2013). Learning through case comparisons: a meta-analytic review. *Educ. Psychol.* 48, 87–113. doi: 10.1080/00461520.2013.775712
- Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 522–533. doi: 10.1037/0278-7393.16.3.522
- Bassok, M., and Olseth, K. L. (1995). Object-based representations: transfer between cases of continuous and discrete models of change. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 1522–1538. doi: 10.1037/0278-7393.21.6.1522
- Belenky, D. M., and Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: an organizing framework for categorizing external knowledge representations. *Educ. Psychol. Rev.* 26, 27–50. doi: 10.1007/s10648-014-9251-9
- Bokosmaty, S., Sweller, J., and Kalyuga, S. (2015). Learning geometry problem solving by studying worked examples: effects of learner guidance and expertise. *Am. Educ. Res. J.* 52, 307–333. doi: 10.3102/0002831214549450
- Catrambone, R., and Holyoak, K. J. (1990). Learning subgoals and methods for solving probability problems. *Mem. Cogn.* 18, 593–603. doi: 10.3758/BF03197102
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., and Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cogn. Sci.* 13, 145–182. doi: 10.1207/s15516709cog1302_1
- Cummins, D. D. (1992). Role of analogical reasoning in the induction of problem categories. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 1103–1124. doi: 10.1037/0278-7393.18.5.1103
- Day, S. B., Motz, B. A., and Goldstone, R. L. (2015). The cognitive costs of context: the effects of concreteness and immersiveness in instructional examples. *Front. Psychol.* 6:1876. doi: 10.3389/fpsyg.2015.01876
- Dixon, J. A. (2005). “Mathematical problem solving: the roles of exemplar, schema, and relational representations” in *Handbook of mathematical cognition*. ed. J. I. D. Campbell (New York, NY US: Psychology Press), 379–395.

Ethics statement

The studies involving humans were approved by Research Ethics Committee National Taiwan University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

HL contributed to conception of the study, research design, data collection, and analysis, and manuscript writing.

Funding

This research was supported in part by Tzu Chi University, Grant TCMRC-P-100012.

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.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1241873/full#supplementary-material>

- Dumas, D., Alexander, P. A., and Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: a systematic review of the literature. *Educ. Psychol. Rev.* 25, 391–427. doi: 10.1007/s10648-013-9224-4
- Fyfe, E. R., McNeil, N. M., Son, J. Y., and Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: a systematic review. *Educ. Psychol. Rev.* 26, 9–25. doi: 10.1007/s10648-014-9249-3
- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cogn. Sci.* 7, 155–170. doi: 10.1207/s15516709cog0702_3
- Gentner, D. (1989). “The mechanisms of analogical learning” in *Similarity and analogical reasoning*. eds. S. Vosniadou and A. Ortony (New York, NY: Cambridge University Press), 199–241.
- Gentner, D., Loewenstein, J., and Thompson, L. (2003). Learning and transfer: a general role for analogical encoding. *J. Educ. Psychol.* 95, 393–408. doi: 10.1037/0022-0663.95.2.393
- Gerjets, P., Scheiter, K., and Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learn. Instr.* 16, 104–121. doi: 10.1016/j.learninstruc.2006.02.007
- Gick, M. L., and Holyoak, K. J. (1980). Analogical problem solving. *Cogn. Psychol.* 12, 306–355. doi: 10.1016/0010-0285(80)90013-4
- Goldstone, R. L., and Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *J. Learn. Sci.* 14, 69–110. doi: 10.1207/s15327809jls1401_4
- Goldwater, M. B., and Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychol. Bull.* 142, 729–757. doi: 10.1037/bul0000043
- Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Hum. Dev.* 35, 193–217. doi: 10.1159/000277167
- Kelley, K. (2005). The effects of nonnormal distributions on confidence intervals around the standardized mean difference: bootstrap and parametric confidence intervals. *Educ. Psychol. Meas.* 65, 51–69. doi: 10.1177/0013164404264850
- Kirby, K. N., and Gerlanc, D. (2013). BootES: an R package for bootstrap confidence intervals on effect sizes. *Behav. Res. Methods* 45, 905–927. doi: 10.3758/s13428-013-0330-5
- Koedinger, K. R., Alibali, M. W., and Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: evidence from algebra problem solving. *Cogn. Sci.* 32, 366–397. doi: 10.1080/03640210701863933
- Lachner, A., and Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts’ explanations enhance novices’ deep-learning. *J. Exp. Psychol. Appl.* 21, 101–115. doi: 10.1037/xap0000038
- Lamnina, M., and Chase, C. C. (2019). Developing a thirst for knowledge: how uncertainty in the classroom influences curiosity, affect, learning, and transfer. *Contemp. Educ. Psychol.* 59:101785. doi: 10.1016/j.cedpsych.2019.101785
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberger, O. (2006). *SAS for mixed models*, 2nd ed. Cary, NC: SAS Institute Inc.
- Lopez, E. J., Shavelson, R. J., Nandagopal, K., Szu, E., and Penn, J. (2014). Ethnically diverse students’ knowledge structures in first-semester organic chemistry. *J. Res. Sci. Teach.* 51, 741–758. doi: 10.1002/tea.21160
- Markman, A. B., and Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Mem. Cogn.* 24, 235–249. doi: 10.3758/BF03200884
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *Am. Psychol.* 59, 14–19. doi: 10.1037/0003-066X.59.1.14
- Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., and Cox, G. (2012). Coordinating principles and examples through analogy and self-explanation. *Eur. J. Psychol. Educ.* 28, 1237–1263. doi: 10.1007/s10212-012-0164-z
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 510–520. doi: 10.1037/0278-7393.14.3.510
- Novick, L. R., and Holyoak, K. J. (1991). Mathematical problem solving by analogy. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 398–415. doi: 10.1037/0278-7393.17.3.398
- Rattermann, M. J., and Gentner, D. (1998). More evidence for a relational shift in the development of analogy: children’s performance on a causal-mapping task. *Cogn. Dev.* 13, 453–478. doi: 10.1016/S0885-2014(98)90003-x
- Reed, S. K. (1989). Constraints on the abstraction of solutions. *J. Educ. Psychol.* 81, 532–540. doi: 10.1037/0022-0663.81.4.532
- Reed, S. K., and Bolstad, C. A. (1991). Use of examples and procedures in problem solving. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 753–766. doi: 10.1037/0278-7393.17.4.753
- Reed, S. K., Stebick, S., Comey, B., and Carroll, D. (2012). Finding similarities and differences in the solutions of word problems. *J. Educ. Psychol.* 104, 636–646. doi: 10.1037/a0027181
- Renkl, A. (1997). Learning from worked-out examples: a study on individual differences. *Cogn. Sci.* 21, 1–29. doi: 10.1207/s15516709cog2101_1
- Revelle, W. R. (2019). *Psych: Procedures for personality and psychological research*.
- Richland, L. E., and McDonough, I. M. (2010). Learning by analogy: discriminating between potential analogs. *Contemp. Educ. Psychol.* 35, 28–43. doi: 10.1016/j.cedpsych.2009.09.001
- Richland, L. E., and Simms, N. (2015). Analogy, higher order thinking, and education. *Cogn. Sci.* 6, 177–192. doi: 10.1002/wcs.1336
- Rittle-Johnson, B., and Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: does one lead to the other? *J. Educ. Psychol.* 91, 175–189. doi: 10.1037/0022-0663.91.1.175
- Rittle-Johnson, B., and Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *J. Educ. Psychol.* 99, 561–574. doi: 10.1037/0022-0663.99.3.561
- Ross, B. H., and Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 42–55. doi: 10.1037/0278-7393.16.1.42
- Ross, B. H., and Kilbane, M. C. (1997). Effects of principle explanation and superficial similarity on analogical mapping in problem solving. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 427–440. doi: 10.1037/0278-7393.23.2.427
- Schwartz, D. L., Chase, C. C., Oppezzo, M. A., and Chin, D. B. (2011). Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *J. Educ. Psychol.* 103, 759–775. doi: 10.1037/a0025140
- Sloutsky, V. M., Kaminski, J. A., and Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychon. Bull. Rev.* 12, 508–513. doi: 10.3758/BF03193796



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Ana Marjanovic-Shane,
Dialogic Pedagogy Journal, United States
María Martínez López,
University of Castilla La Mancha, Spain

*CORRESPONDENCE

Angela Hadjipanteli
✉ hadjipanteli.a@unic.ac.cy

RECEIVED 13 April 2023

ACCEPTED 07 August 2023

PUBLISHED 29 August 2023

CITATION

Hadjipanteli A and Vrikki M (2023)
Conceptualising teacher dialogic artistry: a
nexus of pedagogical dispositions and
language qualities.
Front. Educ. 8:1205501.
doi: 10.3389/feduc.2023.1205501

COPYRIGHT

© 2023 Hadjipanteli and Vrikki. 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.

Conceptualising teacher dialogic artistry: a nexus of pedagogical dispositions and language qualities

Angela Hadjipanteli* and Maria Vrikki

Department of Education, School of Education, University of Nicosia, Nicosia, Cyprus

Teacher pedagogical artistry is fundamental to good teaching. This paper attempts to conceptualise one perspective of pedagogical artistry, that of *teacher dialogic artistry*. This conceptualisation is built on two dimensions of teachers' conduct aesthetically embodied in dialogic pedagogy: the pedagogical dispositions and language qualities. In this light, we suggest three dispositions that teachers need to develop a dialogic ecology in their class. These are togetherness, provisionality, and exploratory disposition. It is also essential that each of these dispositions be accompanied by a teacher's code of verbal communication thoroughly harmonised with them. In the context of such dialogic ecology, students are offered the potential to shape their dialogic ethos.

KEYWORDS

dialogic pedagogy, teacher dialogic artistry, togetherness, provisionality, exploratory disposition, dialogic ethos

1. Introduction

In modern educational theory, teacher pedagogical artistry tends to be a neglected notion and its merit as an inherent quality of good teaching (Eisner, 2004) is paradoxically oblivious. As Bellezza (2020) characteristically remarks, 'teacher artistry is grossly undervalued and greatly malnourished' (p. 27). This occurs within an era, where the teaching profession is primarily taught as 'a technological, knowledge-based practice (Hadjipanteli, 2021, p. 15), which means that the understanding of teaching as a poetic and art-full practice whose aesthetic energies have a critical effect on the learning ecology (e.g., Rubin, 1985; Eisner, 2004; Lutzker, 2021), seems languid.

When considering teaching as an art, what seems to be the most important in our opinion, is that a teacher works as an artist (Dewey, 1910), and pedagogy is intertwined with artistry. Therefore, a teacher-artist predominantly cares for the design and implementation of teaching 'within a purposeful, informed, ethical and aesthetic framework' (Ewing and Smith, 2001, p. 16). In this light, pedagogical artistry in teaching could go far beyond a teacher's knowledgeable presence, focusing also on the 'shape and feel of the relationships' (Thompson, 2015, p. 439) that students compose throughout their learning. This suggests that pedagogical artistry could bring the aesthetic custody and affective manipulation of teaching together, boosting the formation of an engaging and ethically well-structured learning ecology. For this purpose, teachers will need to energise a synthesis of poetical, intellectual and ethical dispositions to unify their pedagogical competence with artistic sensitivity.

In this article, we concentrate on a distinct outlook of a teacher's pedagogy, namely their competence to enact and encourage dialogic interactions with their students. In particular,

we are interested in the different qualities of dialogue that a teacher both could use and promote in class, as a catalyst for students' active participation in the construction of meaning and learning. We draw on the growing body of research on dialogic pedagogy, which shows how teachers can become facilitators in students' co-construction of knowledge and critical engagement with multiple ideas; we argue that this is an important aspect of a teacher's artistry and we specify it as *dialogic teacher artistry*. Specifically, we examine the following research questions:

- What pedagogical dispositions does the teacher embody within the process of dialogic pedagogy, which manifest their dialogic artistry?
- What language abilities and linguistic elements demonstrate dialogic teacher artistry?

We view this topic as particularly important and timely, given the growing diversity within student populations across western societies. The increased heterogeneity in classrooms, due to the students' wide range of cultural and linguistic backgrounds, calls for more inclusive and student-centred approaches to pedagogy that would take advantage of the students' rich backgrounds. Additionally, as explained in Section 3, we view dialogue as an instrumental tool that can be used for the teaching of all cognitive areas. This broad applicability justifies the need for teachers to develop what we call dialogic teacher artistry for the successful implementation of dialogic pedagogies.

2. Teacher pedagogical artistry: a key source of good teaching

That teaching is an art and the true teacher an artist is a familiar saying. Now the teacher's own claim to rank as an artist is measured by his ability to foster the attitude of the artist in those who study with him, whether they be youth or little children (Dewey, 1910, p. 220).

Identifying the influential 'ability' of a 'true teacher' that is essential to their motivational pedagogy, Dewey (1910) indicates pedagogical artistry. This ability originates from the Aristotelian concept *poetry*, which means *ποιεῖν* (making) and, therefore, accompanies the practice of *technê* (Aristotle, 1999). For Aristotle, *technê* is an intellectual virtue that helps us know 'how to bring into existence a thing which may either exist or not and the efficient cause of which lies in the maker and not in the thing made' (1999, 1140a12–1140a14). Following this conceptualisation, Sennett (2009) elaborates on *technê* with the notion of craftsmanship—'the skill of making things well' (p. 8)—which necessitates, as he expounds, all the mental and ethical power of one's self. However, if we ponder *technê* as equivalent to the meaning of art, in modern thinking, this tends to provoke misinterpretations, since *technê* is rather only associated with technology and technical skills. This interconnection signifies a separation of the rational from the expressiveness and aesthetics, aspects that in contemporary perception are meant to be at the heart of art (Ingold, 2000), which is typically seen as a creative activity.

Drawing upon the broader and more original sense of art as *technê*, we adopt both Eisner's (2002) thesis that artistry 'pertains to

the *crafting* of action' (p. 382, italics original), and Schön's (1987) conception that it stands for one's professional practice knowledge. Therefore, recognising teaching as a poetic and artful practice, this incites a critical consideration of a teacher's artistry (e.g., Eisner, 2002; Bellezza, 2020; Lutzker, 2021). As Hadjipanteli (2020) highlights, in teaching 'every pedagogical choice and action' is intertwined with 'a wide range of aesthetic acts: instructive, organisational, communicative and administrative' (p. 204), within which a teacher can infuse their pedagogical artistry. This circumstance ascribes to pedagogical artistry a complexity and multimodality (Eisner, 2002; Andrews, 2017), making it look like a kaleidoscopic phenomenon that can inspire the learning ecology within a class. Eisner (2002) enhances this view, by emphasising that '[g]ood teaching depends on artistry and aesthetic considerations, and that a teacher needs to know how to appropriately respond to 'all aspects of teaching that follow no rule', but 'they need to be felt' (p. 382). Here, Eisner reminds us that teaching is largely a social practice dependent on real-life aesthetics; a given condition that requires teachers' vigilance and pedagogical prudence in order that they sense, for example, the unsaid messages from students' learning and behavioural responses within an activity, or yet, to discern what emotional-ethical conditions influence their learning process at a specific moment.

In this line of thinking, we might acknowledge a few key characteristics of pedagogical artistry. Taking as granted that a teaching/learning ecology varies from class to class, potentially, owing to students' different profiles, idiosyncrasies and cultures, there is not solely a single approach or a fixed method for teachers to put in praxis their pedagogical artistry (Eisner, 2002; Piazzoli, 2018). In addition, their artistry can be embedded in the construction of a lesson in regard to students' concrete learning interests; the design of a curriculum unit in conjunction with a specific methodology suitable for students' educational needs; the selection and creation of teaching materials that address students' learning needs; the teacher's talk and motion; and moreover, the exploitation of classroom space. By these indicative paradigms, it becomes obvious that pedagogical artistry can function both as process and product (Sinclair and Kelman, 2013). As a process, it can shape the aesthetic elements of the form and content of an activity and as a product, it is perceptible within the live performance itself of teachers' instructive energies.

Although this multidimensionality of pedagogical artistry insinuates that each teacher could exercise it in various and alternative manners, what really undergoes in every case is that pedagogical artistry is performed with the synergy both of poetical and ethical dispositions (Eisner, 2002; Sennett, 2009; Andersen, 2020; Hadjipanteli, 2020). Rubin (1985) contends that a teacher's pedagogical artistry entails 'a personal style' of teaching energised by flair, originality, 'genuine commitment and an educated mind' (p. 165). Likewise, Eisner (2002) argues that this harmonised ethical-intellectual exertion required by a teacher exacts 'sensibility, imagination, technique, and the ability to make judgements about the feel and significance of the particular' (382). Creativity, playfulness, empathy and togetherness are some additional pedagogical dispositions of a teacher's artistry (Hadjipanteli, 2020). Carr's (2003) depiction of teacher-artist 'as a good jazz musician [...] whose techniques adapt and evolve in constant sensitive response to the needs of the musical moment', enable us to infer that teaching as art is a practice that comprises a 'constant creative interplay with the needs and challenges' (p. 27) of every pedagogical occasion.

In literature, the significance of a teacher's pedagogical artistry is intimately connected with the optimisation of students' vigorous learning, exciting their attentiveness, affective response and embodied participation (e.g., [Hadjipanteli, 2020](#); [Rawson, 2022](#)). From this perspective, a teacher's pedagogical artistry is transmuted into a vital source of students' 'eager and poetical engagement' ([Hadjipanteli, 2020](#), p. 204). In [Greene's \(1995\)](#), view such a result forms the cornerstone for significant learning, because it is driven by the desire of exploration and search. In this conceptual context, a teacher's pedagogical artistry denotes a critical formative factor of students' learning ecology; and concurrently, it constitutes a powerful and impactful property of a teacher's pedagogy that inextricably lies at the core of the notion of good teaching. Furthermore, it is critical to note that a teacher's pedagogical artistry could be considered, as entailed by the discussion above, a culturally driven ability. This affirms, on the one hand, the likely diversity and heterogeneity of teachers' pedagogical artistry in different cultures and suggests, on the other hand, that teachers need to be familiar with their students' culture, so that they can easier discern their learning preferences, predispositions and needs.

3. Teachers' dialogic dispositions: a pivotal characteristic of dialogic pedagogy

Dialogic approaches to teaching and learning have received increased attention in recent decades due to their association with student active participation, and improved understanding and achievement (e.g., [Howe et al., 2019](#)). The field is characterised by two main strands of dialogic pedagogies: (a) instrumental dialogic pedagogies, which view dialogue as a tool of teaching that is independent to the subject-matter or the participants, and (b) non-instrumental dialogic pedagogies, which view meaning making as inherently dialogic ([Matusov, 2018](#)). In this paper, we adopt the instrumental dialogic pedagogy approach because we conceptualise it as a tool that teachers can use in order to achieve learning goals of the curricula. This tool can be used in the teaching of all cognitive areas, including the teaching of language, citizenship education and the natural sciences. Following this, despite the existence of various approaches, and thus terms, to the study of productive forms of dialogue (i.e., 'dialogic teaching' by [Alexander, 2017](#); 'accountable talk' by [Michaels and O'Connor, 2015](#); 'dialogic inquiry' by [Wells, 1999](#)), there seems to be consensus on the characteristics of productive classroom dialogue ([Howe and Abedin, 2013](#); [Hennessy et al., 2016](#)). A dialogic lesson is one where students are encouraged to share their ideas, elaborate on them, and justify them. They have opportunities to listen to and engage with other ideas and perspectives, build on them or challenge them. Working in small groups is a fundamental setting in this environment because this provides students with more space to express themselves and work with a certain number of ideas collaboratively. In this environment, students actively develop their learning as they co-construct meaning with their peers.

Conducting a dialogic lesson requires a skillful teacher with dialogic dispositions, who can develop the required settings for productive dialogues to occur and guide students through this dialogic process. First, the teacher needs to build a suitable classroom culture that promotes students' *dialogic ethos* ([Alexander, 2017](#)). This entails

that students should feel safe knowing that they belong in a group where all ideas are respected and that all challenges and criticisms are towards ideas and not the person who is expressing them ([Littleton and Mercer, 2013](#)). Similarly, students should be trained for dialogic behaviour when working in groups; they should respect one another, make sure that all group members contribute to the discussion and that all ideas are being considered ([Littleton and Mercer, 2013](#)).

Second, teachers should not act as an authoritative figure in the dialogue, which seeks to evaluate students' knowledge through questioning, as the traditional teacher role would require. Instead, the teacher should know how to re-purpose the goal of classroom dialogue so that they become the ground for students to express their ideas, even their half-based ideas, and co-construct their thinking. In this setting, the teacher should act as a facilitator and guide of students' interactions ([Mortimer and Scott, 2003](#)). In particular, teachers are expected to initiate such interactions by asking open-ended questions (e.g., [Nystrand et al., 2003](#); [Wells and Arauz, 2006](#); [Vrikki and Evagorou, 2023](#)), modelling dialogic talk, and identifying differences between students' opinions so that they could be resolved ([Alexander, 2017](#); [Wilkinson et al., 2017](#)). The language that teachers use has a key role in achieving the kind of dialogue that is productive. In particular, teachers should use language of reasoning (e.g., 'because', 'so', 'but', 'if', 'how', and 'why'), possibility (e.g., 'would', 'maybe', 'could' and 'might') and vague language (e.g., 'like', 'kind of/sort of' and 'or something/or anything') in order to achieve the provisionality and possibility thinking that productive discussions require ([Maine and Čermáková, 2022](#)).

For these reasons, we conceptualise the teachers' skills and dispositions required for developing dialogic pedagogy to be part of a teacher's artistry; we call it *teacher dialogic artistry*. Building on these key presuppositions of dialogic pedagogy in combination with the theory of pedagogical artistry, we conceptualise teacher dialogic artistry based on a nexus of essential dispositions and language qualities.

4. Teacher dialogic artistry

4.1. Togetherness: a premise for students' democratic engagement

By its nature, dialogue is a communal practice. This becomes explicit by the Greek etymology of *διάλογος* (dialogue), which encompasses two concepts: *δια*—'gather together'—and *λόγος*—'the flow of meaning' ([Isaacs, 1999](#)). Following this interpretation, dialogic pedagogy premises a learning ecology of synergy, wherein teacher and students know to create a common interactive space for thinking and exchanging views and feelings; it is an ecology of an active learning community that stimulates students' eagerness for communicating their convictions and values. The practical fabrication of such an ecology indicates the original prospect of teachers' dialogic artistry. For the achievement of this purpose, teachers mainly need to interweave two conventions. First, the adoption of a student-centred approach ([Alexander, 2017](#)), where students are given the potential to 'take control of classroom dialogue' ([Vrikki et al., 2019](#), p. 482). Second, the formation of a relational culture that helps students manifest those 'positive-energy driven emotions' ([Hadjipanteli, 2020](#), p. 202), that can

constructively affect the dialogic process such as safety, trust, pleasure, curiosity and empathy, for instance.

In the context of both these conditions, the aesthetics of teachers' dialogic artistry presupposes the embodiment of a plexus of social-ethical dispositions and, also, the practice of a linguistic code of communication. Fundamental to the first presupposition is teachers' respect for students' views and beliefs in conjunction with the virtues of democracy. In drawing upon Neelands' (2009) notion of democracy the virtues that teachers' artistry will amalgamate in dialogic process is *isonomia* (the right of equality), *isegoria* (the right to speak), *isopsephia* (the right to equal representation), *parrhesia* (the ethical obligation to speak your mind) and *autonomia* (the right to self-determination). Teachers are called to give students an empirical and intimate acquaintance with all these democratic virtues within the practice of dialogue. In this attempt, teachers' dialogic approach itself needs to witness a serious and responsible harnessing of these democratic virtues.

Importantly, the familiarity of students with this democratic dialogic ethos is not irrespective of teachers' linguistic conduct; instead, this can trigger students' sense that dialogue is a shared and democratic activity in an ecology of togetherness. In this regard, teachers' linguistic communication may integrate vocabulary and phrases that make all students feel that they have their own personal voice, resulting from a collective and reciprocal dialogic process (see Figure 1). This comes in contrast with the traditional authoritative language used on the part of the teachers and shown in the indicative non-recommended examples in Figure 1. This kind of language tends to establish control over the classroom discourse and can possibly lead students' thinking towards a specific idea or conclusion (Aguiar et al., 2010).

4.2. Provisionality: a space for students' intrinsic understanding

In dialogic pedagogy, it is essential for students to be competent to articulate their own genuine dialogue between possibility and their thinking, thus experiencing a philosophical search. This necessity transforms dialogue into a space of investigation, where students undertake 'to pose questions, seek out explanations, to look for reasons, to construct meanings' (Greene, 1995, p. 26). For such a thoughtful dialogic process, teachers' dialogic artistry needs to bring together a sequence of intellectual dispositions like open-mindedness, social imagination and ethical imagination. These dispositions compose a teacher's provisionality in dialogic pedagogy.

Open-mindedness 'involves a willingness to form and revise one's view as impartially and as objectively as possible in the light of available evidence and argument' (Hare, 1981, p. 3). It proceeds synergistically with imagination, whilst the latter motivates the insight of possible, 'deals in unpredictabilities, in the unexpected' and 'requires reflectiveness' (Greene, 1995, pp. 124–125). Notably, social imagination, as defined by Greene (1995), is the 'capacity to invent visions of what should be and what might be in our deficient society, on the streets where we live, in our schools' (p. 5). Correspondingly, Moore (2011) describes ethical imagination as the unique human ability that permits the possibility of connecting to other people, by developing emotional bonds with them and respectfully confronting them.

The combination of these dispositions makes up a functional agent of teachers' dialogic artistry, as it strengthens students' cogitation and deliberative process, enabling them to shape meanings and interpretations. Therefore, teachers' provisionality can significantly contribute to the constitution of a dialogic culture, which impels students to widen their understanding of the world and life with different modes of looking at things. In this fashion, students can configure their belief-holding identity and self-knowledge, components that possess an underlying place in good teaching (Sokkett, 2012).

Practically speaking, what is indispensable in this dialogic process is the harmonisation of teachers' provisionality with a befitting usage of language from their side. This compatibility is manifested by suitable open-ended and dilemmatic questions that are implemented as 'belief questions' or 'self-revelation questions' (Belenky et al., 1986, p. 1). A few recommended examples of such questions are shown in Figure 1, which in their wholeness challenge students' reflection and prompt them to mould a clearer picture of their perceptions.

4.3. Exploratory disposition: openness for students' joint construction of knowledge

The third perspective of dialogic teacher artistry refers to its exploratory and enquiry-based nature. In the *Routledge International Handbook of Research on Dialogic Education*, Mercer et al. (2020) write that dialogic pedagogy 'implies that education should be designed to engage students in an ongoing process of shared enquiry that takes the form of dialogue' (p. 2). Alexander's (2017) notion of dialogic teaching implies that students should participate in a process of joint construction of knowledge. In this context, teachers' artistry can be revealed in the way that they deal with the multiplicity of ideas and thoughts being expressed. It is not enough to simply collect students' ideas; rather, they should transform the ideas into a joint inquiry by coordinating them in ways that allow students' exploration. One example of how teachers can coordinate multiple ideas appears in Figure 1 under 'Exploratory disposition', namely 'Could you identify common or contradictory elements between the ideas/opinions that have been heard so far?'. With such questions, teachers encourage students to actively engage with all ideas being heard and to begin identifying links between them. This can be the beginning of a co-construction process that helps students evaluate and refine ideas.

First, teachers can direct students' attention to ideas that are pivotal to the progression of the joint construction of understanding. For example, by intervening in students' exchanges with phrases like 'Have you heard Alex's thought?' or 'Alex has an objection here,' teachers can draw students' attention to new or critical perspectives that can help shape their exploration. Second, teachers can highlight links between ideas. For example, in the following quote from our database we can see how this teacher combines two ideas expressed by two different students in order to help the class to move their exploration forward. The quote is expressed by a primary school teacher, who discusses with her students a story about diversity.

Remember Alex what you said before about diversity? That it is also about us having different activities. I wonder. Can Alex's thought that everyone can have different activities, and Andrew's thought

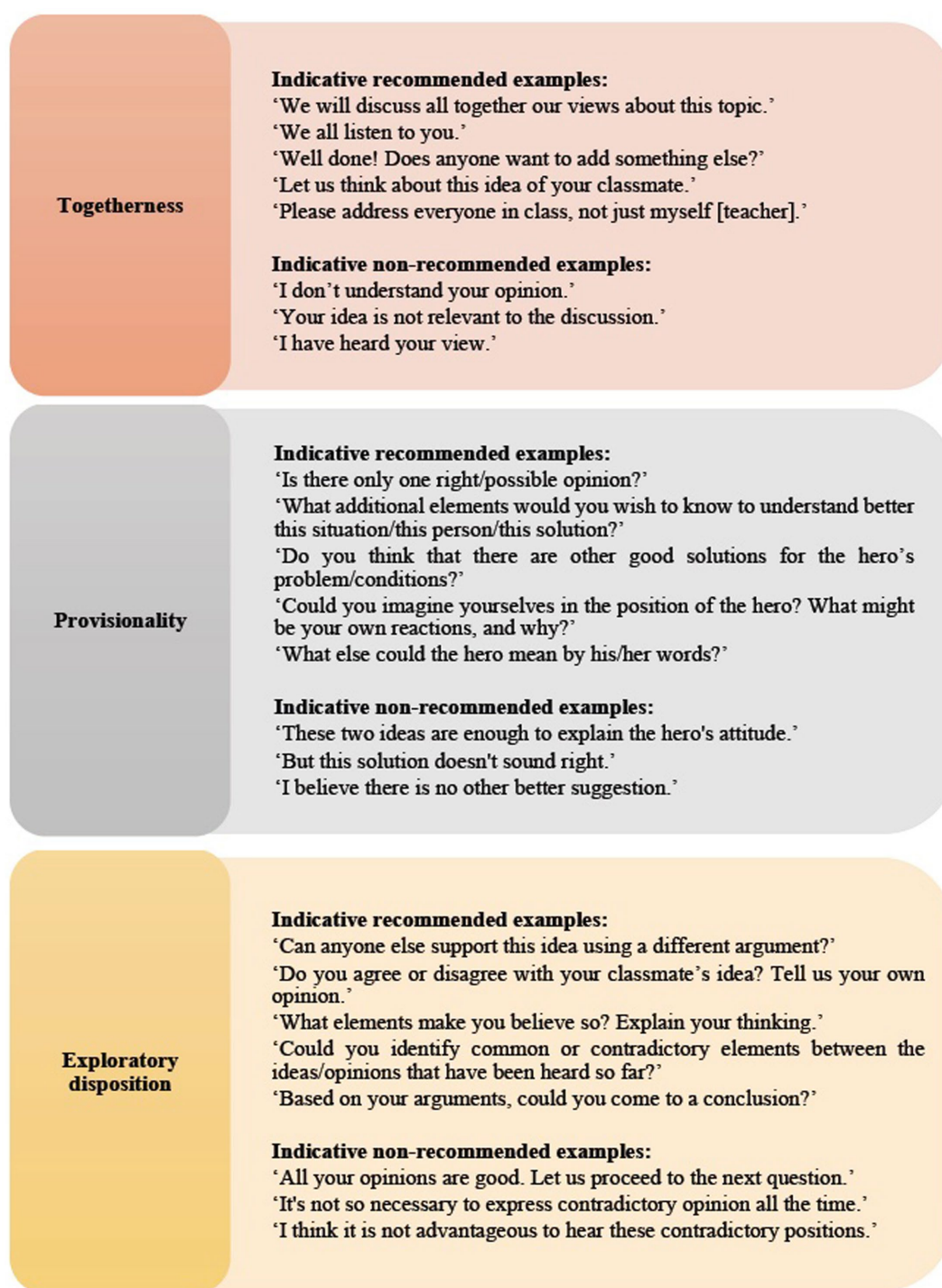


FIGURE 1
Recommended and non-recommended language examples of the three pedagogical dispositions of teacher dialogic artistry.

that the mouse's father didn't let him enjoy his hobby of dancing, be connected in some way?

Such techniques reveal teachers' artistry in facilitating and supporting students' exploration in classroom talk. Additional recommended examples, as well as non-recommended examples are presented in Figure 1.

5. Discussion

Teacher dialogic artistry, like dialogue *per se*, is a cognitive, poetical and normative concept. 'Genuine dialogue is extremely demanding', as Kazepides (2012) argues; 'it requires respect, trust, open-mindedness, a willingness to listen and to risk one's own preconceptions, fixed beliefs, biases and prejudices in the pursuit of

truth' (p. 915). In harmony with this spectrum of prerequisites, dialogic pedagogy is also challenging, as it requests from teachers to build a learning ecology infused by similar intellectual and ethical dispositions conducive to students' personal and spirited involvement in the dialogic process.

In this paper, we have discussed both from a theoretical and practical perspective, how teachers are expected to configure a dialogic ecology in their class concerted by their dialogic artistry. In this respect, we recommend dialogic pedagogy as a poetic and art-full practice guided by a teacher, who cares for helping students extend their existing 'dialogical nature of the self' as a 'human achievement' (Kazepides, 2012, p. 914). The three essential qualities of teacher dialogic artistry proposed could significantly affect students' dialogic performance, encouraging them to construct good relationships with the democratic, synergistic, insightful and investigative processes incurred in dialogue. Therefore, we might assume that a dialogic space built on a teacher's togetherness, provisionality and exploratory disposition communicated by their behavioural and linguistic actions can inspire students to form an aesthetics of good dialogue and an ethics of dialogic self.

Implications for teacher training and teacher professional development can be drawn. To cultivate dialogic artistry, teachers should develop relevant pedagogical dispositions and possess certain language qualities. Training programmes should aim to raise teachers' awareness of dispositions of togetherness, provisionality and exploration, and provide appropriate tools for translating these dispositions into practice. A code of verbal communication should accompany these dispositions to help teachers develop an appropriate dialogic ecology for students.

References

- Aguiar, O. G., Mortimer, E. F., and Scott, P. (2010). Learning from and responding to students' questions: the authoritative and dialogic tension. *J. Res. Sci. Teach.* 47, 174–103. doi: 10.1002/tea.20315
- Alexander, R. (2017). *Towards Dialogic Teaching: Rethinking Classroom Talk* (5th Edn.). Thirk, UK: Dialogos.
- Andersen, J. (2020). Understanding the artistry of actors who create theatre with and for children. *RiDE J. Appl. Theatre Perform.* 25, 484–497. doi: 10.1080/13569783.2020.1779586
- Andrews, K. (2017). "Artistry in teaching: a choreographic approach to studying the performative dimensions of teaching" in *Arts-Based Methods in Education Around the World*, eds. T. Chemi and X. Du (Denmark/The Netherlands: River Publishers), 15–27.
- Aristotle (1999). *Nicomachean Ethics*. Cambridge/London: Harvard University Press/William Heinemann.
- Belenky, M., Clinchy, B., Goldberger, N., and Tarule, J. (1986). *Women's Ways of Knowing: The Development of Self, Voice, and Mind*. New York: Basic Books.
- Bellezza, A. (2020). Developing performative competence and teacher artistry: a pedagogical imperative in the multicultural classroom. *L2 J.* 12, 23–42. doi: 10.5070/L20046068
- Carr, D. (2003). *Making Sense of Education: An Introduction to the Philosophy and Theory of Education and Teaching*. London: Routledge Falmer.
- Dewey, J. (1910). *How We Think*. D. C. United States: D. C. Heath and Co., Publishers.
- Eisner, E. W. (2002). From episteme to phronesis to artistry in the study and improvement of teaching. *Teach. Teach. Educ.* 18, 375–385. doi: 10.1016/S0742-051X(02)00004-5
- Eisner, E. W. (2004). *The Arts and the Creation of Mind*. New Heaven & London: Yale University Press.
- Ewing, R., and Smith, D. (2001). "Doing, knowing, being and becoming: the nature of professional practice" in *Professional Practice in Health, Education and the Creative Arts*, eds. J. Higgs and A. Titchen (Oxford, UK: Blackwell Science), 16–28.
- Greene, M. (1995). *Realising the Imagination*. New York: Jossey-Bass.
- Hadjipanteli, A. (2020). Drama pedagogy as aretaic pedagogy: the synergy of a teacher's embodiment of artistry. *RiDE J. Appl. Theatre Perform.* 25, 201–217. doi: 10.1080/13569783.2020.1730168
- Hadjipanteli, A. (2021). A study of a person-centred approach to teacher education within a drama course: its contribution to pre-service teacher's professional ethics. *Eur. J. Teach. Educ.* 1–19. doi: 10.1080/02619768.2021.1970741
- Hare, R. M. (1981). *Moral Thinking*. Oxford, UK: Clarendon Press.
- Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., et al. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learn. Cult. Soc. Interact.* 9, 16–44. doi: 10.1016/j.lcsi.2015.12.001
- Howe, C., and Abedin, M. (2013). Classroom dialogue: a systematic review across four decades of research. *Camb. J. Educ.* 43, 325–356. doi: 10.1080/0305764X.2013.786024
- Howe, C., Hennessy, S., Mercer, N., Vrikki, M., and Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: does it really impact on student outcomes? *J. Learn. Sci.* 28, 462–512. doi: 10.1080/10508406.2019.1573730
- Ingold, T. (2000). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. London and New York: Routledge.
- Isaacs, W. N. (1999). *Dialogue and the Art of Thinking Together*. New York: Anchor Books.
- Kazepides, T. (2012). Education as dialogue. *Educ. Philos. Theory* 44, 913–925. doi: 10.1111/j.1469-5812.2011.00762
- Littleton, K., and Mercer, N. (2013). *Interthinking: Putting Talk to Work*. London: Routledge.
- Lutsker, P. (2021). Teaching as an art in an age of science and technology. *Tsinghua J. Educ.* 37, 1–64. doi: 10.6869/THJER.202012_37(2).0001
- Maine, F., and Čermáková, A. (2022). Thinking aloud: the role of epistemic modality in reasoning in primary education classrooms. *Lang. Educ.* 37, 428–443. doi: 10.1080/09500782.2022.2129979

Data availability statement

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

Author contributions

AH and MV conceived the idea of teacher dialogic artistry, suggested the three dimensions of teacher dialogic artistry, and presented the discussion of the article. AH developed the theory of pedagogical artistry. MV discussed the key presuppositions for dialogic pedagogy. All 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.

- Matusov, E. (2018). "Mapping dialogic pedagogy: instrumental and non-instrumental education" in *The Cambridge Handbook of Sociocultural Psychology*. eds. A. Rosa and J. Valsiner. 2nd Edn. (Cambridge, UK: Cambridge University Press), 274–301.
- Mercer, N., Wegerif, R., and Major, L. (eds.). (2020). *The Routledge International Handbook of Research on Dialogue Education*. Abingdon, Oxon and New York, NY: Routledge.
- Michaels, S., and O'Connor, C. (2015). "Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion" in *Socializing Intelligence Through Talk and Dialogue*. eds. L. B. Resnick, C. Asterhan and S. N. Clarke (Washington DC: American Educational Research Association), 347–361.
- Moore, H. L. (2011). *Still Life: Hopes, Desires and Satisfaction (1st Edn.)*. Cambridge, UK: Polity.
- Mortimer, E. F., and Scott, P. H. (2003). *Meaning Making in Secondary Science Classrooms*. Maidenhead and Philadelphia: Open University Press.
- Neelands, J. (2009). Acting together: ensemble as a democratic process in art and life. *RiDE J. Appl. Theatre Perform.* 14, 173–189. doi: 10.1080/13569780902868713
- Nystrand, M., Wu, L., Gamoran, A., Zeiser, S., and Long, D. A. (2003). Questions in time: investigating the structure and dynamics of unfolding classroom discourse. *Discourse Process*. 35, 135–198. doi: 10.1207/S15326950DP3502_3
- Piazolli, E. (2018). *Embodying Language in Action: The Artistry of Process Drama in Second Language Education*. London: Palgrave Macmillan.
- Rawson, M. (2022). "Steiner Waldorf teacher education" in *Holistic Teacher Education: In Search of a Curriculum in Troubled Times*. ed. R. C. Collister (Newcastle, UK: Cambridge Scholars Publishing), 54–73.
- Rubin, L. J. (1985). *Artistry in Teaching*. New York, US: Random House.
- Schön, D. A. (1987). *Educating the Reflective Practitioner*. San Francisco, CA: Jossey-Bass Publishers.
- Sennett, R. (2009). *The Craftsman*. London: Penguin Books.
- Sinclair, C., and Kelman, D. (2013). "Drama, Cultural Leadership and Reflective Practice: Taking the Road to Zamunda" in *How Drama Activates Learning: Contemporary Research and Practice*. eds. M. Anderson and J. Dunn (London/New York: Bloomsbury Academic), 30–46.
- Sockett, H. (2012). *Knowledge and Virtue in Teaching and Learning: The Primacy of Dispositions*. New York: Routledge.
- Thompson, J. (2015). Towards an aesthetics of care. *RiDE J. Appl. Theatre Perform.* 20, 430–441. doi: 10.1080/13569783.2015.1068109
- Vrikki, M., Brindley, S., Abedin, M., and Riga, F. (2019). Exploring dialogic space: a case study of a religious education classroom. *Lang. Educ.* 33, 469–485. doi: 10.1080/09500782.2019.1604741
- Vrikki, M., and Evagorou, M. (2023). An analysis of teacher questioning practices during dialogic lessons. *Int. J. Educ. Res.* 117, 102–107. doi: 10.1016/j.ijer.2022.102107
- Wells, G. (1999). *Dialogic Inquiry: Towards a Sociocultural Practice and Theory of Education*. Cambridge, UK: Cambridge University Press.
- Wells, G., and Arauz, R. M. (2006). Dialogue in the classroom. *J. Learn. Sci.* 15, 379–428. doi: 10.1207/s15327809jls1503_3
- Wilkinson, I., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., et al. (2017). Toward a more dialogic pedagogy: changing teachers' beliefs and practices through professional development in language arts classrooms. *Lang. Educ.* 31, 65–82. doi: 10.1080/09500782.2016.1230129



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Juan Manuel Trujillo Torres,
University of Granada, Spain
Anouschka Van Leeuwen,
Utrecht University, Netherlands

*CORRESPONDENCE

Slavica Šimić Šašić
✉ ssimic@unizd.hr

RECEIVED 22 August 2023

ACCEPTED 11 September 2023

PUBLISHED 10 October 2023

CITATION

Šimić Šašić S, Nikčević-Milković A and
Cindrić M (2023) How much do teachers know
about self-regulated learning? To what degree
and in what way do they encourage it in
students?

Front. Educ. 8:1281438.

doi: 10.3389/feduc.2023.1281438

COPYRIGHT

© 2023 Šimić Šašić, Nikčević-Milković and
Cindrić. 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.

How much do teachers know about self-regulated learning? To what degree and in what way do they encourage it in students?

Slavica Šimić Šašić^{1*}, Anela Nikčević-Milković² and Maja Cindrić¹

¹Department of Teacher and Preschool Teacher Education, University of Zadar, Zadar, Croatia,

²Department of Teacher Studies in Gospić, University of Zadar, Gospić, Croatia

Introduction: Self-regulated learning (SRL) is a multidimensional process that involves personal (cognitive and emotional), behavioral, and contextual components. Teachers, as mediators in socialization, can influence SRL in various direct and indirect ways: by instructing students on effective learning strategies and structuring the learning environment. Most teachers agree that students should be helped to become self-regulated in their learning, however, they are unsure of how to do this, which is why they encourage it to a limited degree. Therefore, the objectives of the research were: (1) to examine how much teachers know about SRL, to what degree and how they encourage it in students; (2) to examine the differences in encouraging SRL with regard to gender, seniority, type of school and status of school subjects they teach.

Method: The research was conducted using an online questionnaire at two measurement points. A total of 251/179 primary and secondary school teachers in Croatia participated in the study. To assess teaching knowledge and methods of encouraging SRL, the questionnaire employed both open-ended questions ("Provide your own subjective definition of self-regulated learning."; "How you promote self-regulated learning in your practice? Which approaches are most effective?") and the Self-Regulated Learning (SRL) Encouragement Scale (encouragement of learning planning and learning organization strategies, metacognitive monitoring of learning, elaboration and evaluation, understanding and effort investment).

Results: Teachers only partially know what self-regulated learning constitutes. About two-thirds of teachers have heard of the SRL, but only 14% accurately defines the construct. When defining SRL, they most often emphasize independently regulated/directed learning, and the definitions are mostly partially correct. On the other hand, teachers estimate that they encourage SRL to a significant extent. There are certain discrepancies in the data regarding the methods of encouraging SRL depending on the methodology (qualitative/quantitative). Teachers encourage SRL to a limited extent and provide implicit and indirect SRL encouragement. The results indicate that female teachers and teachers in humanities encourage SRL to a greater extent when compared to natural science and technical subject teachers, and the same goes for elective subject teachers when compared to compulsory subject teachers.

Discussion: The results have theoretical, but also great practical implications when it comes to the implementation of this multidimensional and complex construct in the educational system. They corroborate the theoretical foundations of SRL and SRL encouragement, whilst indicating the need for improving teachers' knowledge and competencies in regard to SRL encouragement in students.

KEYWORDS

encouragement of self-regulated learning, primary and secondary school teachers, self-regulation of learning, teacher knowledge, differences in encouraging self-regulated learning

1. Introduction

Self-regulated learning (SRL) is a multidimensional process that involves personal (cognitive and emotional), behavioral, and contextual components. It is a process that integrates metacognitive aspects (planning, goal-setting, organization, self-monitoring, self-evaluation), cognitive aspects (selecting learning strategies, structuring the environment), as well as motivational aspects (self-efficacy, task interest, self-attribution), leading students to effectively regulate their own learning process (Zimmerman, 2002). According to Zimmerman (2002), SRL unfolds through three phases: during the preparation phase (prior to learning), the student analyses the task, sets goals, plans, selects learning strategies, and defines motivational beliefs. This includes assessing self-efficacy, setting expectations, goal orientation, and intrinsic interest. The performance phase (during learning) involves self-control and self-observation processes. The student focuses on the task, adjusts effort, monitors performance, analyses the conditions under which the performance occurs, and the consequences it entails. This involves applying different ways to enhance performance. The third phase, self-reflection (after learning), includes self-evaluation and self-directed action. The students assess their own achievement, reflect on the reasons behind that achievement, and identify the emotions triggered by that achievement (satisfaction/dissatisfaction) along with adaptive or defensive reactions (Zimmerman, 2002).

The terms SRL and SRL models primarily emerged within the socio-cognitive perspective, with Hadwin and Oshige (2011) asserting that the social context plays a crucial role in shaping and influencing SRL. Teachers, as intermediaries in socialization, can impact SRL in various direct and indirect ways: by teaching students effective learning strategies, or by structuring the learning environment to allow students to discover effective learning strategies on their own (Kistner et al., 2010; Dignath-van Ewijk and van der Werf, 2012; Karlen et al., 2020). Teachers may promote self-regulated learning through direct instruction of learning strategies, both implicit and explicit ones. Implicit instruction involves demonstrating specific behaviors, such as modelling strategy usage or verbalizing thought processes by the teacher. In this case, students are not explicitly informed that the behavior can be an effective learning strategy. This type of instructional strategy is referred to as blind training (Brown et al., 1981). Teachers can also engage their students in strategic behavior by asking questions. On the other hand, through direct instruction, teachers are able to explicitly demonstrate and tell students that a certain activity is a learning strategy that can enhance their performance. They can advise why it's important to use a particular strategy, how to apply it, when or in which situations it's appropriate, and what skills are necessary for its use. In that case, students receive some information about the meaning and importance of that strategy, and Brown et al. (1981) refer to this explicit strategy as informed training. Blind training can improve the use of a specific

strategy, but it does not generalize, whereas informed training enables the transfer of strategy application to the appropriate environment, which is particularly useful for less-able students (Veenman, 2007). The highest level of instruction is self-control training, combining guidance in strategy usage and informed training with explicit instructions on how to apply, monitor, check, and evaluate a specific strategy. This type of training facilitates the transfer of strategy application to relevant environments in the most sustainable way (Brown et al., 1981). Pintrich (2002) also emphasizes the need for explicit teaching of metacognitive knowledge. In most cases, teachers instruct metacognition in a rather implicit manner, assuming that students will autonomously acquire knowledge and skills. However, students need to be informed about the significance of strategies and how to use, monitor, and evaluate them. When it comes to encouraging SRL, another possibility comes indirectly through creating a supportive learning environment. The learning environment encompasses not only the characteristics of students and teachers but also the learning content, tasks, and teaching methods. An essential prerequisite for practicing self-regulation in classrooms is a learning environment that enables and encourages students to learn in ways they determine themselves (Kistner et al., 2010). In indirect instruction, the teacher creates a learning environment based on constructivist principles: presenting students with complex, authentic, and meaningful learning activities that promote specific subject knowledge and knowledge about SRL; granting students autonomy/choice in terms of what to learn, when, with whom, and for how long; facilitating self-regulated and meaningful learning directed towards clear objectives; adapting support and feedback to individual student needs in challenging situations; implementing forms of assessment and self-assessment in service of learning enhancement and monitoring, as well as actively involving students in evaluating their own learning; supporting positive self-beliefs related to learning and problem-solving (Karlen et al., 2020). De Corte et al. (2004) emphasize several main principles in their CLIA model (Competence, Learning, Intervention, Assessment): social interaction among students (collaboration), active knowledge construction (constructivism), learning embedded in authentic situations to encourage transfer (situation), and the development of self-regulation skills (self-guidance). Various studies empirically verify the positive effects of strong CLIA-based learning environments, such as improved student problem-solving competencies (Verschaffel et al., 1999), increased self-regulation activity in students, and improved academic achievement (Masui and De Corte, 2005). Vandeveldt et al. (2012) highlight modelling, prompting and scaffolding as key tools for fostering learning self-regulation, creating a supportive environment, and explicitly teaching effective self-regulation strategies, while also transitioning from maladaptive to adaptive strategies as needed. In conclusion, we can deduce that there are various approaches to fostering learning self-regulation: (a) direct instruction, modelling, verbalization of thinking, which encourage cognitive and

metacognitive aspects of SRL, such as learning strategies, planning, progress monitoring, comprehension, evaluation, etc.; (b) cognitive-behavioral modification, promoting attention direction and retention, modifying student beliefs and goals, dealing with negative emotions; (c) classroom environment modification, through task types, authority, rewards, student grouping methods (collaborative learning), evaluation methods, altering interpersonal relationships to create a more positive classroom/school atmosphere, etc. [Perry et al. \(2006\)](#) state that students develop effective forms of SRL when engaged in complex, meaningful tasks, when they have a say in their learning, products, and evaluation criteria, when working in groups and seeking feedback from peers, and when evaluating their own learning. Combining direct and indirect approaches yields the best results ([Paris and Paris, 2001](#)). In sum, the listed means of promoting SRL show that teachers are able to encourage all components of SRL in all phases of SRL. Furthermore, in the process of self-regulated learning, students seek the support of teachers, value their feedback, and critically reflect on the teaching strategies used ([Chaves et al., 2015, 2016](#)).

Teacher knowledge is typically classified into three categories: content knowledge, pedagogical knowledge, and pedagogical content knowledge. [Askill-Williams et al. \(2012\)](#) suggest that teachers should also possess content and pedagogical content knowledge about fostering SRL, including cognitive and metacognitive learning strategies. [Karlen et al. \(2020\)](#) discuss teacher professional competencies in SRL, which encompass teachers' personal competencies in SRL and experience as a self-regulated learner (teacher as a self-regulated learner), as well as competencies in teaching, diagnosing, and supporting SRL in the classroom (teacher as an agent of SRL). Teacher professional competencies in the field of SRL are related to teaching practices and the development of SRL in students ([Wilson and Bai, 2010; Moos and Ringdal, 2012; Spruce and Bol, 2015](#)). Studies have shown that teachers rarely integrate SRL into their everyday classroom teaching, that they tend to provide limited direct instructions on strategies, and seldom emphasize metacognitive aspects of SRL ([Dignath-van Ewijk and van der Werf, 2012; Vandevelde et al., 2012; Kistner et al., 2015; Spruce and Bol, 2015; Dignath and Büttner, 2018; Karlen et al., 2020](#)). Moreover, targeted SRL training occurs very infrequently ([Hamman et al., 2000](#)). While most teachers agree that students need help to develop self-regulated learning, they are unsure about how to achieve this ([Perry et al., 2008](#)). [Vandevelde et al. \(2012\)](#) found that teachers stimulate SRL only to a limited extent. Research into the connection between teacher beliefs, knowledge, and teaching practices indicates that teachers express positive beliefs about SRL, but their knowledge about SRL and its implementation in the classroom is generally weak ([Spruce and Bol, 2015; Karlen et al., 2020](#)). [Dignath-van Ewijk and van der Werf \(2012\)](#) found that teachers perceive and understand concepts of SRL and "learning to learn" as separate concepts, i.e., they associate SRL with creating a supportive environment and "learning to learn" with teaching learning strategies, while also exhibiting a more positive attitude toward constructivism compared to SRL. Teachers typically grant students freedom in self-regulation but fail to prepare them to handle new challenges ([Bolhuis and Voeten, 2001](#)). They prefer to use active teaching methods but do not instruct students in how to learn ([de Kock et al., 2005](#)). Similarly, [Dignath-van Ewijk and van der Werf \(2012\)](#) investigated teachers' knowledge about SRL and fostering SRL, finding that most teachers emphasize student autonomy and a constructivist learning environment, while only a smaller portion focus on teaching learning strategies. [Wilson and Bai \(2010\)](#) identified a weak correlation

between teachers' knowledge about SRL and their teaching practices when it comes to encouraging SRL. Even teachers with greater knowledge did not consistently exhibit a high level of teaching that supports SRL strategies. Therefore, it is extremely important to promote the development of teacher competencies in the field of SRL so that teachers can effectively encourage it, especially through informed instruction and self-control training for students.

Encouraging SRL depends on teacher characteristics, such as gender, age, teaching experience, teacher beliefs, teacher competencies for their own self-regulation, etc. ([Hargraves, 2005](#)). However, research results on this matter are not consistent. [Lombaerts et al. \(2009\)](#) found that the gender of elementary school teachers does not significantly impact the degree to which SRL is encouraged, while some other authors discovered that female teachers tend to use approaches that promote SRL more ([Elmas et al., 2011; Yan, 2018](#)). [Peeters et al. \(2015\)](#) determined that older teachers are less likely to support SRL. [Lombaerts et al. \(2007\)](#) established that teaching experience influences the teaching of metacognition in the classroom, while [Yan \(2018\)](#) found that teaching experience is not a significant predictor of SRL encouragement practices in instruction. Yan also noted that elementary school teachers perceive more benefits from SRL than high school teachers. In general, the research confirms that promoting SRL tends to be more beneficial for elementary school students compared to high school students ([De Smul et al., 2018](#)). Teachers in middle school and high school create opportunities for learning self-regulation but rarely engage in direct instruction on learning strategies. Teachers in lower grades of elementary school (up to 6th grade) predominantly encourage learning self-regulation ([Moos and Ringdal, 2012](#)). Teachers who possess a strong understanding and value the content of their school subject are more likely to use advanced, student-cantered teaching methods ([Baumert and Kunter, 2013](#)). [Fauzi and Widjanti \(2018\)](#) discovered that mathematics teachers more frequently use teaching techniques that promote SRL compared to teachers of other subjects. They also found that mathematics teachers provide more opportunities for students to solve problems using SRL strategies. [Chatzistamatiou et al. \(2013\)](#) emphasized that mathematics teachers employ teaching strategies that facilitate independent learning and creative problem-solving. They also found that, compared to less successful mathematics teachers, the best ones more often utilize techniques that encourage SRL. However, [Coggin \(2020\)](#) did not find differences in teaching SRL among teachers of different subjects (languages, mathematics, natural sciences, social sciences, and others/multiple subjects). Therefore, we wanted to examine how much teachers in Croatia know about SRL and how they encourage it in students. Considering the inconsistent results, we also sought to examine the differences in promoting SRL based on gender, teaching experience, type of school, and school subject status (science field; compulsory/elective). We assume that the knowledge about SRL will be weak, that teachers will rarely promote SRL, especially directly, and we do not expect differences based on gender, teaching experience, type of school, and school subject status.

2. Method

2.1. Participants

The research was conducted as part of a wider research project titled "The Role of Teachers in Self-Regulated Learning

TABLE 1 Sample characteristics with respect to gender, teaching experience, type of school, science field of the school subject, and subject's status.

	%
Gender	
Women	80.88
Men	19.12
Years of teaching experience	
0–7	22.71
8–15	38.65
16–23	24.30
Over 24	14.34
Type of school	
Elementary	47.41
Grammar (high) school	21.12
Vocational (high) school	31.47
Subject's science field	
Humanities	46.61
Social	9.96
Natural	27.89
Technical	15.54
Subject's status	
Compulsory	90.44
Elective	9.56

Encouragement,” and the data were collected across two measurement points. In the first measurement point, 251 and in the second, 179 primary and secondary school teachers from 17 counties of the Republic of Croatia participated. [Table 1](#) shows the sample structure in the first measurement point (the sample structure in the second measurement point was similar). The average age of the teachers was 42 years.

2.2. Measurement instrument

General Information and SRL Knowledge Questionnaire – used to collect data pertaining to gender, teaching experience, type of school where teachers work, field, and the status of the subject they teach. In the first measurement point, we also asked teachers if they had heard of the concept of self-regulated learning (YES/NO) and to provide their own subjective definition of self-regulated learning. In the second measurement point, among other things, we asked teachers how they promote self-regulated learning in their practice? Which approaches are most effective? (open-ended questions). Definitions of SRL were evaluated as accurate if they emphasized the majority of SRL elements (e.g., “...an active process in which students set learning goals, use various learning strategies, monitor their progress, and adapt to their environment; “Planning and setting learning goals while consistently monitoring and evaluating what has been learned”), partially accurate if they highlighted several SRL elements (e.g., “Students independently direct their learning and develop learning skills”; “They plan study time, break

learning content into smaller units, schedule time for review”), and inaccurate (e.g., “You regulate learning on your own, learn as much as you want”; “Students learn... what they themselves or the teacher think needs to be learned and known”; “Process of directing students towards independent learning”). Teacher responses on the means of SRL encouragement were coded based on the processes tied to individual SRL components and phases (see the categories in [Table 2](#)). Two assessors separately analyzed the teachers' responses to open-ended questions, with consensus being ultimately achieved.

Self-Regulated Learning Encouragement Scale ([Šimić Šašić et al., in press](#)) – It was constructed based on the existing models of self-regulated learning and the literature on self-regulated learning encouragement, ultimately incorporating 51 statements in total. The analysis revealed a five-factor structure, meaning that the scale measures teacher encouragement of self-regulated learning in the following five areas:

- Encouragement of Learning Planning and Learning Organization Strategies (ELPLO) – measures the degree to which the planning of learning is encouraged (setting learning goals, allocating time, using learning strategies, organizing the learning environment, encouraging task assessment, identifying causes of success/failure, etc.) and the encouragement of learning organization strategies (breaking content into smaller meaningful units, identifying key concepts, summarizing, asking questions). One example is the following statement: “I ask students to create a study plan or set goals for their learning.”
- Encouragement of Metacognitive Monitoring of Learning (EMML) – the statements pertain to encouraging students to maintain and direct their attention during learning, try various learning/problem-solving approaches, and engage in activities when motivation wanes or negative emotions arise. This includes supporting interest in learning, investing extra effort, reminding students of task value, etc. One such example is the following statement: “I encourage students to direct and maintain their attention on the content they are learning.”
- Encouragement of Elaboration and Evaluation (EEE) – measures the encouragement of explanation/discussion among students, graphical representation of information, application of knowledge/creation, assigning less structured tasks, independent task solving, working in pairs/small groups, and involving students in setting evaluation criteria, self-evaluation, and evaluating the work of others. One such example is the following statement: “I ask students to explain the content they are learning to each other or to discuss the content amongst themselves.”
- Encouragement of Understanding (EU) – the statements pertain to activating previously acquired knowledge, initiating teaching with intriguing tasks, encouraging drawing conclusions, connecting information from various sources, correcting misunderstandings, and linking information to everyday life situations. One such example is the following statement: “I give students enough time to explore and gain understanding of new concepts/content.”
- Encouragement of Effort Investment (EEI) – the statements relate to encouraging students that they can accomplish tasks, motivating effort investment, attributing success to effort, and emphasizing the value of knowledge. One such example is the following statement: “I encourage students suggesting they can learn/accomplish a task.”

TABLE 2 Analysis of teachers' responses concerning the SRL definition (N = 251).

Areas emphasized in their definitions	Examples of their claims	f
SRL	<i>...an active process, where students set goals, utilize various learning strategies, monitor their progress, and adapt to their environment...</i>	36
Learning autonomy	<i>students learn what they themselves or the teacher believe needs to be learned... you learn as much as you want... regular and independent learning... with the application of their own learning methods and self-discipline...</i>	48
Learning organization	<i>the student organizes their own learning... time, responsibilities</i>	24
Planning (of learning, time)	<i>...students independently plan their learning... study time... learning methods...</i>	35
Goal-setting	<i>setting personal goals... the student determines the goal...</i>	13
Intrinsic motivation	<i>students self-motivate... their motivation is to learn well... they learn without external motivation...</i>	25
Learning strategies	<i>students independently find the best methods to learn...</i>	46
Learning supervision	<i>students self-direct their learning... self-regulate/manage learning... monitor progress... adjust learning methods...</i>	93
Active learning	<i>active learning... students active in the learning process...</i>	7
Efficient learning	<i>students independently and effectively organize their learning and achieve success... with the goal of personal development...</i>	
Evaluation	<i>students value progress... learning... self-evaluation</i>	5
Effort investment	<i>students are willing to invest extra effort...</i>	1
Responsibility for one's own progress	<i>...students take responsibility for their learning...</i>	4
Additional learning	<i>...students expand their knowledge...</i>	1
Behaviors directed to goals	<i>...behaviors that achieve the goal...</i>	1
Seeking help	<i>...ask for help when something is not clear...</i>	2
Positive motivation and emotions	<i>...have positive motivational beliefs and emotions...</i>	1
Adjusted learning	<i>...adjusted to the students' abilities, motivation, interests, and capabilities... students' needs...</i>	2
External regulation	<i>...we guide students towards independent learning...</i>	1
Long-term learning	<i>...students have the goal of attaining knowledge on something permanently...</i>	1
Do not know		3
No response		4

Teachers were asked to rate their agreement with these statements on a scale from 1 (strongly disagree) to 5 (strongly agree), and the subscales showed satisfactory reliability coefficients with Cronbach's alpha as follows (respectively): 0.90, 0.89, 0.83, 0.85, and 0.71.

2.3. Procedure

The research was conducted in the form of an online questionnaire. Approval for conducting the research was obtained from [Ministry of Science and Education of the Republic of Croatia \(2019\)](#), the Ethics Committee of the University of Zadar, as well as school principals. Teachers were informed about the research's purpose, objectives, and methodology, and their participation was voluntary and anonymous. The questionnaire link was sent to school coordinators who then forwarded it to the teachers. Filling out the questionnaire took approximately 15 min.

2.4. Data processing

Considering that research tends to report different results in relation to teachers' knowledge and beliefs about SRL when using

qualitative and quantitative methods ([Dignath-van Ewijk and van der Werf, 2012](#); [Šimić Šašić et al., 2021](#)), in accordance with the advantages and disadvantages of each, we measured the teachers' knowledge of SRL and their encouragement of SRL using both open-ended questions and the SRL Encouragement Scale. We analyzed the responses to open-ended questions qualitatively, conducted content analysis, and presented response categories and frequencies. Quantitative data were processed using descriptive and inferential statistics (t-test, analysis of variance, Bonferroni test for *post hoc* analyses).

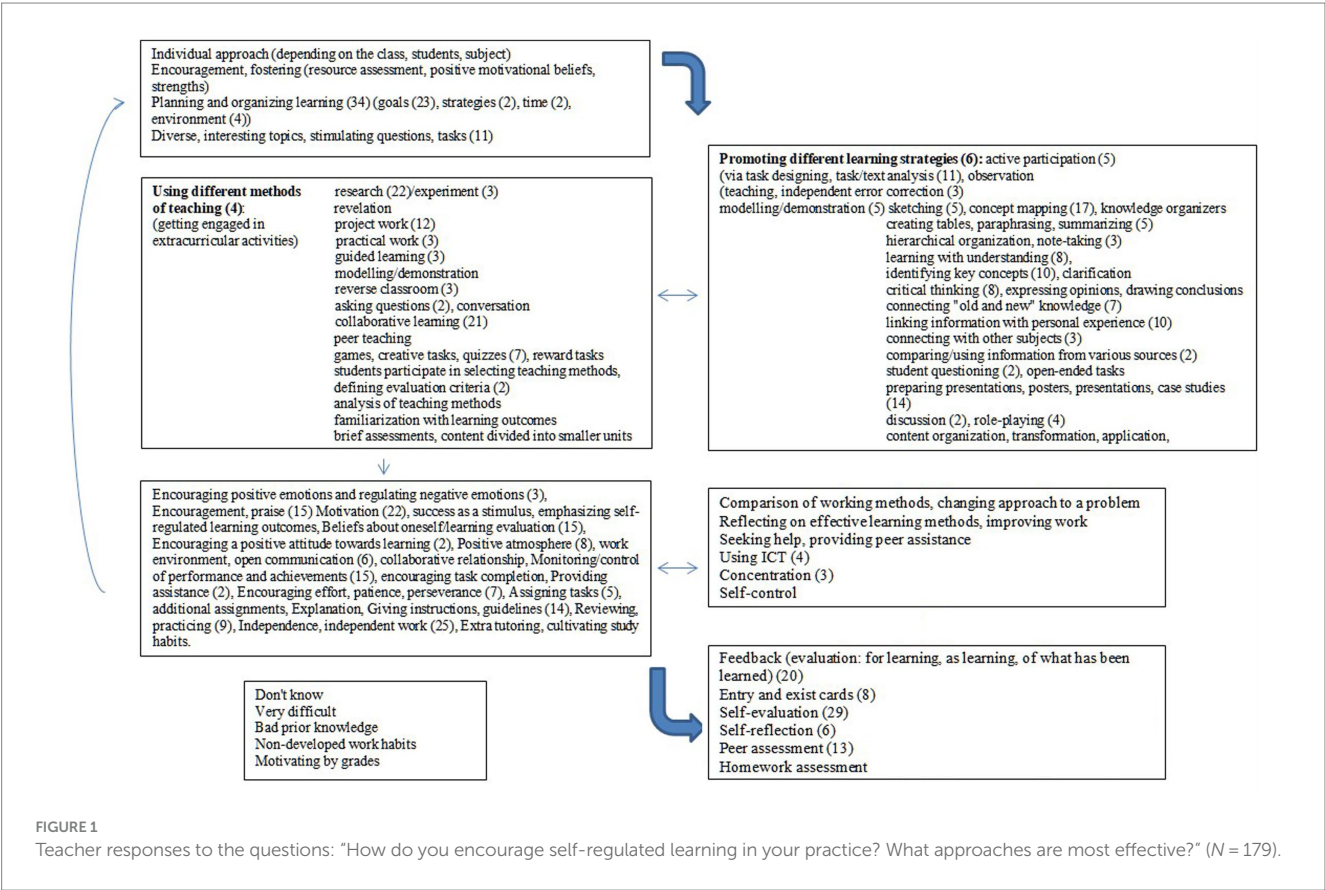
3. Results

Table 3 presents the teacher responses (%) to the question of whether they have heard of self-regulated learning, along with the analysis of teachers' subjective definitions of SRL. Table 2 provides an analysis of teacher response categories when it comes to defining SRL.

One-third of the teachers stated that they have not heard of SRL. Based on the data and the number of teachers who did not respond or wrote that they did not know, it was found that 21.12% of teachers are not familiar with self-regulated learning. In their

TABLE 3 Analysis of familiarity and accuracy teachers' definitions of SRL.

Have you ever heard of self-regulated learning	f	%	Definition	f	%
Yes	169	67.60	Accurate definition	36	14.34
No	81	32.40	Partially accurate definition	162	64.54
			Inaccurate definition	46	18.33
			No response/do not know	7	2.79



definitions of learning, teachers predominantly emphasized metacognitive monitoring/self-direction, independence, use of learning strategies, and learning planning.

In similar fashion, we analyzed the teachers' responses to the questions in the second measurement point: "How do you encourage self-regulated learning in your practice? What approaches are most effective?." The results of the analysis are presented in Figure 1.

The teachers' responses could have been grouped according to the components of SRL (metacognitive, cognitive, motivational, and contextual), the phases of SRL (preparation, performance, and self-reflection), and the methods of encouraging SRL (direct/indirect). Teachers most commonly referred to encouraging planning and organization of learning, self-evaluation, autonomy, goal-setting, inquiry-based learning, motivating, collaborative learning, providing feedback, creating mind maps, promoting monitoring/performance control, beliefs about the actor, and preparing presentations (Figure 1).

Teachers provided relatively high to very high assessments when it comes to them encouraging SRL. Their estimates suggested that

they mostly encouraged effort investment, followed by metacognition and understanding, while having least encouraged learning planning and organizational strategies, as well as elaboration and learning evaluation (Table 4).

The Levene's test showed that all variables have homogeneous variance except for Encouragement of Understanding (which is why a non-parametric test was additionally conducted ultimately yielding the same results; for consistency in presentation, see the included the parametric test). Statistically significant differences were found in terms of gender in all components of promoting SRL except for Encouragement of Effort Investment. Female teachers tended to promote SRL to a greater extent (Table 5).

No differences in encouraging SRL were found based on teaching experience and educational level (primary/secondary school – grammar/vocational school).

The Levene's test showed that all variables have homogeneous variance except for Encouragement of Effort Investment (therefore, an additional non-parametric test was conducted which yielded the same

TABLE 4 Indicators of descriptive statistics for the SRL Encouragement Scale in the 1st measurement point ($N = 251$) and the 2nd measurement point ($N = 179$).

	M_1	SD_1	M_2	SD_2
Encouragement of learning planning and learning organization strategies (ELPLO)	3.85	0.61	4.04	0.37
Encouragement of metacognitive monitoring of learning (EMML)	4.48	0.48	4.57	0.42
Encouragement of elaboration and evaluation (EEE)	3.77	0.61	3.89	0.34
Encouragement of understanding (EU)	4.43	0.41	4.51	0.62
Encouragement of effort investment (EEI)	4.68	0.40	4.75	0.56

TABLE 5 Testing the significance of differences in encouraging SRL with respect to gender (1st measurement point).

	M_M	M_W	t	p
	$N = 48$	$M = 203$		
Encouragement of learning planning and learning organization strategies (ELPLO)	3.66	3.90	-2.48	0.01
Encouragement of metacognitive monitoring of learning (EMML)	4.28	4.53	-3.29	0.00
Encouragement of elaboration and evaluation (EEE)	3.44	3.85	-4.28	0.00
Encouragement of understanding (EU)	4.20	4.48	-4.36	0.00
Encouragement of effort investment (EEI)	4.60	4.70	-1.59	0.11

M, male teachers; W, female teachers.

TABLE 6 Testing the significance of differences in encouraging SRL based on subject status – science field (1st measurement point).

	M_1	M_2	M_3	M_4	F	p
	$N = 25$	$N = 117$	$N = 70$	$N = 39$		
Encouragement of learning planning and learning organization strategies (ELPLO)	4.03	3.97	3.71	3.64	5.45	0.00
Encouragement of metacognitive monitoring of learning (EMML)	4.53	4.57	4.37	4.37	3.40	0.02
Encouragement of elaboration and evaluation (EEE)	3.83	3.88	3.54	3.82	4.67	0.00
Encouragement of understanding (EU)	4.43	4.51	4.37	4.27	4.12	0.01
Encouragement of effort investment (EEI)	4.65	4.73	4.66	4.59	1.47	0.22

M_1 = social; M_2 = humanities; M_3 = natural; M_4 = technical.

results; for consistency in presentation, see the included the parametric test). Statistically significant differences were found based on the science field which the subject in question belongs to, in all components of promoting SRL except for Encouragement of Effort Investment. The Bonferroni test revealed that the difference in ELPLO was statistically significant in relation to teachers whose subject belongs to the humanities field and teachers in the natural and technical field. The difference in EMML and EEE was significant between teachers in the humanities and natural science field, while the difference in EU was observed between teachers in the humanities and technical field. Teachers in the humanities field tend to promote SRL more than teachers in the natural and technical fields (Table 6).

The Levene's test further showed that the variables Encouragement of Elaboration and Evaluation and Encouragement of Understanding exhibited homogeneous variance, while the other three did not (hence, for variables without homogeneous variance, nonparametric tests were additionally conducted, which yielded the same results, but for

consistency of presentation, see the included parametric test). Statistically significant differences were found based on the subject status (compulsory/elective) in ELPLO, EMML, and EEE, while differences in EU and EEI were not statistically significant. Teachers of elective subjects more strongly encouraged SRL (Table 7).

4. Discussion

Given the accelerated scientific, technological, and societal advancements that require competencies which enable lifelong learning and equip students to face the challenges of modern life, the self-regulated learning (SRL) construct becomes extremely important in the educational context, as well as outside of it. In order for teachers to promote SRL, they need to comprehend the dynamics of SRL and the dynamics of creating a learning environment that fosters SRL (Boekaerts, 2002), which means they need to possess substantial content

TABLE 7 Testing the significance of differences in promoting SRL based on subject status – compulsory/elective (1st measurement point).

	M_o	M_i	t	p
	$N = 227$	$M = 24$		
Encouragement of learning planning and learning organization strategies (ELPLO)	3.83	4.11	−2.18	0.03
Encouragement of metacognitive monitoring of learning (EMML)	4.46	4.70	−2.40	0.02
Encouragement of elaboration and evaluation (EEE)	3.73	4.15	−3.30	0.00
Encouragement of understanding (EU)	4.41	4.57	−1.81	0.07
Encouragement of effort investment (EEI)	4.67	4.77	−1.15	0.25

M_o , compulsory school subject; M_i , elective school subject.

and methodological knowledge about SRL (Askell-Williams et al., 2012; Karlen et al., 2020). To gain a better understanding of teachers' knowledge about SRL, we employed both qualitative (open-ended questions) and quantitative (assessment scale) methods. Due to limitations of scale-based self-assessments, which can prompt teachers to evaluate aspects they have not previously considered, or lack sufficient knowledge about and find unfamiliar, thus yielding desirable or acceptable responses, we also implemented open-ended questions in the questionnaire. To address the question of how much teachers know about SRL, we asked if they have heard of the SRL construct and to provide a subjective definition of SRL. The results of the quantitative methodology revealed that one-third of the teachers had not heard of the SRL construct. Although 68% of teachers have heard of SRL, an analysis of teachers' subjective definitions indicates that only 14% of teachers accurately define SRL, while as many as 65% of teachers provide partially accurate definitions of SRL. A slightly smaller percentage (21%) of those who are unsure or inaccurately define SRL compared to the percentage of those who have not heard of SRL could be a consequence of the partially accurate definitions, given that the substantial number of teachers stated that SRL is best defined as "... learning that the student personally designs/regulates/directs..." An analysis of the categories teachers use when defining SRL reveals that teachers most frequently emphasize metacognitive strategies of monitoring, regulating, and self-directing their own learning, as well as autonomy in learning. Similar findings were reported by Dignath and Sprenger (2020). In general, these results align with other authors' findings that teachers' knowledge of SRL is generally weak (Spruce and Bol, 2015; Karlen et al., 2020) and that most teachers encourage student autonomy (Dignath-van Ewijk and van der Werf, 2012). Teachers mainly grant students autonomy in self-regulation, but they do not adequately prepare them to handle new challenges (Bolhuis and Voeten, 2001). If students are only given autonomy without providing them with the means to employ strategies, it is not beneficial for them (Kirschner et al., 2006). When defining SRL, teachers also highlighted learning methods (cognitive strategies), planning and organizing learning, as well as intrinsic motivation.

In order to see the ways in which, and the degree to which the teachers promote SRL in students, we asked them the following: "How do you encourage self-regulated learning in your practice? Which methods are most effective?" (open-ended question), and we applied the Self-Regulated Learning Encouragement Scale. In teachers' responses, we can identify metacognitive, cognitive, motivational, and contextual aspects of SRL, encouragement processes for SRL in different learning phases (Zimmerman, 2002), as well as strategies for direct and indirect promotion of SRL (Dignath-van Ewijk and van der Werf, 2012). Teachers noted metacognitive aspects (planning,

goal-setting, organization, monitoring, self-evaluation), cognitive aspects (repetition, organization – mind mapping, tables, hierarchical organization, elaboration – paraphrasing, summarizing, connecting information, identifying key concepts, asking questions, mnemonic techniques...), motivational aspects (motivation encouragement, emotion regulation, beliefs about oneself...), and contextual aspects (teaching methods: research, project-based learning, collaborative learning, classroom atmosphere, etc.) of SRL. Individual approach, planning and organizing learning, initiating teaching with stimulating and interesting tasks and questions can be categorized as a process of encouraging SRL before learning. Teachers predominantly listed strategies for promoting SRL that can be associated with the learning phases. Teachers reported using various teaching methods, encouraging different learning strategies, fostering positive emotions, regulating negative emotions, motivation, positive beliefs, performance control, encouragement, praise, utilizing information and communication technology, enhancing student concentration, promoting self-control and independent work, giving instructions, guidance, repetition or practice, assigning tasks, and so on. Student self-evaluation, peer assessment, and providing feedback are processes that can be linked to the post-learning phase. Given the complexity, process-oriented nature, continuous timeline, variability, and cyclical nature of self-regulated learning, some of these processes can occur in all phases of SRL (encouragement, feedback, and self-reflection). In terms of direct encouragement, processes include modelling, guided learning, task/teaching analysis, comparing approaches, contemplating effective learning methods, metacognitive monitoring, etc. Indirect processes involve creating a learning environment (active teaching and learning methods, encouraging positive and regulating negative emotions, motivation, resource checking, beliefs about oneself, classroom environment, etc.). Teachers most frequently mentioned promoting planning and organization of learning, self-evaluation, independence, goal-setting, inquiry-based learning, motivation enhancement, collaborative learning, providing feedback, creating mind maps, encouraging monitoring/performance control, fostering beliefs about oneself, and preparing presentations. On the other hand, when analyzing teachers' self-assessments, they estimate that they significantly promote SRL. They mostly encourage effort investment, followed by metacognitive monitoring of learning and understanding, while they least promote learning planning, organizational strategies, elaboration, and learning evaluation. We can observe a certain discrepancy between qualitative and quantitative data. It should be noted that only a small number of teachers in their responses covered a greater number of various SRL-encouragement processes, while on an individual level, the answers were partial (the teachers noted only a few means and processes they encouraged). A

comprehensive image of the processes teachers can encourage is obtained through collective analysis of all teacher responses. Conversely, teachers provide more favorable self-assessments, but a limitation is evident in the content of statements in the SRL Encouragement Scale. In open-ended questions, teachers most frequently mentioned promoting learning planning and organization, self-evaluation, and independence in learning, while the assessment scale showed the opposite trend. It is possible that the previous examination using the SRL Encouragement Scale influenced teachers' responses regarding the ways of promoting SRL in the second measurement point, as there was an increase in the arithmetic means on the SRL Encouragement Scale in the second measurement point. However, this influence may not be strong since the order of results across subscales did not change. In addition, individual teacher responses continued to indicate a lack of knowledge about ways to promote SRL. Despite this methodological limitation, teachers' responses will serve for the revision of the SRL Encouragement Scale. It should be highlighted that teachers did not spontaneously highlight direct teaching of effective learning strategies. They mention modelling and guided learning but in the context of content learning. Teachers more frequently noted implicit and indirect SRL-encouragement approaches. This is in line with the results of previous research, which suggest that teachers prefer using active teaching methods but do not instruct students on how to learn (de Kock et al., 2005; Dignath-van Ewijk and van der Werf, 2012). Students' self-regulation is predicted by teaching metacognitive strategies (planning, organization, goal-setting, self-monitoring, self-evaluation) and creating a learning environment that demands and enables SRL (Dignath-van Ewijk and van der Werf, 2012; Dignath-van Ewijk and Büttner, 2013). Although most teachers agree with the concept of supporting their students to become self-regulated learners, many express uncertainty about how to do so (Perry et al., 2008). Therefore, understanding whether teachers know how to enhance their students' SRL, and at what level to initiate this education is crucial for teaching SRL. The more teachers know about SRL, the greater the likelihood that they will be able to promote it in their students (Paris and Winograd, 2001; Spruce and Bol, 2015). In other words, if teachers do not possess well-developed knowledge about learning, it is unlikely that they will effectively guide their students in developing knowledge of cognitive and metacognitive learning strategies (Askill-Williams et al., 2012). In general, we can conclude that teachers partially understand what SRL is, and within a limited scope, they tend to encourage SRL by emphasizing self-direction and autonomy, which reflects implicit and indirect encouragement. In other words, the first hypothesis has been confirmed.

The testing of differences in promoting SRL based on gender in this study revealed that female teachers significantly encourage all components of SRL, except for effort investment, which is in line with the results obtained by Elmas et al. (2011) and Yan (2018). Chen (2000) discussed gender roles in teaching with respect to feminine and masculine characteristics, emphasizing that women tended to be "caring teachers" who more often used student-centered, active, and constructivist teaching methods, while men tended to adopt a teacher-centered approach, maintaining an authoritative figure that aligns with their societal role. Although some studies show that older teachers provide less support for self-regulated learning (Peeters et al., 2015) and that teachers in lower grades of primary school (up to the 6th grade) mostly encourage self-regulated learning (Moos and Ringdal, 2012), our

study did not find differences when it comes to encouraging SRL with respect to teaching experience and educational level (primary/secondary school – gymnasium/vocational). While Fauzi and Widjajanti (2018) found that mathematics teachers more frequently use teaching techniques that promote SRL, Coggin (2020) did not find differences among teachers of different subjects; our research shows that SRL is most encouraged by teachers in the humanities, followed by the social sciences, then the natural sciences, and finally the technical fields, although the differences are mostly statistically significant between teachers in the humanities and those in the natural sciences and technical fields. It's possible that teachers in the natural sciences and technical fields lack pedagogical and methodological knowledge more than teachers in the humanities and social sciences. Teachers of elective subjects tend to encourage learning planning and organizational strategies, metacognition, elaboration, and evaluation more than the teachers of compulsory subjects. This finding is not surprising since elective subjects attract motivated students who choose the subject based on their abilities and interests. It's possible that teachers adopt a more relaxed approach and are more willing to experiment with various teaching methods in said elective courses.

The research provides insight into teachers' knowledge and ways of encouraging SRL among teachers in Croatia. Teachers exhibit weak knowledge of SRL, i.e., lack of competencies in the field of SRL. A collective analysis of responses of all teachers regarding ways to encourage SRL suggests that teachers are capable of promoting all SRL components (cognitive, metacognitive, motivational, and contextual) and processes across different SRL phases (preparation, execution, and reflection), even though this is not evident in their individual responses. These results confirm the theoretical foundations of SRL and the ways to encourage it. The results also indicate insufficient use of direct explicit demonstration of self-regulation strategies with explanations of why, when, and how to apply specific strategies in terms of informed instruction and self-control training. The data also indicates that SRL encouragement depends on teacher characteristics (gender), subject area, and subject status. Its advantage is that it combines both qualitative and quantitative methods, which provides us with a clearer insight into their SRL knowledge and ways of encouragement. The limitations of this research primarily refer to a convenience sampling method and contextual specificity, making the results difficult to generalize. However, considering the heterogeneity of the sample (in terms of gender, teaching experience, school type, subject field, and subject status), and the fact that the results on SRL knowledge and the ways of SRL encouragement align with the findings of other authors, the data has indicative value. Of course, encouraging SRL among teachers also depends on the educational policies of each country, such as the transition from a traditional content- and teacher-centered approach to a student-centered, constructivist approach, and these differences across educational contexts are to be taken into consideration. Future research could focus on comparing SRL encouragement across different cultures or educational contexts. Nevertheless, the conducted research has significant practical implications. In 2019, the Ministry of Science and Education in Croatia introduced a curriculum decision on the topic "Learning to Learn" for primary and secondary schools, and yet only a small number of teachers possess the necessary knowledge about SRL and ways of encouraging it. In other words, it's important to develop models for enhancing and strengthening teachers' knowledge and competencies in fostering SRL within the educational system. These models should be integrated into training programs for both current

teachers and future educators. These programs ought to be aimed at developing teachers' content knowledge about the nature and dynamics of SRL, as well as improving their pedagogical knowledge and teacher competencies: both their own competencies in SRL (teachers as self-regulated learners) and their competencies in teaching, diagnosing, and supporting SRL in the classroom (teachers as agents of SRL). There should be a particular emphasis on direct and explicit teaching, informed training, and self-control training for students. The results also suggest that greater attention should be devoted to enhancing teaching competencies among male teachers, teachers in technical and natural sciences fields, as well as teachers of compulsory school subjects. Future research should continue to explore the role of different factors at the level of teachers, students, schools, and educational systems when it comes to encouraging SRL.

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 humans were approved by Ministry of Science and Education of the Republic of Croatia; The Ethics Committee of the University of Zadar. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because the subjects were of legal age, participation in the research was voluntary, anonymous, and they could withdraw their participation at any time.

References

- Askill-Williams, H., Lawson, M. J., and Skrzypiec, G. (2012). Scaffolding cognitive and metacognitive strategy instruction in regular class lessons. *Instr. Sci.* 40, 413–443. doi: 10.1007/s11251-011-9182-5
- Baumert, J., and Kunter, M. (2013). "The COACTIV model of teachers' professional competence" in *Cognitive activation in the mathematics classroom and professional competence of teachers*. eds. M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss and M. Neubrand (New York, NY: Springer), 25–48.
- Boekaerts, M. (2002). Bringing about change in the classroom: strengths and weaknesses of the self-regulated learning approach – EARLI presidential address, 2001. *Learn. Instr.* 12, 589–604. doi: 10.1016/S0959-4752(02)00010-5
- Bolhuis, S., and Voeten, M. J. (2001). Toward self-directed learning in secondary schools: what do teachers do? *Teach. Teach. Educ.* 17, 837–855. doi: 10.1016/S0742-051X(01)00034-8
- Brown, A., Campione, J. C., and Day, J. D. (1981). Learning to learn: on training students to learn from texts. *Educ. Res.* 10, 14–21. doi: 10.3102/0013189X010002014
- Chatzistamatiou, M., Dermitzaki, I., and Bagiatis, V. (2013). Self-regulatory teaching in mathematics: relations to teachers' motivation, affect and professional commitment. *Eur. J. Psychol. Educ.* 29, 295–310. doi: 10.1007/s10212-013-0199-9
- Chaves, E., Trujillo, J. M., and López, J. A. (2015). Autorregulación del aprendizaje en entornos personales de aprendizaje en el grado de educación primaria de la Universidad de Granada, España. *Form. Univ.* 8, 63–76. doi: 10.4067/S0718-50062015000400008
- Chaves, E., Trujillo, J. M., and López, J. A. (2016). Acciones para la autorregulación del aprendizaje en entornos personales. *Pixel-Bit. Rev. Medios Educ.* 48, 67–82.
- Chen, Y. M. (2000). *Feminization in writing pedagogy - a study of teacher's gender at EFL university composition classrooms*. Research Report, Taiwan: National Chung Cheng University.
- Coggin, J. (2020). Self-regulated learning Instruction's relationships with teacher subject area, teacher beliefs, and teacher efficacy. Dissertations University of Southern Mississippi. 1793. Available at: <https://aquila.usm.edu/dissertations/1793> (Accessed July 23, 2023).
- De Corte, E., Verschaffel, L., and Masui, C. (2004). The CLIA-model: a framework for designing powerful learning environments for thinking and problem solving. *Eur. J. Psychol. Educ.* 19, 365–384. doi: 10.1007/BF03173216
- de Kock, A., Slegers, P., and Voeten, M. J. M. (2005). New learning and choices of secondary school teachers when arranging learning environments. *Teach. Teach. Educ.* 21, 799–816. doi: 10.1016/j.tate.2005.05.012
- De Smul, M., Heirweg, S., Devos, G., and Van Keer, H. (2018). School and teacher determinants underlying teachers' implementation of self-regulated learning in primary education. *Res. Pap. Educ.* 34, 701–724. doi: 10.1080/02671522.2018.1536888
- Dignath, C., and Büttner, G. (2018). Teachers' direct and indirect promotion of self-regulated learning in primary and secondary school mathematics classes - insights from video-based classroom observations and teacher interviews. *Metacogn. Learn.* 13, 127–157. doi: 10.1007/s11409-018-9181-x
- Dignath, C., and Sprenger, L. (2020). Can you only diagnose what you know? The relation between teachers' self-regulation of learning concepts and their assessment of students' self-regulation. *Front. Educ.* 5:585683. doi: 10.3389/feduc.2020.585683
- Dignath-van Ewijk, C., and Büttner, G. (2013). Assessing how teachers enhance self-regulated learning: a multiperspective approach. *J. Cogn. Educ. Psychol.* 12, 338–358. doi: 10.1891/1945-8959.12.3.338
- Dignath-van Ewijk, C., and van der Werf, G. (2012). What teachers think about self-regulated learning: investigating teacher beliefs and teacher behavior of enhancing students' selfregulation. *Educ. Res. Int.* 2012, 1–10. doi: 10.1155/2012/741713
- Elmas, R., Demirdöğen, B., and Geban, Ö. (2011). Preservice chemistry teachers' images about science teaching in their future classrooms. *Hacettepe Univ. J. Educ.* 40, 164–175.

Author contributions

SŠ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing. AN-M: Investigation, Methodology, Writing – review and editing. MC: Investigation, Methodology, Writing – review and editing. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was financed by the University of Zadar as an institutional project under the code IP.01.2021.18.

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.

- Fauzi, A., and Widjajanti, D. B. (2018). Self-regulated learning: the effect on student's mathematics achievement. *J. Phys. Conf. Ser.* 1097:012139. doi: 10.1088/1742-6596/1097/1/012139
- Hadwin, A., and Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: exploring perspectives of social in self-regulated learning theory. *Teach. Coll. Rec.* 113, 240–264. doi: 10.1177/016146811111300204
- Hamman, D., Berthelot, J., Saia, J., and Crowley, E. (2000). Teachers' coaching of learning and its relation to students' strategic learning. *J. Educ. Psychol.* 92, 342–348. doi: 10.1037/0022-0663.92.2.342
- Hargraves, A. (2005). Educational change takes ages: life, career and generational factors in teachers' emotional responses to educational change. *Teach. Teach. Educ.* 21, 967–983. doi: 10.1016/j.tate.2005.06.007
- Karlen, Y., Hertel, S., and Hirt, C. N. (2020). Teachers' professional competences in self-regulated learning: an approach to integrate teachers' competences as self-regulated learners and as agents of self-regulated learning in a holistic manner. *Front. Educ. Teach. Educ.* 5:159. doi: 10.3389/feduc.2020.00159
- Kirschner, P. A., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86. doi: 10.1207/s15326985ep4102_1
- Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Büttner, G., and Klieme, E. (2010). Promotion of self-regulated learning in classrooms: investigating frequency, quality, and consequences for student performance. *Metacogn. Learn.* 5, 157–171. doi: 10.1007/s11409-010-9055-3
- Kistner, S., Rakoczy, K., Otto, B., Klieme, E. I., and Büttner, G. (2015). Teaching learning strategies: the role of instructional context and teacher beliefs. *J. Educ. Res. Online* 7, 176–197. doi: 10.25656/01:11052
- Lombaerts, K., Engels, N. I., and Van Braak, J. (2009). Determinants of teachers' recognitions of self-regulated learning practices in elementary education. *J. Educ. Res.* 102, 163–174. doi: 10.3200/JOER.102.3.163-174
- Lombaerts, K., Engels, N., and Vanderfaeillie, J. (2007). Exploring teachers' actions to promote self-regulated learning practices in primary school. *Educ. Dev. Psychol.* 24, 4–24. doi: 10.1017/S0816512200029187
- Masui, C., and De Corte, E. (2005). Learning to reflect and to attribute constructively as basic components of self-regulated learning. *Br. J. Educ. Psychol.* 75, 351–372. doi: 10.1348/000709905X25030
- Ministry of Science and Education of the Republic of Croatia (2019). Decision on the adoption of the curriculum for the cross-curricular topic learning how to learn for primary and secondary schools in the Republic of Croatia. Available at: https://narodne-novine.nn.hr/clanci/sluzbeni/2019_01_7_154.html (Accessed July 28, 2023).
- Moos, D. C., and Ringdal, A. (2012). Self-regulated learning in the classroom: a literature review on the teacher's role. *Educ. Res. Int.* 2012, 1–15. doi: 10.1155/2012/423284
- Paris, S. G., and Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educ. Psychol.* 36, 89–101. doi: 10.1207/S15326985EP3602_4
- Paris, S. G. I., and Winograd, P. (2001). The role of self-regulated learning in contextual teaching: principles and practices for teacher preparation. Office of Educational Research and Improvement. Available at: <https://files.eric.ed.gov/fulltext/ED479905.pdf> (Accessed July 15, 2023).
- Peeters, J., De Backer, F., Kindekens, A., Jacquet, W., and Lombaerts, K. (2015). Mechanisms underlying primary school teachers' promotion of self-regulated learning. Annual meeting of the American Educational Research Association, Chicago, SAD.
- Perry, N. E., Hutchinson, L., and Thauberger, C. (2008). Talking about teaching self regulated learning: scaffolding student teachers' development and use of practices that promote self-regulated learning. *Int. J. Educ. Res.* 47, 97–108. doi: 10.1016/j.ijer.2007.11.010
- Perry, N. E., Phillips, L., and Hutchinson, L. (2006). Mentoring student teachers to support self-regulated learning. *Elem. Sch. J.* 106, 237–254. doi: 10.1086/501485
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Pract.* 41, 219–225. doi: 10.1207/s15430421tip4104_3
- Šimić Šašić, S., Kalajžić, A., and i Nikčević-Milković, A. (2021). Poticanje samoreguliranog učenja – uloga nekih sociodemografskih čimbenika, učiteljskih percepcija o poučavanju i učenju, predanosti poslu i sagorijevanja na poslu [Promoting Self-Regulated Learning – the Role of some Sociodemographic Factors, Teachers' Perceptions of Teaching and Learning, Commitment to Work and Work Burnout]. *Psihologijske teme*. 30, 467–487.
- Šimić Šašić, S., Nikčević-Milković, A., and Klarin, M. (in press). Validacija Skale poticanja samoreguliranog učenja [Validating the Self-Regulated Learning Encouragement Scale]. *Suvremena psihologija*.
- Spruce, R., and Bol, L. (2015). Teacher beliefs, knowledge, and practice of self-regulated learning. *Metacogn. Learn.* 10, 245–277. doi: 10.1007/s11409-014-9124-0
- Vandeveld, S., Vandenbussche, L., and Van Keer, H. (2012). Stimulating self-regulated learning in primary education: encouraging versus hampering factors for teachers. *Procedia Soc. Behav. Sci.* 69, 1562–1571. doi: 10.1016/j.sbspro.2012.12.099
- Veenman, M. V. J. (2007). The assessment of metacognition: a matter of multi-method designs. EAPA Newsletter of the European Association of Psychological Assessment, 1, 8–9. Available at: http://www.eapa-homepage.org/eapa/_downloads/eapa-newsletter-200710.pdf (Accessed July 15, 2023).
- Verschaffel, L., De Corte, E., Lasure, S., van Vaerenbergh, G., Bogaerts, H., and Ratinckx, E. (1999). Learning to solve mathematical application problems: a design experiment with fifth graders. *Math. Think. Learn.* 1, 195–229. doi: 10.1207/s15327833mtl0103_2
- Wilson, N. S., and Bai, H. (2010). The relationships and impact of teachers' metacognitive knowledge and pedagogical understandings of metacognition. *Metacogn. Learn.* 5, 269–288. doi: 10.1007/s11409-010-9062-4
- Yan, Z. (2018). How teachers' beliefs and demographic variables impact on self-regulated learning instruction. *Educ. Stud.* 44, 564–577. doi: 10.1080/03055698.2017.1382331
- Zimmerman, B. (2002). Becoming a self-regulated learner: an overview. *Theory Pract.* 41, 64–70. doi: 10.1207/s15430421tip4102_2



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Ingrid Pramling Samuelsson,
University of Gothenburg, Sweden
Gül Kadan,
Cankiri Karatekin University, Türkiye

*CORRESPONDENCE

Cornelia Rüdisüli
✉ cornelia.ruedisueli@phzh.ch

[†]These authors share senior authorship

RECEIVED 01 September 2023

ACCEPTED 27 September 2023

PUBLISHED 30 October 2023

CITATION

Rüdisüli C, Duss I, Lannen P and Wustmann
Seiler C (2023) External assessment of teachers'
roles during children's free play and its relation
to types of children's play.
Front. Educ. 8:1287273.
doi: 10.3389/feduc.2023.1287273

COPYRIGHT

© 2023 Rüdisüli, Duss, Lannen and Wustmann
Seiler. 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.

External assessment of teachers' roles during children's free play and its relation to types of children's play

Cornelia Rüdisüli^{1*}, Isabelle Duss², Patricia Lannen^{2†} and
Corina Wustmann Seiler^{1†}

¹Department of Research and Development, Zurich University of Teacher Education, Zurich, Switzerland, ²Research Department, Marie Meierhofer Children's Institute, Associated Institute of the University of Zurich, Zurich, Switzerland

Teachers in early childhood education and care (ECEC) assume various roles in children's free play, such as tutor, classroom manager, co-player, and director. Recent research has shown that teacher-related characteristics such as work experience and classroom-related characteristics such as children's age are significant predictors of teachers' roles. However, these roles have mostly been assessed through self-report rather than external observation using standardized scales. Furthermore, it is unclear how teachers' roles depend on the types of children's play, such as construction play and dramatic play. To address these research gaps, we observed 80 teachers during 291 observation cycles in childcare centers and kindergartens in Switzerland. We developed a standardized observation scale to simultaneously assess both teachers' roles during children's free play and the types of children's play. Additional teacher- and classroom-related predictors were assessed with an online questionnaire for teachers. The results of a multilevel regression analysis showed that the roles observed differed in the type of play in which children engaged. However, teacher- and classroom-related predictors were more relevant and explained more variance in teachers' roles than the types of children's play. The findings may stimulate self-reflection by ECEC teachers on the roles they assume during children's free play.

KEYWORDS

play involvement, teachers' roles, free play, childcare center, kindergarten, type of play, early childhood education and care, multilevel regression analysis

1. Introduction

The natural way young children learn is through free play, defined as a voluntary, intrinsically motivated activity that is initiated and directed by the children (Whitebread et al., 2017; Zosh et al., 2018). Free play time in early childhood education and care (ECEC) is characterized by the children's free choice of activities in contrast to structured educational activities led by the teachers, such as circle time (Slot et al., 2015). Currently, teachers' play support is considered as a key competence of teachers in ECEC (Göncü et al., 2000; Jensen et al., 2019). Teachers' play support refers both to the provision of an optimal learning environment and to teachers' involvement in children's play. However, the intensity and intention of teachers' play involvement that is most appropriate remains empirically unclear (Skene et al., 2022). Moreover, the characteristics that affect teachers' play involvement have yet to be identified. Previous research defined various roles that teachers assume during children's free play, such as co-player, director,

tutor, onlooker, and uninvolved (Enz and Christie, 1993; Johnson et al., 2005). The results of recent questionnaire studies on self-reported teachers' roles provide information about the frequency of roles and explained their variation with teacher-related characteristics such as work experience and classroom-related ones such as the age of children (Ivrendi, 2020; Grigoropoulos, 2021; Wustmann Seiler et al., 2022). However, these questionnaire studies do not report on the roles that teachers assume in specific situations during free play. Instead, they indicate the overall subjective preferences of the teachers for the roles they assume during free play according to their perceptions and intentions. Observational studies that focus on teachers' roles in specific situations are currently sparse (Enz and Christie, 1993; Jung, 2013; Gaviria-Loaiza et al., 2017; Tajik and Singer, 2018) but would be crucial in further developing our understanding of teachers' practice during children's free play. Moreover, if situational play-related characteristics such as the types of free play that children engage in are also considered when attempting to explain why and when teachers assume a certain role, we can move beyond the more stable predictors such as teachers' work experience and children's ages to include the specific dynamics of given situations. Therefore, the present study uses a standardized observation scale to describe teachers' roles during free play while considering the types of children's play as predictors.

1.1. Teachers' play involvement in early childhood education and care

Teachers in ECEC take on central and diverse parts in children's play activities by organizing space, materials, and time for children's play, supporting children's learning through play, helping children to engage in play, and participating in children's play (Wood, 2008; Jensen et al., 2019). However, opinions differ on how teachers should be involved in children's free play (Trawick-Smith and Dziurgot, 2011). From a social constructivist perspective stemming from Vygotsky's work (Wood and Bennett, 1998), teachers' proactive play involvement is required to promote children's development and stimulate their play and learning (Yang, 2013; Zosh et al., 2018). However, concerns have been raised about direct adult involvement in children's free play (Trawick-Smith and Dziurgot, 2011): evidence indicates that too much teacher involvement is perceived by children as intrusive, reduces their interest in the activity, and impairs children's play behavior (Gmitrova and Gmitrov, 2003). As a result, teachers should only be involved in a responsive way when children ask for or need support (Trawick-Smith and Dziurgot, 2011). In contrast, the constructive perspective arising from Piaget's theory (Wood and Bennett, 1998) emphasizes the child's active role and the need for the teacher to remain in the background so as to promote children's self-determination and control in play and to strengthen children's motivation to play (Sandberg, 2002; Frost et al., 2005; McInnes et al., 2011). Reflecting this controversial discussion, teacher play involvement is described on a continuum with different levels of control by children and teachers (Sylva et al., 2011; Trawick-Smith and Dziurgot, 2011; Pyle and Danniels, 2016; Zosh et al., 2018).

To more precisely classify teachers' play involvement, previous studies have examined teachers' positioning in children's free play (inside or outside) or according to dimensions of interaction (e.g., behavioral, emotional, or learning support; e.g., Sylva et al., 2011;

Kucharz et al., 2014; Fleeer, 2015; Broekhuizen et al., 2017). These studies showed that most teachers were positioned outside of children's play and mainly interacted during non-play activities, such as classroom management tasks to monitor children's behavior (Kucharz et al., 2014; Fleeer, 2015). However, it is crucial to consider that teachers' positioning was only assessed during dramatic play (Fleeer, 2015). Moreover, dimensions of interaction reflect professional quality more generally and may not be exclusive to teachers' play involvement (Sylva et al., 2011). In addition, Slot et al. (2015) found that the quality of teacher-child interaction was rated more highly in guided activities than in children's free play. The lack of structure in open-ended free-play situations and broad variety of possible play activities may be challenging for teachers trying to respond appropriately to children's needs. In these situations, teachers may also experience difficulties in providing opportunities to promote children's play and learning without following a script as in guided activities (Reyhing et al., 2019). Therefore, describing the roles that teachers intuitively assume in these unplanned free play situations in detail and understanding what predicts these various roles is an essential first step in finally understanding the value of these roles. For this purpose, an assessment of roles should include not only the positioning of teachers in children's play but also their degree of guidance and their intended behavior to reflect the entire continuum of teachers' involvement in children's free play more accurately.

1.2. Teachers' roles during free play

Six roles adopted by teachers involved during free play have been empirically identified and described in various studies (Enz and Christie, 1993; Johnson et al., 2005; Gaviria-Loaiza et al., 2017; Ivrendi, 2020): (1) The *co-player* participates in children's play by taking small and passive roles without dominating the play. (2) The *play tutor*, also called play leader, enriches children's play with suggestions, questions, and demonstrations and supports the children in case of difficulties or if they cannot find their way into play. The teacher can accomplish this either by assisting from within or from outside children's play. (3) The *director* takes control and direction of children's play by making decisions and giving instructions. (4) The *redirector* guides children's attention to promote academic learning by asking didactic questions. (5) The *stage manager* adapts the spatial and material play and learning environment to children's play needs and discusses and supports the implementation of children's play ideas. (6) The *classroom manager* supports children in emotion regulation or conflict resolution and ensures compliance with behavior rules during children's play. In addition, two roles have been described in which teachers do not actively interact with children: The *onlooker* is available and consciously observes children's play activities, and *uninvolved* teachers are busy with other tasks such as administrative and preparatory activities while children are playing. Previous studies have examined these roles through external assessment in small-scale live observations and video analysis (e.g., Jung, 2013; Gaviria-Loaiza et al., 2017; Tajik and Singer, 2018), and via self-assessments with questionnaires and interviews (e.g., Yang, 2013; Ivrendi, 2020; Wustmann Seiler et al., 2022). To date, the roles of onlooker, stage manager, and classroom manager have most frequently been reported and observed, and the roles of play tutor, uninvolved, and redirector have been observed least frequently (Kontos, 1999; Jung, 2013; Aras,

2016; Gaviria-Loaiza et al., 2017; Ivrendi, 2020; Wustmann Seiler et al., 2022). Depending on the situation, teachers change their roles. But it is assumed that role preferences are stable to a certain extent and can be explained with various predictors (Enz and Christie, 1993).

1.3. Predictors of teachers' roles during free play

Predictors of teachers' roles in children's free play have been related in previous studies to the teacher (e.g., work experience), to the classroom (e.g., children's age), and to children's play. To the best of our knowledge, only one study to date has specifically examined teachers' roles during free play considering situational play-related characteristics such as types of children's play. This was done using audio recordings, and results were reported primarily with descriptive statistics (Kontos, 1999). Three studies recently examined teacher- and classroom-related predictors of teachers' roles during free play in more complex analyses of questionnaire data: a Turkish study in public kindergartens in homogenous age groups for children aged 3–4 years, 5 years, and 6 years (Ivrendi, 2020); a Greek study in ECEC settings with young children with a mean of 2.8 years of age (Grigoropoulos, 2021); and a Swiss study in public kindergartens with mixed-age groups from 4 to 7 years (Wustmann Seiler et al., 2022). The results were heterogeneous in the frequency with which the roles were assumed and in which predictors were relevant. This might be due to the heterogeneous educational settings of these studies in children's age, group size, and teachers' educational backgrounds. Moreover, the geographical regions in which the studies were conducted were also likely to play a role in the heterogeneous results, as another comparative study found significant differences in teachers' play involvement among countries (Van der Aalsvoort et al., 2015).

1.3.1. Teacher-related predictors

Previous studies have identified teacher-related characteristics such as work experience, age, educational background, and professional self-efficacy as significant predictors of teachers' roles during free play. Teachers with more work experience were more likely to report taking an uninvolved role by doing something else without interacting with the children (Ivrendi, 2020; Grigoropoulos, 2021). They were also more likely to report that they assume the role of tutor, enriching children's play (Grigoropoulos, 2021). In addition, these experienced teachers were less likely to report being involved in children's play as co-players and as classroom managers (Wustmann Seiler et al., 2022). Older and better educated teachers were both less likely to report assuming a tutor role and more likely to report the uninvolved role (Grigoropoulos, 2021). Teachers who reported higher levels of professional self-efficacy were also more likely to report assuming the roles of onlooker and stage manager (Wustmann Seiler et al., 2022). This is in line with findings that teachers' professional self-efficacy, such as the belief that they can support even children with heterogeneous needs adaptively, has been shown to affect teachers' pedagogical actions (Zee and Koomen, 2016; Perren et al., 2017).

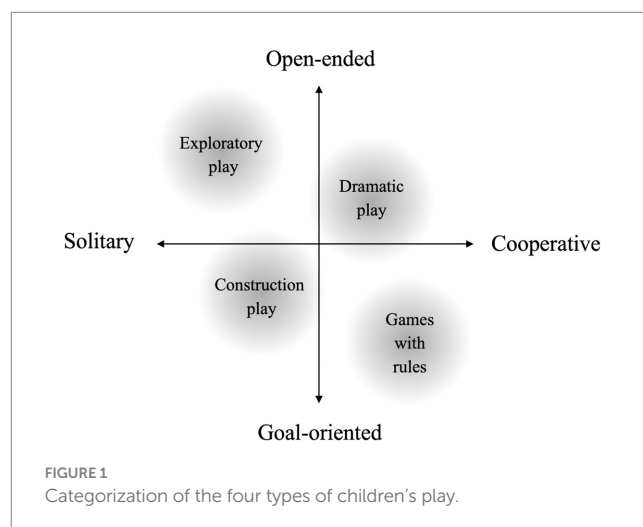
1.3.2. Classroom-related predictors

Several classroom-related characteristics have previously been examined as predictors of teachers' roles during free play, including group composition and play environment. It is generally assumed that the higher the number of children in a classroom, the fewer

teacher–child interactions take place (Howes et al., 2011). However, previous studies have not produced consistent results on how group size predicts the different roles that teachers assume during children's free play. According to Ivrendi (2020), teachers more often reported taking the uninvolved and the co-player role when there were fewer children in the group. However, Grigoropoulos (2021) found that the uninvolved role was more frequently reported when there were more children in the group. In addition, she reported that the more children were in a classroom, the less likely teachers reported assuming the role of tutor (Grigoropoulos, 2021). Another study found that the tutor role was more frequently reported when the group consisted of more children who spoke a foreign language (Wustmann Seiler et al., 2022). Children's age was also shown to predict teachers' roles during free play. Teachers reported assuming the role of director more frequently for 3- and 4-year-olds than for 5- and 6-year-olds (Ivrendi, 2020). The amount of free play time and the number of learning centers with different play and learning opportunities in the classroom have also been identified as predictors for the role a teacher assumes. Teachers who provided more time for free play also reported playing the role of tutor more frequently and therefore providing more frequent support for children's play activities. Teachers who provided less time for free play were more likely to report acting as classroom managers and therefore instead provided more frequent support for the children's social–emotional and behavioral problems (Wustmann Seiler et al., 2022). The number of learning centers was linked to the self-reported roles of onlooker and stage manager: with higher numbers of learning centers, teachers more frequently reported observing children's play behavior and supporting children in organizing play materials or implementing their ideas on what and how to play (Ivrendi, 2020).

1.3.3. Play-related predictors

During free play, children engage in a variety of play types (Johnson et al., 2005). Four basic types of children's play can be distinguished: exploratory play, dramatic play, construction play, and games with rules (Mogel, 2008; Whitebread et al., 2017). These types of children's play differ in the social setting, for instance in solitary play or cooperative play and the purpose of the play activity, which may be open-ended or goal-oriented. Figure 1 represents an attempt to categorize the four types according to the Revised Peer Play Scale (Howes and Matheson, 1992) and the Social/Cognitive Play



Scale (Johnson et al., 1999). However, the classification does not illustrate distinct categories and should be understood as a continuum between the extreme manifestations with overlaps. For example, when children play alone, at the left on the continuum, they play without other playmates around. But it is also possible for children to play on their own next to each other, as in exploratory or construction play, which would be positioned further to the right on the continuum. The same is true for the other side of the continuum. For example, children can play explicitly cooperatively with the same play objectives or content, which requires social negotiations as in dramatic play and games with rules, or they may play together with simple social interactions in their own flow. The continuum of the purpose of the play activity illustrates similar overlapping. A play activity can have a clearly defined goal such as playing a board game, can be goal-oriented with changing goals during the play process such as building a tower, or can be open-ended with many different options in how the play activity develops, but in a defined direction or framework such as pretending to celebrate a birthday party.

Children in exploratory play are basically exploring the world and themselves in solitary play without pursuing a specific goal. This type of children's play is considered to have low cognitive and social demands and can be detected in infants (Pellegrini, 2009). Dramatic play generally presents itself as social play in which children play with other children, adults, or representatives such as dolls in an open-ended way, developing their play gradually and not following a script as in theatre. In contrast, construction play is defined as a goal-oriented activity in which no playmates need be included. Both dramatic and construction play are considered types of play with higher cognitive demands than exploratory play and have been more frequently observed in children's play from the age of 3 to 4 years onwards (Johnson et al., 1999; Mogel, 2008). Games with rules are generally played in company with playmates by pursuing a defined goal. Most of these games place high demands on cognitive skills involving working memory or developing strategies and on social skills involving taking defeats and respecting play rules. Games with rules have been increasingly observed in children's play from the age of 4 to 5 years onwards (Johnson et al., 1999; Mogel, 2008). Qualitative analyses have shown that teachers are most likely to participate in children's play in concrete and goal-oriented play situations such as construction play and games with rules and least likely to participate in dramatic play (Shin and Spodek, 1991). In addition, quantitative analyses of audio recordings showed that teachers spent more of their time participating in product-oriented constructive play than in open-ended play (Kontos, 1999). Further results demonstrated that teachers' proactive play involvement increased children's complexity in social pretend play for instance by decontextualization and role-taking, whereas only managing children's play had no effect (Perren et al., 2019). Moreover, when teachers participated more actively in children's play, a more complex play was observed: social play such as cooperative pretend play and cognitive demanding play such as games with rules (Vu et al., 2015).

1.4. Summary and aims of the study

The roles defined above describe teachers' pedagogical practice in play involvement on a continuum of control by the teacher and the child. Heterogenous findings have been found on

teacher- and classroom-related predictors of these roles. However, previous studies were conducted in diverse countries and ECEC settings (Ivrendi, 2020; Grigoropoulos, 2021; Wustmann Seiler et al., 2022). In addition, most of these studies used questionnaires and do not capture the roles teachers actually assume during free play. To date, situational play-related characteristics have rarely been examined as a predictor, although Johnson et al. (2005) have referred to the variety of play contexts in which children engage during free play and how this complex dynamic may place high demands on teachers and affect teachers' play involvement.

The aim of the present study was to extend previous research by (1) developing an observation scale to capture the range of roles teachers assume during free play in a standardized way, (2) describing the frequency of the roles observed without evaluating the roles' value for children's development, and (3) analyzing how teacher-, classroom-, and play-related characteristics predict these roles. Two different ECEC settings, childcare centers and kindergartens, were included to test for potential setting-specific differences regardless of geographical region. The two settings differ in children's ages, group sizes and teachers' educational backgrounds. Teacher-related predictors included teachers' work experience and teachers' professional self-efficacy. Time for free play and ECEC setting were used as classroom-related predictors. Finally, the four types of children's play constituted the play-related predictors. Due to the heterogeneous findings in previous studies, no hypotheses were formulated to predict the specific roles that teachers assume during free play. However, we assume according to Perren et al. (2019) and Vu et al. (2015) that children need more active play support as the complexity of their play increases. In addition, many of the variables used as teacher- and classroom-related predictors differ between the two ECEC settings. As a result, we expect the distribution of frequency with which teachers assume the various roles to differ between the two settings.

2. Materials and methods

2.1. Procedure

This study was conducted in spring 2021. The recruitment of the participants was part of the "Playfulness in Early Childhood (Playful)" project, funded by the Swiss National Science Foundation. The study involved 80 live observations in childcare centers and kindergartens in Switzerland. We developed a standardized observation scale to externally assess teachers' situational roles during free play in a nonparticipatory observation. In addition, we observed and recorded children's types of play as teachers interacted with children. Data from several observation cycles were collected per teacher, and additional variables were assessed with an online questionnaire.

The study was reviewed and approved by the Ethics Committee of Faculty of Philosophy at the University of Zurich, Switzerland, in January 2021 (Ethics approval number 20.12.13). Teachers were informed about the aims and procedures of the study with a written study description. In addition, they were informed about their right to terminate their participation at any time without stating any reason and were assured that their data would be collected anonymously and

evaluated exclusively for scientific research purposes. All participants provided informed consent before participation.

2.2. The context of ECEC in Switzerland

In Switzerland, two ECEC institutions are distinguished: childcare centers and kindergartens. Childcare centers are private institutions for children from infancy onwards, mostly until entry of kindergarten, and are attended by around 33% of children aged 0 to 3 years ([Bundesamt für Statistik, 2021](#)). Kindergartens are part of the public school system and compulsory for children approximately from the age of 4 to 6 years. Both settings also differ in teachers' education. Teachers in kindergarten hold a bachelor level degree whereas teachers in childcare centers have completed vocational training. However, in both settings a distinction is made between activities in child-centered free play and activities that are led by teachers ([Stamm, 2009](#)).

2.3. Sample

The sample was recruited through a broad call for applications in German-speaking cantons of Switzerland via different channels such as newsletters, school administrations, and organizations of ECEC. When selecting the sample, we stratified by urban and rural areas, cantons, and different pedagogical settings. Eighty teachers in the German-speaking part of Switzerland participated in the study. Of these, 93.8% were female, and 88.8% were Swiss citizens. They were between 20 and 63 years of age ($M = 37.85$, $SD = 12.1$) and had between 1 and 38 years of work experience ($M = 14.1$, $SD = 10.6$). They worked in two ECEC settings: kindergartens ($n = 47$, including 4 kindergartens for children with special needs) and childcare centers ($n = 33$). All the participating teachers were observed, and 97.5% completed the online questionnaire.

2.4. Study measures

2.4.1. Observation scale to assess teachers' roles during free play

An external observation scale was developed to assess teachers' situational roles during children's free play [Teacher Roles during Free play – Observation scale (TRFP-O)] in standardized live observations. The scale was developed from the items of the Teacher Roles in Free Play (TRFP) questionnaire ([Ivrendi, 2020](#)) and its German translation and extension ([Wustmann Seiler et al., 2022](#)). First, a conceptual framework was developed to describe the various roles of the teachers that were observed in specific situations during free play (see [Table 1](#)).

TABLE 1 Conceptual framework of teachers' roles during free play.

	Proactive		Reactive
With reference to children's play	Co-player	Director	Tutor & stage manager
Without reference to children's play	Redirector		Classroom manager

These roles are classified by whether teachers' behavior refers to children's play. The roles are also differentiated according to teachers' activity level, whether the teacher acted proactively or reacted to the immediate needs of the children during their play. Because the focus of the observation was on the teachers and the type of children's play while interacting with the children, the two passive roles without clear interaction, uninvolved and onlooker, were excluded for this purpose. Second, roles that did not differ from each other when classified by activity level and reference to children's play were combined. This was the case for the roles of tutor and stage manager, which are therefore referred to below as *tutor and stage manager*. Similarly, co-player and director were both classified the same way in the two dimensions. However, we decided to keep the two roles separate, because the teachers' positioning is inside the play as co-player and outside the play as director. We adapted the description of the director role to the Swiss context, because the literature describes the director as taking the leadership of the play situation and making decisions about the theme, material, or group composition ([Gaviria-Loaiza et al., 2017](#); [Ivrendi, 2020](#)). However, pilot tests using video observations revealed that in Swiss ECEC settings, this does not correspond to the understanding of child-centered free play and was therefore seldomly observed. As a result, we included teachers' efforts to affect children's play without direct instruction, for example, by proactively giving hints or comments about children's play without the children asking for it, as an additional characteristic of the director role. The description of each role and coding examples can be found in [Table 2](#).

Roles were rated by trained researchers on a five-point scale (1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always) by how much of the time the teacher spent in one of the roles for free-play situations in each observation cycle of between 15 and 20 min. In this study, each teacher was observed for 3 to 5 observation cycles ($M = 3.7$, $SD = 0.82$; $N = 291$) during three to four hours. A total of five trained researcher were involved in the data assessment. Ratings were usually done by one trained researcher. For this study, 19% of the sample were rated by a second rater for quality control. Interrater reliability for teachers' roles was satisfactory ($ICC = 0.72$ – 0.93). The training for the researchers consisted of a one-day session including various inputs covering the procedure and theoretical framework of the observation scale. The training also involved coaching, exercises with videos of teachers' play involvement, and comparing the ratings of these video-based exercises.

2.4.2. Observation scale to assess the type of children's play

During the same observation cycles, the trained researchers noted the type of play children were engaged in while interacting with the teacher. When several types of play occurred during a single cycle, the one that was most prevalent during that time was noted dichotomously (1 = yes, 0 = no). The types of play were derived from the four categories of the Social/Cognitive Play Scale ([Johnson et al., 1999](#)): (1) Exploratory play is characterized by a sensomotor process with or without objects or using objects in a stereotypical way (e.g., rough and tumble play, playing with water, sand, or other objects without pursuing a specific goal, dancing without choreography). (2) Dramatic play is characterized by dramaturgical elements or acting as-if (e.g., imitating everyday actions by dress up or playing at the doll's house). (3) Construction play is characterized by a goal-oriented and constructive process of building with play materials (e.g., building a

TABLE 2 Operationalization of the five active roles of teachers during free play.

Roles	Role description	Coding examples
Co-player	As <i>co-players</i> , teachers participate in children's play in minor functions or roles, at children's request or on their own initiative. They leave the leadership of the play to the children and join in as a play partner at the children's level, without directing children's play. Parallel play, when teachers play for themselves next to the children, is also considered co-play.	The teacher proactively participates in children's play (at the children's request or on his/her own initiative): • Takes on a role in pretend play. • Plays a rule game with the children. • Dances or sings along.
Director	In the role of <i>director</i> , teachers give hints, comments, and instructions about the children's play on their own initiative - without prompting, demand, or identifiable need for assistance. They take over the control and guidance of the children's play and try to influence it by making decisions (about the topic, material or group composition).	The teacher proactively supports children's play (on his/her own initiative): • Determines group compositions or assigns children to different play corners/offers or roles. • Makes comments or asks questions about children's play. • Suggests other play strategies or new play opportunities.
Tutor and stage Manager	In the role of <i>tutor and stage manager</i> , the teachers support the children's play when it gets bogged down, when the children cannot find their way into a play, or when they send out signals of support (with questions or gestures). The teachers then react with hints, suggestions, or questions. The teachers support and advise the children during play and adapt the physical play and learning environment to the children's interests and needs.	The teacher reacts to the request or needs of the children regarding their play: • Provides assistance when children ask for it or when play cannot continue. • Provided additional play materials or creating space. • Is eager to integrate children who are not playing.
Redirector	In the role of <i>redirector</i> , the teachers implement activities with individual children to promote specific skills during the free play time. Didactic questions from the teachers (with a focus on specific learning content) and general conversations with the children during play time are also evaluated as redirectors.	The teacher proactively interacts with the children without reference to their play (on his/her own initiative): • Chats with individual children about their daily routines without reference to their current play. • Directs an activity of the children specifically (e.g., a handicraft, telling a book). • Asks didactic questions or talks to children about specific learning content (e.g., letters, numbers, and colors).
Classroom manager	In their role as <i>classroom managers</i> , the teachers react to the children's needs, which are not related to their current play, but to their immediate behavior. They intervene in a regulating manner in the event of conflicts, rule violations or behavioral problems. When teachers gain an overview of the group's activities, this is also seen as a classroom manager.	The teacher reacts to the request or needs of the children without reference to their play: • Accompanies children in conflict situations and supports them in emotion regulation. • Intervenes in the event of inappropriate or dangerous behavior. • Gives children practical everyday support, e.g., undoing a trouser button, putting on shoes, blowing their nose.

tower or a hut, painting a picture, laying patterns, modeling with playdough). (4) Lastly, games with rules are characterized by clear game procedures and objectives (e.g., board games, puzzles, hide-and-seek games).

2.4.3. Questionnaire to assess teacher-related and classroom-related characteristics

The teacher-related characteristics of work experience and professional self-efficacy and the classroom-related characteristics of ECEC setting and duration of free play time were assessed with an online questionnaire via the Survalyzer program. Professional self-efficacy was evaluated with the German version of the Professional Self-Efficacy Scale (Perren et al., 2017). The scale consists of nine items, which are assessed on a five-point scale (1 = not at all true, 5 = completely true, e.g., "I can structure the learning environment in such a way that all children in the group find suitable challenges, even when their needs are very different"). A confirmatory factor analysis

confirmed a one-dimensional scale with a good model fit with two error correlations [$X^2(36) = 137.06$, $p < 0.001$, CFI = 1.00, RMSEA = 0.014, SRMR = 0.058] and satisfactory reliability (Cronbach's $\alpha = 0.76$). ECEC setting and amount of free play time were recorded as dummy variables. Free play time was divided into less or more than 2 h per morning (0 = less than 2 h; 1 = more than 2 h). Settings included childcare center or kindergarten (0 = childcare center, 1 = kindergarten).

2.5. Analytic plan

First, we computed simple descriptive statistics and bivariate correlations to describe the prevalence of the variation in teachers' roles during free play as well as the predictors (see Table 3). Second, we estimated multivariate regression models in which we predicted teachers' roles in free play by the type of children's play and by

teacher-related and classroom-related characteristics. Because the observation cycles (level 1) were nested within teachers and groups (level 2), we calculated intraclass correlation coefficients (ICCs) to explore outcome variation across level 2 units before finally estimating multivariate regression models in a multilevel framework (see Table 4). All externally assessed predictors from each observation cycle (level 1) and self-reported predictors (level 2) were modeled as manifest independent variables. Teachers' roles in free play measured at level 1 entered the model as a manifest dependent variable. The analyses were conducted using MPLUS version 8.8 (Muthén and Muthén, 1998-2018).

3. Results

3.1. Descriptive statistics

There were differences in the frequency of the observed roles and the types of children's play across the overall 291 observation cycles (see Table 3). The reactive roles of tutor and stage manager ($M=2.66$; $SD=1.00$) and classroom manager ($M=2.22$; $SD=0.97$) were the most frequently observed roles, followed by the proactive roles of the director ($M=1.83$; $SD=0.99$) and the redirector ($M=1.82$; $SD=1.00$). The co-player ($M=1.69$; $SD=0.94$) was the least frequently observed role. The mean values are rather low for a five-point scale because the 5 (always) was only rated if no other role could be observed in this cycle, which was very rarely the case. Exploratory play ($M=0.31$; $SD=0.46$) occurred most in interaction with teachers, followed by construction play ($M=0.29$; $SD=0.45$) and dramatic play ($M=0.11$; $SD=0.32$). Games with rules ($M=0.08$; $SD=0.28$) were rarely observed as the most prevalent type of play.

3.2. Bivariate correlations

3.2.1. Intercorrelations of the roles

As the correlation matrix in Table 3 shows, the intercorrelations of the roles demonstrated that the role of classroom manager was significant negatively correlated at a low level with the three roles with reference to children's play of co-player ($r=-0.13$, $p<0.05$), director ($r=-0.12$, $p<0.05$), and tutor and stage manager ($r=-0.13$, $p<0.05$). In addition, the role of redirector correlated significantly negatively with those of co-player ($r=-0.16$, $p<0.01$) and tutor and stage manager ($r=-0.38$, $p<0.001$) at a low to moderate level. No other roles correlated significantly with each other.

3.2.2. Teacher-related characteristics

The correlations between work experience and the roles of co-player ($r=-0.12$, $p<0.05$) and classroom manager ($r=-0.18$, $p<0.01$) were significantly negative, whereas the correlations with the role of director ($r=0.14$, $p<0.05$) and redirector ($r=0.12$, $p<0.05$) were significantly positive. All these significant correlations between teachers' roles and work experience were at a low level. There was no significant correlation between teachers' work experience and the role of tutor and stage manager. Professional self-efficacy was negatively correlated with the tutor and stage manager ($r=-0.25$, $p<0.001$) and positively correlated with the redirector ($r=0.24$, $p<0.001$), both at a low level but highly significant. We found no significant correlation between professional self-efficacy and the other roles.

3.2.3. Classroom-related characteristics

Free play time correlated significantly positively with the classroom manager role ($r=0.13$, $p<0.05$), whereas the setting was significantly positively correlated with the director role ($r=0.14$, $p<0.05$), both at a low level. No other roles had any significant correlation with the classroom-related predictors.

3.2.4. Play-related characteristics

Exploratory play was significantly negatively correlated with the director ($r=-0.19$, $p<0.01$) and positively with the classroom manager ($r=0.15$, $p<0.05$). A small significant positive correlation was found between dramatic play and co-player ($r=0.18$, $p<0.01$). Construction play was significantly negatively correlated with the classroom manager ($r=-0.14$, $p<0.05$) role and significantly positively correlated with the director ($r=0.17$, $p<0.01$) and the tutor and stage manager ($r=0.15$, $p<0.01$) roles. Games with rules correlated significantly positively with the co-player role ($r=0.13$, $p<0.05$) and significantly negatively with that of classroom manager ($r=-0.13$, $p<0.05$). These correlations reported at a low level. No other roles were significantly correlated with play-related characteristics.

3.3. Multilevel regression analysis

The data had a hierarchical structure with several observation cycles per teacher. A random intercept model was chosen to account for the nesting of the data. First, an intercept-only model was run to calculate the ICC values for each role: co-player ($ICC=0.12$), director ($ICC=0.45$), tutor and stage manager ($ICC=0.48$), redirector ($ICC=0.36$), and classroom manager ($ICC=0.20$). These values showed a substantial amount of variance for teacher- and classroom related characteristics at the between-level that needs to be controlled for. The resulting model with 286 observations cycles and 78¹ teachers fitted the data very well [$X^2(60)=167.42$, $p<0.001$, CFI=1.00, RMSEA=0.00, SRMR_{within}=0.001, SRMR_{between}=0.011].

3.3.1. Teacher- and classroom-related predictors at the between-level

Teachers' work experience and professional self-efficacy predicted a few roles with moderate to strong effects (see Table 4). Teachers' work experience was negatively related to the classroom manager ($\beta=-0.47$, $p<0.001$), but not to other roles. Teachers with a higher professional self-efficacy assumed less the role of tutor and stage manager ($\beta=-0.35$, $p<0.01$), and more the role of redirector ($\beta=0.40$, $p<0.01$). No other roles were linked to teachers' professional self-efficacy. Free play time was marginally significantly related to the classroom manager ($\beta=0.31$, $p<0.10$) at a moderate level. However, the setting was moderately negatively associated with the co-player ($\beta=-0.45$, $p<0.05$) and demonstrated that teachers in kindergarten less frequently assumed the co-player role than teachers working in childcare centers. No other significant relations were found. In summary, teacher- and classroom related predictors explained 9–36% of the variance in the roles.

¹ The number of cases decreased due to two missing questionnaires.

TABLE 3 Descriptive statistics and bivariate correlations of all study variables at the level of observation cycles.

		<i>N</i>	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12
Teacher roles during free play																
1	Co-player	289	1.69	0.94	1											
2	Director	286	1.83	0.99	0.00	1										
3	Redirector	284	1.82	1.00	−0.16**	−0.00	1									
4	Tutor and stage manager	290	2.66	1.00	0.09	−0.08	−0.38***	1								
5	Classroom manager	291	2.22	0.97	−0.13*	−0.12*	−0.06	−0.13*	1							
Teacher-related characteristics																
6	Work experience	286	13.07	10.36	−0.12*	0.14*	0.12*	−0.11	−0.18**	1						
7	Professional self-efficacy	286	4.14	0.37	−0.12	0.45	−0.25***	0.24***	0.02	0.19**	1					
Classroom-related characteristics																
8	Free play time ^b	286	0.42	0.49	−0.05	−0.11	0.08	0.02	0.13*	0.03	0.22***	1				
9	Setting ^c	291	0.60	0.49	−0.10	0.14*	0.04	−0.01	−0.05	0.25**	−0.04	−0.17**	1			
Play-related characteristics																
10	Exploratory play ^a	291	0.31	0.46	−0.09	−0.19**	−0.11	0.00	0.15*	−0.05	0.07	0.01	−0.19***	1		
11	Dramatic play ^a	291	0.11	0.32	0.18**	−0.00	0.09	−0.01	−0.03	0.10	−0.02	−0.06	0.03	−0.24***	1	
12	Construction play ^a	291	0.29	0.45	−0.01	0.17**	−0.19	0.15**	−0.14*	0.02	−0.08	−0.05	0.17**	−0.43***	−0.23***	1
13	Games with rules ^a	291	0.08	0.28	0.13*	0.09	−0.05	−0.05	−0.13*	−0.05	0.04	0.00	0.14*	−0.20***	−0.11	−0.19***

^a 0 = No; 1 = Yes, ^b 0 = Less than 2h; 1 = More than 2h, ^c 0 = Childcare center; 1 = Kindergarten.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

TABLE 4 Multilevel regression model predicting teachers' externally assessed roles during free play.

		Co-player	Director	Redirector	Tutor and stage manager	Classroom manager
		β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
BETWEEN-level						
Teacher-related characteristics						
	Work experience	-0.29 (0.21)	0.15 (0.15)	0.07 (0.16)	-0.11 (0.14)	-0.47*** (0.15)
	Professional self-efficacy	0.07 (0.07)	0.14 (0.15)	0.40** (0.14)	-0.35** (0.11)	0.06 (0.18)
Classroom-related characteristics						
	Free play time ^a	-0.16 (0.17)	-0.18 (0.12)	0.04 (0.14)	0.14 (0.12)	0.31 [†] (0.17)
	Setting ^b	-0.45* (0.20)	0.08 (0.12)	0.12 (0.12)	-0.06 (0.12)	0.15 (0.18)
	R ²	0.36 [†] (0.21)	0.09 (0.07)	0.20 [†] (0.11)	0.15 [†] (0.09)	0.28 [†] (0.15)
WITHIN-level						
Play-related characteristics						
	Exploratory play ^c	0.02 (0.07)	-0.05 (0.07)	-0.16 (0.10)	0.15 (0.10)	0.01 (0.08)
	Dramatic play ^c	0.25* (0.10)	0.05 (0.08)	-0.07 (0.10)	0.21* (0.10)	-0.10 (0.08)
	Construction play ^c	0.15 [†] (0.09)	0.24** (0.09)	-0.09 (0.09)	0.32*** (0.09)	-0.19* (0.08)
	Games with rules ^c	0.23** (0.08)	0.17 [†] (0.09)	-0.17** (0.06)	0.13 (0.10)	-0.21*** (0.06)
	R ²	0.09* (0.04)	0.09 [†] (0.05)	0.03 (0.02)	0.07 [†] (0.04)	0.06 [†] (0.03)

^a 0 = Less than 2 hours; 1 = More than 2 hours; ^b 0 = Childcare center; 1 = Kindergarten; ^c 0 = No; 1 = Yes,

β = Standardized regression coefficient; SE = Standard error.

CFI = 1.00, RMSEA = 0.00, SRMR_{within} = 0.001, SRMR_{between} = 0.011. [†] $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.3.2. Situational play-related predictors at the within-level

The multilevel regression analysis (see Table 4) showed that almost all types of children's play except exploratory play predicted different teachers' roles to varying extents. Dramatic play was significantly positively related to the roles of co-player ($\beta = 0.25$, $p < 0.05$) and tutor and stage manager ($\beta = 0.21$, $p < 0.05$) at a low level but not to other roles. Construction play significantly positively predicted the roles of director ($\beta = 0.24$, $p < 0.01$) and tutor and stage manager ($\beta = 0.32$, $p < 0.001$) at a low to moderate level, and the co-player role marginally at a low level ($\beta = 0.15$, $p < 0.10$). Moreover, construction play was significantly negatively linked to the classroom manager ($\beta = -0.19$, $p < 0.05$) at a low level. Games with rules predicted teachers' roles with low to moderately significant effects. When children were playing games with rules, teachers were more likely to act as a co-player ($\beta = 0.23$, $p < 0.01$) and less in the roles of redirector ($\beta = -0.17$, $p < 0.01$), and classroom manager ($\beta = -0.21$, $p < 0.001$). In addition, a marginally significant relation was found between games with rules and the director role ($\beta = 0.17$, $p < 0.10$). In summary, play-related predictors explained 6–9% of the variance in the roles.

4. Discussion

The aim of the present study was to externally assess the roles of teachers during children's free play in two ECEC settings, childcare centers and kindergartens, and to investigate their relations to teacher-, classroom-, and play-related predictors. For this purpose, we developed a standardized observation scale (TRFP-O) that distinguishes five active roles that teachers may assume during free

play: co-player, director, redirector, tutor and stage manager, and classroom manager. Teacher- and classroom-related predictors included teachers' work experience, teachers' professional self-efficacy, available free play time, and the ECEC setting. In addition, a play-related predictor was included and operationalized as the type of children's play. This contained four types of children's play: exploratory play, dramatic play, construction play, and games with rules. To the best of our knowledge, this is the first study to address three levels of predictors of teachers' roles in free play and the situational characteristics of the type of children's play with which they engage. In addition, teachers' roles in free play were analyzed using a multilevel approach that considered the hierarchical data structure. The study succeeded in observing and analyzing data from 80 teachers, a significantly larger sample than previously reported in observational studies about teachers' roles during free play (e.g., File, 1994; Tajik and Singer, 2018). Finally, data was assessed using a newly developed standardized observation scale, adding to previous studies that either assessed primarily qualitative, unstandardized data (Jung, 2013; Gaviria-Loaiza et al., 2017) or self-reported questionnaire data (e.g., Ivrendi, 2020; Wustmann Seiler et al., 2022). Our results demonstrate that teachers' roles during free play are mainly explained by teacher- and classroom-related predictors and less by the type of children's play.

4.1. Assessment of teachers' roles during free play using the TRPF-O

The study showed that it is possible to reliably distinguish and assess the five active roles assumed by teachers during free play using our external observation scale (TRFP-O). The interrater reliability

with intraclass coefficients from 0.72 to 0.93 was satisfactory. Findings on the frequency of the roles were congruent with previous studies showing that the tutor, stage manager, and classroom manager roles were observed most frequently (Aras, 2016; Gaviria-Loaiza et al., 2017; Ivrendi, 2020; Wustmann Seiler et al., 2022). However, there were also discrepancies with previous studies, probably due to different types of assessment such as questionnaires and observations. For example, the role of co-player was reported as the second most preferred role in the questionnaire study by Ivrendi (2020) but was observed least frequently in the present study. It seems that teachers overestimate the amount of time they actually participate as co-players in children's play. In addition, the discrepancy might be due to the difference in samples in group size (smaller groups; Ivrendi, 2020) or children's ages (younger children; Grigoropoulos, 2021). The director role was also reported more frequently when assessed with questionnaires (Grigoropoulos, 2021) than in our study which assesses the roles through observation in live situations. This was the case even though this particular role was already softened and operationalized as less directive as part of the live observation. However, the director role is also favored mainly in the context of younger children (Ivrendi, 2020; Grigoropoulos, 2021). In contrast, the role of redirector was not assessed in the questionnaire studies, hence no data is available for a comparison. In previous qualitative observational studies, the redirector role was rarely observed, just as it was in the present study (Gaviria-Loaiza et al., 2017).

4.2. The significance of teacher-related predictors

The teacher-related characteristics of work experience and professional self-efficacy were found to be strong predictors for certain roles. For example, work experience proved to be a significant predictor of the role of classroom manager: the more experience a teacher had, the less the role of classroom manager was observed. This is in line with Wustmann Seiler et al.'s (2022) findings and indicates that experienced teachers might react less to children's rule violations or socio-emotional problems. It is also possible that teachers with more work experience are better able to prevent such situations altogether, for example through an elaborated organization of the play and learning environment or through more effective promotion of the children's social-emotional competences. If teachers have fewer day-to-day classroom management tasks as a result, they have more time for other activities that support children's play or learning. However, the result of the present study does not indicate which roles these teachers then assume more often in this extra time. Previous studies have indicated that experienced teachers interact less with children during free play and more frequently report the passive uninvolved role (Ivrendi, 2020; Grigoropoulos, 2021). This finding aligns with the constructive perspective according to Piaget's theory and rather reflects a more traditional image of older and thus more experienced teachers who believe that adults should stay outside of children's free play so as not to interrupt children's natural learning (Wood and Bennett, 1998; O'Connor, 2014). More recently, Vygotsky's social constructivist theory has been favored and discussed in relation to the concept of guided play (Zosh et al., 2018). Professional self-efficacy predicted the role of the redirector in the present study. Teachers who rated themselves higher in professional self-efficacy

were more likely to focus on children's academic learning during free play. This aligns with the existing body of evidence in which professional self-efficacy is positively related to students' academic adjustment, mediated by instructional support (Zee and Koomen, 2016). In the present study, however, this instructional support related more to academic learning than to children's learning through play, because teachers who rated themselves higher in their professional self-efficacy were observed less in the reactive role of tutor and stage manager, which helps children to implement their play ideas or to find their way into play if necessary. In contrast, in Wustmann Seiler et al.'s (2022) questionnaire study, teachers with higher self-efficacy reported acting more frequently as tutors. It seems that teachers perceive themselves differently in the role of tutor than they appear in the external assessment.

4.3. The relationship between teachers' roles and classroom-related predictors

The ECEC setting proved to be the strongest classroom-related predictor. Results showed that teachers in kindergarten act less as co-players who play actively with children than teachers in childcare centers. This may be due to the different characteristics that are inherent to the two settings. For example, children are *per se* older in kindergarten, the group size is larger, and the teachers have a higher educational qualification than is the case in childcare centers. The correlation of group size with the co-player role is congruent with Ivrendi's (2020) finding that teachers were more actively involved when there were fewer children in the group. However, it is possible that teachers are generally more proactive in supporting younger children's play due to their developmental age and possibly less play experience. This was also shown in Ivrendi's (2020) questionnaire study, which found that teachers of older children were less likely to take the proactive role of the director giving children suggestions and making decisions about their play. Another classroom-related predictor, the amount of time children spent in free play, proved to be less significant for externally assessed teachers' roles, contrary to Wustmann Seiler et al.'s (2022) findings in a questionnaire study. In other words, teachers who provide more than 2 h of free play time on average were not involved differently in children's play than teachers who provide less time for free play.

4.4. Differences in teachers' roles during free play depending on the type of children's play

The roles teachers assumed during free play were explained by three types of children's play which were mainly categorized as goal-oriented or cooperative: dramatic play, construction play, and games with rules. During more goal-oriented types of children's play, construction play and games with rules, more roles were observed with reference to children's play, co-player and director, and fewer roles without, classroom manager and redirector. Moreover, during more cooperative types of children's play, dramatic play and games with rules, teachers were more likely to assume the proactive role of co-player. The reactive role of the tutor and stage manager was observed most frequently during dramatic and construction play. It

appears that children repeatedly show a need for support during more goal-oriented and cooperative play without specific rules, for example, by asking the teachers for advice or play materials. The findings are consistent with current research that states that most teachers were involved in types of children's play that are goal- and product-oriented (Kontos, 1999; Tsai, 2015). In contrast to Shin and Spodek (1991) and File (1994), teachers were also reactively and proactively involved in open-ended dramatic play. The two roles without reference to children's play, redirector and classroom manager, were observed less frequently during goal-oriented types of play: construction play and games with rules. This indicates that children in goal-oriented types of play require less classroom management. In addition, teachers seem to focus less on academic content in these types of children's play because children are already engaged in cognitively demanding play. In contrast, exploratory play, which makes the fewest social and cognitive demands, was not confirmed as a predictor of teachers' roles. Exploratory play focuses on children's sensomotor processes (Johnson et al., 1999), so teachers may want to deliberately withdraw to enable children to have a variety of experience in such play. Also, it is possible that no specific need for teacher involvement arises in these open-ended types of children's play.

4.5. Strengths and limitations

The present study was the first to use a standardized observation scale to assess teachers' roles during free play. The study also examined various characteristics related to the teacher, the classroom, and the type of play in which children are engaged that may influence teachers' roles in this play. We drew data from 291 observation cycles in two Swiss ECEC settings: childcare centers and kindergartens. In addition, the situational play-related predictors at the observation cycle level were examined in a multilevel analysis for the first time. The five roles were rated on the developed scale (TRFP-O) to ensure that they were clearly distinguishable and achieve the most objective recording of roles and to allow observer agreement. The observation scale was used on a relatively large sample with satisfactory interrater reliability. The five roles varied in how teachers positioned themselves during play and their teachers' activity level, representing their guidance and control in play. Furthermore, the roles differed in the reference of teachers' behavior to children's play reflecting teachers' intention, such as supporting children's play processes or supporting academic learning or social-emotional behavior. However, this classification is highly complex and can sometimes seem confusing. Further studies could examine how to reduce complexity when categorizing the roles without losing critical information. In further quantitative studies, for instance on the effectiveness of play involvement on children's development, the complexity could be reduced by assessing the roles in combination or by focusing solely on selected roles, depending on the research question. In addition, future studies should also include the passive roles of onlooker and uninvolved so that the whole continuum of teachers' play involvement can be represented. The advantage of external assessment over questionnaire studies is the absence of bias arising from teachers' subjective perceptions and social desirability. The disadvantage is that the ratings only reflect what was evident at that specific time of observation. It is quite likely that atypical situations

occur, or participants behave differently than usual during the observation due to social desirability. However, because the observers stayed in the background and were present for several hours, it can be assumed that individual atypical situations were less salient. However, to demonstrate the reliability of the TRFP-O observation scale, its test-retest reliability should be analyzed in future studies. Furthermore, it is recommended to redefine the highest value of the five-point scale from *always* to *very often*, so that the whole scale can be increasingly exploited. In addition, it remains open whether the correlations with the setting were due to the age of the children, the group size, or teachers' educational background. Moreover, further studies should capture the specific characteristics of the children during observation, including age, playfulness, and support needs, to examine the extent to which teachers adapt their play involvement to children's needs. A good fit between the children's needs and the level of adult guidance has been described as a prerequisite for children's independent play (Trawick-Smith and Dziurgot, 2011). The authors showed that teachers with a higher level of education succeeded in doing so better than teachers with a lower level of education. In addition, the participants had registered for the study voluntarily. Therefore, it is possible that teachers who were particularly interested in the topic of play and playfulness were involved. Teachers who are less interested in play may act differently during children's free play, so a risk of selection bias cannot be excluded. Finally, the sample size is limited by the extensive and complex data collection involved in external observation.

4.6. Implications for practice

The results described above stimulate reflection on the roles the teachers assumed in free play and indicate potential for development in teacher education and training. The following questions can, for example, be addressed: Why do teachers with increasing professional experience and higher self-efficacy not increasingly take on active roles with reference to children's play: co-player, director, tutor and stage manager? What are the reasons for teachers' restraint in open-ended types of children's play? Is it justifiable for teachers to foster academic learning during free play if it interrupts children's play? Presenting teachers with such questions increases their awareness of their own role in free play. In this way, high-quality interactions that promote both children's play and development can be initiated intentionally.

5. Conclusion

The present external assessment of teachers' roles in free play showed that teachers were involved in children's play in ECEC in various ways. Using a newly developed scale (TRFP-O) enabled observation of five active roles: co-player, director, redirector, tutor and stage manager, and classroom manager. We were able to show that teachers were more proactively involved, for example as co-player and director in goal-oriented types of children's play than in open-ended types. However, the roles that teachers assumed during children's free play depend more on the teachers' work experience and professional self-efficacy and less on the type of play

in which children are engaged. The setting, whether childcare center or kindergarten, was a significant predictor only for the role of co-player, but not for the other roles. This indicates that setting differences are not relevant for most of the roles that teachers assume during free play.

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 humans were approved by Ethics Committee of the Faculty of Philosophy of the University of Zurich. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CR: Data curation, Formal Analysis, Investigation, Writing – original draft. ID: Data curation, Investigation, Writing – review & editing. PL: Conceptualization, Funding acquisition, Methodology, Writing – review & editing, Supervision. CW: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – review & editing, Supervision.

References

- Aras, S. (2016). Free play in early childhood education: a phenomenological study. *Early Child Dev. Care* 186, 1173–1184. doi: 10.1080/03004430.2015.1083558
- Broekhuizen, M. L., Slot, P. L., van Aken, M. A., and Dubas, J. S. (2017). Teachers' emotional and behavioral support and preschoolers' self-regulation: relations with social and emotional skills during play. *Early Educ. Dev.* 28, 135–153. doi: 10.1080/10409289.2016.1206458
- Bundesamt für Statistik. (2021). Familien in der Schweiz. [Families in Switzerland]. Available at: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/familien/familienergaenzende-kinderbetreuung.assetdetail.17084546.html> [Accessed July 29, 2023].
- Enz, B., and Christie, J. F. (1993). "Teacher play interaction styles and their impact on children's oral language and literacy play." in *Paper presented at the annual meeting of the National Reading Conference, Charleston, USA*. Available at: <https://files.eric.ed.gov/fulltext/ED366015.pdf> [Accessed May 25, 2023].
- File, N. (1994). Children's play, teacher-child interactions, and teacher beliefs in integrated early childhood programs. *Early Child. Res. Q.* 9, 223–240. doi: 10.1016/0885-2006(94)90007-8
- Fleer, M. (2015). Pedagogical positioning in play – teachers being inside and outside of children's imaginary play. *Early Child Dev. Care* 185, 1801–1814. doi: 10.1080/03004430.2015.1028393
- Frost, J. L., Wortham, S. C., and Reifel, S. (2005). *Play and child development* (2nd ed.). New York: Pearson Education.
- Gaviria-Loaiza, J., Han, M., Vu, J. A., and Hustedt, J. (2017). Children's responses to different types of teacher involvement during free play. *J. Childh. Stud.* 42, 4–19. doi: 10.18357/jcs.v42i3.17890
- Gmitrova, V., and Gmitrov, J. (2003). The impact of teacher-directed and child-directed pretend play on cognitive competence in kindergarten children. *Early Childhood Educ. J.* 30, 241–246. doi: 10.1023/A:1023339724780
- Göncü, A., Mistry, J., and Mosier, C. (2000). Cultural variations in the play of toddlers. *Int. J. Behav. Dev.* 24, 321–329. doi: 10.1080/01650250050118303
- Grigoropoulos, I. (2021). Greek early childhood educators' engagement practices and facilitative roles during free play. *Contemp. Sch. Psychol.* 27, 1–10. doi: 10.1007/s40688-021-00400-6
- Howes, C., Guerra, A. W., Fuligni, A., Zucker, E., Lee, L., Obregon, N. B., et al. (2011). Classroom dimensions predict early peer interaction when children are diverse in ethnicity, race, and home language. *Early Child. Res. Q.* 26, 399–408. doi: 10.1016/j.ecresq.2011.02.004
- Howes, C., and Matheson, C. C. (1992). Sequences in the development of competent play with peers: social and social pretend play. *Dev. Psychol.* 28, 961–974. doi: 10.1037/0012-1649.28.5.961
- Ivrendi, A. (2020). Early childhood teacher's roles in free play. *Early Years* 40, 273–286. doi: 10.1080/09575146.2017.1403416
- Jensen, H., Pyle, A., Zosh, J. M., Ebrahim, H. B., Scherman, A. Z., Reunamo, J., et al. (2019). *Play facilitation: The science behind the art of engaging young children (white paper)*. The LEGO Foundation, DK.
- Johnson, J., Christie, J., and Wardle, F. (2005). *Play, development, and early education*. Upper Saddle River, NJ: Pearson.
- Johnson, J., Christie, J., and Yawkey, T. (1999). *Play and early childhood development*. 2nd ed. New York: Longman.
- Jung, J. (2013). Teachers' role in infants play and its changing nature in a dynamic group care context. *Early Child. Res. Q.* 28, 187–198. doi: 10.1016/j.ecresq.2012.05.001
- Kontos, S. (1999). Preschool teachers' talk, roles, and activity settings during free play. *Early Child. Res. Q.* 14, 363–382. doi: 10.1016/S0885-2006(99)00016-2
- Kucharz, D., Mackowiak, K., Zioli, S., Kauertz, A., Rathgeb-Schnierer, E., and Dieck, M. (2014). Professionelles Handeln im Elementarbereich (PRIMEL). Eine deutsch-schweizerische Videostudie. [Professional action in elementary school.] Münster u.a.: Waxmann.
- McInnes, K., Howard, J., Miles, G., and Crowley, K. (2011). Differences in practitioners' understanding of play and how this influences pedagogy and children's perceptions of play. *Early Years* 31, 121–133. doi: 10.1080/09575146.2011.572870
- Mogel, H. (2008). *Psychologie des Kinderspiels: Von den frühesten Spielen bis zum Computerspiel, die Bedeutung des Spiels als Lebensform des Kindes, seine Funktion und Wirksamkeit für die kindliche Entwicklung* (3. Auflage). [Psychology of children's play]. Heidelberg: Springer.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Swiss National Science Foundation (Grant No. 100019_192422).

Acknowledgments

The authors thank all teachers for their participation in the study and the numerous research assistants for their contributions to data assessment and data preparation. Our thanks also go to Simon Milligan for editing the paper for language.

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.

- Muthén, L. K., and Muthén, B. O. (1998–2018). *MPlus user's guide (8th ed.)*. Los Angeles, CA: Muthén & Muthén.
- O'Connor, K. (2014). Spiel und Pädagogik im Kindergarten. Eine qualitative Studie zu Einstellungen von Erzieherinnen. [Play and pedagogy in kindergarten. A qualitative study on the attitudes of educators] München: Deutsches Jugendinstitut e.V.
- Pellegrini, A. D. (2009). *The role of play in human development*. Oxford University Press, USA.
- Perren, S., Herrmann, S., Iljuschin, I., Frei, D., Körner, C., and Sticca, F. (2017). Child-centred educational practice in different early education settings: associations with professionals' attitudes, self-efficacy, and professional background. *Early Child. Res. Q.* 38, 137–148. doi: 10.1016/j.ecresq.2016.07.001
- Perren, S., Sticca, F., Weiss-Hanselmann, B., and Burkhardt Bossi, C. (2019). Let us play together! Can play tutoring stimulate children's social pretend play level? *J. Early Child. Res.* 17, 205–210. doi: 10.1177/1476718X19849248
- Pyle, A., and Danniel, E. (2016). A continuum of play-based learning: the role of the teacher in play-based pedagogy and the fear of hijacking play. *Early Educ. Dev.* 28, 274–289. doi: 10.1080/10409289.2016.1220771
- Reyhing, Y., Frei, D., Bossi, C. B., and Perren, S. (2019). Die Bedeutung situativer Charakteristiken und struktureller Rahmenbedingungen für die Qualität der unterstützenden Fachkraft-Kind-Interaktion in Kindertagesstätten [The importance of situational characteristics and structural framework conditions for the quality of supportive teacher-child interaction in childcare centers]. *Zeitschrift Pädagogische Psychol.* 33, 33–47. doi: 10.1024/1010-0652/a000233
- Sandberg, A. (2002). Children's concepts of teachers' ways of relating to play. *Aust. J. Early Childh.* 27, 18–22. doi: 10.1177/183693910202700405
- Shin, E., and Spodek, B. (1991). "The relationship between children's play patterns and types of teacher intervention." in *Paper presented at the annual conference of the American Educational Research Association, Chicago*.
- Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., and Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Dev.* 93, 1162–1180. doi: 10.1111/cdev.13730
- Slot, P. L., Leseman, P. P., Verhagen, J., and Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Child. Res. Q.* 33, 64–76. doi: 10.1016/j.ecresq.2015.06.001
- Stamm, M. (2009). Frühkindliche Bildung in der Schweiz. Eine Grundlagenstudie im Auftrag der UNESCO-Kommission Schweiz (Monograph). [Early childhood education in Switzerland.] Fribourg: University of Fribourg. Available at: <http://edudoc.ch/record/31203/files/FBBE.pdf> [Accessed May 25, 2023].
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (2011). Pre-school quality and educational outcomes at age 11: low quality has little benefit. *J. Early Child. Res.* 9, 109–124. doi: 10.1177/1476718X10387900
- Tajik, M., and Singer, E. (2018). Collaborative research between academics and practitioners to enhance play engagement in the free play of two- and three-year olds. *Early Years* 41, 64–78. doi: 10.1080/09575146.2018.1478391
- Trawick-Smith, J., and Dziurgot, T. (2011). 'Good-fit' teacher-child play interactions and the subsequent autonomous play of preschool children. *Early Child. Res. Q.* 26, 110–123. doi: 10.1016/j.ecresq.2010.04.005
- Tsai, C.-Y. (2015). Am I interfering? Preschool teacher participation in children play. *Univ. J. Educ. Res.* 3, 1028–1033. doi: 10.13189/ujer.2015.031212
- van der Aalsvoort, G., Prakke, B., Howard, J., König, A., and Parkkinen, T. (2015). Trainee teachers' perspectives on play characteristics and their role in children's play: an international comparative study amongst trainees in the Netherlands, Wales, Germany and Finland. *Eur. Early Child. Educ.* 23, 277–292. doi: 10.1080/1350293X.2015.1016807
- Vu, J. A., Han, M., and Buell, M. J. (2015). The effects of in-service training on teachers' beliefs and practices in children's play. *Eur. Early Child. Educ. Res. J.* 23, 444–460. doi: 10.1080/1350293X.2015.1087144
- Whitebread, D., Neale, D., Jensen, H., Liu, C., Solis, S.L., Hopkins, E., et al. (2017). *The role of play in children's development: A review of the evidence (research summary)*. The LEGO Foundation, DK.
- Wood, E. (2008). "Developing a pedagogy of play for the 21st century" in *Early childhood education, society and culture*. eds. A. Anning, J. Cullen and M. Fleer (New York, NY: Sage), 19–30.
- Wood, E., and Bennett, N. (1998). Teachers' theories of play: constructivist or social constructivist? *Early Child Dev. Care* 140, 17–30. doi: 10.1080/0300443981400103
- Wustmann Seiler, C. W., Rüdisüli, C., and von Felten, R. (2022). Was braucht ihr für euer spiel – darf ich mitspielen? Selbstwahrgenommene Spielbegleitung von Lehrpersonen in Schweizer Kindergärten. [what do you need for your play – can I play with you? Self-perceived play support by teachers in Swiss kindergartens.]. *Psychol. Erzieh. Unterr.* 69. doi: 10.2378/peu2022.art09d
- Yang, Y. (2013). A qualitative study of teachers' involvement in children's play. *Lit. Inform. Comp. Educ. J.* 4, 1244–1251. doi: 10.20533/licej.2040.2589.2013.0165
- 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
- Zosh, J., Hirsh-Pasek, K., Hopkins, E., Jensen, H., Liu, C., Neale, D., et al. (2018). Accessing the inaccessible: redefining play as a spectrum. *Front. Psychol.* 9:1124. doi: 10.3389/fpsyg.2018.01124



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo,
Portugal

REVIEWED BY

Pitambar Paudel,
Tribhuvan University, Nepal
Sigamoney Naicker,
University of the Western Cape, South Africa

*CORRESPONDENCE

Steven Van Vaerenbergh
✉ steven.vanvaerenbergh@unican.es
Jessica Fernández-Solana
✉ jfsolana@ubu.es

RECEIVED 19 June 2023

ACCEPTED 03 November 2023

PUBLISHED 21 November 2023

CITATION

Arias-Pastor M, Van Vaerenbergh S,
Fernández-Solana J and
González-Bernal JJ (2023) Perceptions and
preparedness of secondary teacher trainees to
foster inclusive schools for all.
Front. Educ. 8:1242623.
doi: 10.3389/educ.2023.1242623

COPYRIGHT

© 2023 Arias-Pastor, Van Vaerenbergh,
Fernández-Solana and González-Bernal. 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.

Perceptions and preparedness of secondary teacher trainees to foster inclusive schools for all

Mercedes Arias-Pastor¹, Steven Van Vaerenbergh^{2*},
Jessica Fernández-Solana^{3*} and Jerónimo J. González-Bernal³

¹Department of Education, University of Cantabria, Santander, Spain, ²Department of Mathematics, Statistics and Computing, University of Cantabria, Santander, Spain, ³Department of Health Sciences, University of Burgos, Burgos, Spain

Introduction: The advancement of inclusive education over the past few decades emphasizes the pivotal role of teachers in transforming the educational landscape. As schools transition toward a more inclusive approach, it is imperative to evaluate the efficacy of initial teacher training programs in preparing educators for this inclusive transition. This study aims to describe the preparedness and perceptions of students in the Master's Degree in Secondary Education, Vocational Training, and Language Teaching (MDSE) regarding inclusive education, guided by the profile developed by the European Agency for the Development of Special Needs and Inclusive Education (AEDNEEI).

Methods: Furthermore, the research analyzes how external factors, such as perceived self-efficacy, influence these perceptions and readiness for inclusive teaching. A total of 218 students enrolled in the MDSE, with an average age of 31.5 years and a standard deviation of 6, were examined. Of the participants, 33% were male and 67% were female. These students came from different Spanish universities and were either in the final stages of their studies or had already completed them. Prior to participating, they had finished the generic module and completed their practice sessions in secondary education centers. The "Teacher Training in Secondary Education: Key Elements for Teaching in an Inclusive School for All" (TTSE-IN) questionnaire was employed, which includes five validated and pertinent instruments, with four of them being employed for the study's objectives: The "Questionnaire for Future Secondary Education Teachers about Perceptions of Diversity Attention," the "Questionnaire for the Evaluation of Teacher Training for Inclusion CEFI-R," the "Revised Scale of Feelings, Attitudes, and Concerns about Inclusive Education" (SACIE-R), and the "Brief Scale of Teacher Self-Efficacy."

Results and Discussion: The results show the presence of positive attitudes in future educators along with a poor overall assessment of the training received, which raises concerns about the development of teaching functions more related to daily work in the classroom and the implementation of inclusive methodologies. At the same time, both regular and close contact with people in situations of special vulnerability, experience in training in nonformal contexts, and the level of teacher self-efficacy, in its different components, are postulated as some of the facilitators of the development of the postulates and values of inclusive education and predictors of greater capacity toward attention to diversity.

KEYWORDS

inclusive teacher, initial teacher education, secondary education, teacher self-efficacy, school for all

1 Introduction

Today's societal demands, are reflected in international documents such as "Transforming our world: the 2030 Agenda for Sustainable Development" (Armando et al., 2019). Of particular relevance to this study is the "Sustainable Development Goal" (SDG) (UNESCO, 2005) "Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (Duk and Murillo, 2020). Additionally, the recent regulatory changes in the educational field of our country, as seen in Organic Law of Education 3/2020 (LOMLOE) (BOE, 2020), offer us a new opportunity to reflect on the education we offer to students in our system. This education should be grounded in principles of social justice moving away from the deficit model. Instead, it should favoring the perception and experience of diversity as a challenge and opportunity to enrich the ways of teaching and learning (UNESCO, 2005). This is a chance to collaboratively design strategies and paths toward achieving an "Education for all." The initial training of teachers should equip them with the knowledge, competencies, values and attitudes that prepare them as inclusive educators capable of successfully facing the challenge of diversity in their classrooms; It is evident that the transformation of schools, stemming from their cultures and values, policies, and practices (Sandoval et al., 2013), requires the development and promotion of a new teaching role to lead this process (Herrera-Seda, 2018).

The Sixth Additional Provision of the LOMLOE addresses "Education for Sustainable Development and Global Citizenship." It establishes that the achievement of SDG 4 will be taken into account in teacher training processes and in access to the teaching profession, with the projection that by the year 2022, the knowledge, skills, and attitudes related to education for sustainable development and global citizenship (UNESCO, 2022) will have been incorporated into the system for accessing the teaching profession; Furthermore, the Seventh Additional Provision of the LOMLOE indicates that, within this timeframe, a regulatory proposal will be presented to govern, among other aspects, initial and ongoing teacher training, access to the teaching profession, and professional development. And up to this point, this maxim has not been developed, although proposals for the transformation of initial training have been put forward. This issue has also been addressed by relevant authors in our country who contributed proposals based on their background in educational research (Durán and Giné, 2017; Imbernón, 2019).

As we can see, the development of Inclusive Education has underscored the pivotal role of teachers in the transformation of the education system as a whole. As pointed out by the Profile of Competencies for Inclusive Teaching, a result of the Project "Teacher Training for Inclusion" (European Agency for Development in Special Needs Education, 2011, 2012), there are four values that articulate the main competencies to be developed in the initial training of future teachers, with their related contents (concepts, procedures, and attitudes), for their "ethical literacy" (Booth et al., 2015) in this process: "1. Valuing student diversity: differences among students are a resource and a value within education; 2. Supporting all students: teachers expect the best from all their students; 3. Working as a team: collaboration and teamwork are essential approaches for all teachers; 4. Professional and personal development: teaching is a learning activity, and teachers have the responsibility to learn throughout their lives."

This profile, and the preceding it, have formed the basis for numerous investigations and articles (González-Gil et al., 2017; Rodríguez, 2019). Furthermore, after reviewing of studies on

perceptions and training needs for the development of inclusive schools both from professional and personal competencies perspectives (Rodríguez et al., 2019; Rodríguez Fuentes et al., 2021), it is evident that the success of education and, consequently, inclusion, depends largely on the preparation of teachers to teach in contexts where diversity is the norm, enriching the teaching and learning processes for all.

In the secondary education stage, this challenge is even greater because the previous trainings before the MDSE open a wide range of positions, knowledge, and previous experiences. These elements serve as a foundation upon which to build a new perspective that gives rise to Universal Design for Learning (Alba Pastor, 2018), which implies "above all, an attitude, a predisposition to think about the learning needs of all students" (Ruiz Rodríguez, 2019); thus, the attitudes that future teachers possess regarding inclusive education and attention to diversity constitute a critical factor influencing the extent of adaptation, thereby directly impacting students' academic performance (Blotnick-Gallant et al., 2015), and are strongly associated with the teachers' level of training (Pegalajar Palomino and Colmenero Ruiz, 2017).

Therefore, knowing the attitudes of MDSE students toward attention to diversity and inclusive education is crucial. Equally important is discerning, their perceived level of competencies, skills, and abilities for their professional development as inclusive teachers. Together, these insights, are important step toward improving the education system and achieving maximum results; furthermore, identifying their strengths and weaknesses will help to enhance the resources that universities, and among them, their different levels and areas of impact on society, can promote and develop to support them in their work from the necessary perspective of the Ecology of Equity (Ainscow et al., 2012, 2013).

In addition, several authors (Holdheide and Reschly, 2008) indicate that if teachers do not feel prepared to work with all students, the challenge is to promote and improve the training processes, based on their teaching needs, increasing the sense of teacher self-efficacy. If this feeling is positive, it will certainly generate negative attitudes toward inclusive schools and attention to diversity, along with higher levels of concern and stress that may result in failure to implement educational practices (Sucaticona, 2016). We cannot ignore that according to research (Prieto Navarro, 2002; Collado-Sanchis et al., 2020), these beliefs of teacher self-efficacy are a key element that influences the development of educational strategies that teachers consciously or unconsciously carry out in their classrooms.

Therefore, it seems that both the perception of attention to diversity and inclusive schools, the acquired competencies and attitudes of future teachers toward these, as well as the level and sign of perceived self-efficacy for their work are key aspects to make inclusion effective as a reality in current education.

Now, we must ask ourselves if the initial training through the MDSE program impacts these aspects through the transmission and exemplification of associated conceptual, attitudinal, and procedural knowledge; if this training is related to a more positive attitude toward diversity, and if it influences a higher perception of self-efficacy, a lower level of concerns, and a greater capacity to deploy strategies in the classroom that are adapted to the reality and diversity of each student. Thus, in the conclusions, as the UNESCO report on "Education for All" (ONU, 2014) and the 2020 Education Monitoring Report (UNESCO, 2020) show, proposals will be collected that can help in the development of one of the key strategies for improving the

school institution: the improvement of teacher training so that they obtain the necessary training that promotes positive attitudes toward diversity, questioning of educational reality, and the search for alternatives that overcome inequalities (Sales Ciges, 2006; Collado-Sanchis et al., 2020).

The aim, therefore, is twofold: to describe attitudes and perceptions toward diversity and the level of acquisition of competencies and knowledge required to become inclusive teachers, and to identify the initial relationship of the results with other variables relevant to this study that may act as facilitators or barriers to the professional and personal development of future educators in 21st-century education. Achieving these goals will allow us to provide data that can contribute to the development of new curricula in our country.

2 Materials and methods

2.1 Study design

This is a cross-sectional research that uses a questionnaire as an instrument to collect information. In addition, a correlational study was carried out on the perceptions of MDSE students related to their attitude, concerns, and training toward attention to diversity in inclusive schools, assessing the influence and interaction of the perception of teacher self-efficacy and other relevant variables related to the study object.

The Bioethics Committee of the University of Burgos approved the research (Reference UBU 032/2021), respecting all the requirements established in the Declaration of Helsinki of 1975.

2.2 Participants

The equations should be inserted in editable format from the equation editor. A total of 72 male and 146 female students enrolled in Master's Degree in Secondary Education (MDSE) participated in the study. The study included students from 44 universities in Spain, which consisted of 31 public universities and 13 private institutions. After identifying both public and private universities offering the MDSE, contact was made via email with the MDSE coordinators and faculty members. They were requested for their collaboration in the study and for the distribution of the questionnaire among their students. Out of the 67 identified universities, 44 agreed to participate.

The participants were individuals who had completed their MDSE studies in any specialization during the academic year 2021–2022 or who had finished the Generic Module of the MDSE and finalized their hand-on training at secondary education schools during the academic year 2022–2023. The only requirement for their inclusion in the study was meeting these criteria. The average age of the sample was 31.5 years, with a standard deviation of 6. Specifically, the mean age for women was 31.2 years, while for men, it was 32.3 years.

2.3 Instruments

To gather the necessary data, a comprehensive questionnaire called “Teacher training in secondary education: key elements for

teaching in an inclusive school for all (TTSE-IN)” was developed, consisting of 99 items. This questionnaire incorporated five validated questionnaires based on Spanish samples. For the specific study discussed in this article, which is part of a larger research project, the following instruments from the overall questionnaire were utilized. On the one hand, to measure student attitudes toward diversity and inclusive education, and the perceived level of competencies, skills, and abilities for professional development, three questionnaires were included, with statements rated on a 4 point Likert scale from 1 (totally disagree) to 4 (totally agree). These are:

“Questionnaire for Future Secondary Education Teachers about Perceptions of Diversity (CFDPAD)” (Colmenero Ruiz and Pegalajar Palomino, 2015). This questionnaire collects information using 43 items that exhibit high reliability across all their factors. Factor 1: *Conditioning elements of the diversity attention process in the classroom* ($\alpha=0.959$); Factor 2: *Curricular and organizational response to diversity in the classroom* ($\alpha=0.915$); Factor 3: *Teacher training toward diversity* ($\alpha=0.870$); Factor 4: *Formative teaching practice in diversity attention* ($\alpha=0.906$); Factor 5: *Teacher perception toward students with specific educational support needs* ($\alpha=0.916$).

“Questionnaire for the Evaluation of Teacher Training for Inclusion (CEFI-R)” (González-Gil et al., 2017). It is a multidimensional instrument composed of 16 items that measure the attitudes, competencies, skills, and abilities of teachers and students in different branches of education, including MUPES. The four dimensions are: *Conception of Diversity*, *Methodology*, *Support*, and *Community Participation*. This instrument complements the previous questionnaire and enriches it by providing a perspective on support and community participation in the development of inclusive education.

“Revised Scale of Feelings, Attitudes, and Concerns about Inclusive Education (SACIE-R)” (Forlin et al., 2011). It is designed for practicing and preservice teachers and consists of 12 items that measure the perception of inclusive education and the concept of students who belong to it, feelings toward people with disabilities, and concerns about having different students in the classroom. In the Spanish adaptation (Rodríguez et al., 2019), the reliability for students was deemed acceptable, with a Cronbach's alpha value of $\alpha=0.67$, closely mirroring the original version ($\alpha=0.74$).

On the other hand, to measure perceived teacher self-efficacy, the following scale has been used in different studies, among others, with the population under investigation (Mérida-López and Extremera, 2020):

“Brief Scale of Teacher Self-Efficacy” (Mérida-López et al., 2018). The Spanish adapted version of the “Teachers' Sense of Efficacy Scale” by Tschannen-Moran and Hoy (2001) was used to evaluate teacher self-efficacy. This scale measures the perception of teacher self-efficacy in three dimensions: perceived efficacy for optimizing one's own instruction, perceived efficacy for managing the classroom, and efficacy for engaging students in learning. This scale has been shown to be very reliable and to have excellent validity (Klassen et al., 2009). The instrument proves adequate psychometric properties in Spanish samples.

The questionnaire was designed to be completed online, ensuring anonymity and voluntary participation. It included a section where participants could freely contribute additional comments. To facilitate its administration, the questionnaire was implemented using the Google Forms platform. To reach a wide range of participants from

Spanish universities, emails containing the questionnaire link were sent to the MDSE units, and letters were sent directly to students. After collecting the data, a matrix was created to evaluate the responses using IBM SPSS (Statistical Package for the Social Sciences) version 25, a statistical software program.

2.4 Statistical analysis

In addition to examining the collected responses in detail, the data were analyzed using the following methods (Blanca et al., 2017). Initially, a bivariate analysis was conducted, employing the Student's *T*-test to assess variances in central tendencies between questionnaire responses when comparing two groups. This test was utilized for variables such as gender (female vs. male) and the presence or absence of close and regular interaction with individuals in situations of special vulnerability (yes vs. no), teaching experience with people in situations of special vulnerability in non-formal contexts (yes vs. no), and type of institution where the MDSE is studied (public vs. private). Second, ANOVAs were employed to assess variances in central tendencies when the comparative criteria involved more than two groups, and subsequent post-hoc DMS analysis was conducted to elucidate the character of statistically significant distinctions between groups as they were identified.

It should be noted that to manage the results of the Brief Scale of Teacher Self-Efficacy, we used the total score as a more appropriate indicator for future teachers (Tschannen-Moran and Hoy, 2001), while also analyzing the results of its components. For cases in which the significance of the variable for the object of study was identified, we chose to classify the scores of each subject into the categories of "low," "medium," or "high" by calculating the 33rd and 66th percentiles for this sample. We established that the "low" level would range from the lowest score to the score corresponding to the 33rd percentile, the "medium" level would range from the score following the 33rd percentile to the score corresponding to the 66th percentile, and the "high" level would range from the score immediately above the 66th percentile to the maximum score obtained.

3 Results

The sample included 218 individuals (146 women and 72 men) over the age of 31 who participated in the study, representing approximately half of the sample. A total of 43.1% of the sample entered the MDSE program out of vocation, and 43.6% did so for the possibility of accessing a stable job, while the remaining participants did so because they had no better option after finishing their undergraduate degree (4.1%), due to the influence of a professor who marked their training (4.6%), or because of the influence of a family member who is or has been a teacher (4.6%).

A total of 29.4% have completed higher education in the area of Social and Legal Sciences, 31.2% in Arts and Humanities, 2 1.6% Engineering and Architecture, 14.7% in Sciences and 3.2% in Health Sciences. 43.6% of the sample took the MDSE for the possibility of accessing a stable job and 43.1% for vocation. A total of 34.4% had regular contact with vulnerable people under diverse parameters and 29.4% had experience in non-formal teaching. Regarding the universities, 70.5% were publicly owned and 29.5% were private.

The following section presents the results regarding the perception of MDSE students on both diversity attention and the level of acquired competencies and knowledge for inclusive teaching. Additionally, the relationship of the data with the main analyzed grouping variables will be shown, including, among others, the feeling of teacher self-efficacy, to answer the posed questions and open new ones.

3.1 Questionnaire for future secondary education teachers on perceptions of diversity attention CFDPAD

In Factor 1, *Conditioning Elements of the Diversity Attention Process in the Classroom* ($M=3.6743$; $SD=0.2661$), the sample presents a very high level of agreement on the elements that should be present in the teaching and learning process for the development of a quality diversity attention process in the classroom.

Regarding the primary grouping variables examined, significant statistical differences were observed between groups, with higher levels of agreement and positive attitudes among female participants ($t=3.265$; $p=0.001$) and those who had regular contact with people in situations of special vulnerability ($t=2.397$; $p=0.017$).

Regarding the perception of teaching self-efficacy (Tables 1–3), statistically significant differences were found between groups in all three components that make up the total score. Greater agreement on key elements for quality teaching is associated with a greater sense of teaching self-efficacy in the areas of efficacy in engaging students in learning "commitment to students" ($F=2.837$; $p<0.001$), with statistically significant differences between high versus low ($p<0.001$) and medium ($p=0.010$) level groups; perceived efficacy in optimizing one's own instruction "instructional strategies" ($F=2.248$; $p=0.003$), with differences between high and low perceived self-efficacy levels with a significance of $p<0.001$; and perceived efficacy in managing the classroom "classroom management" ($F=1.896$; $p=0.011$).

In Factor 2 ($M=2.289$; $SD=0.773$), *Curricular and Organizational Response to Diversity in the Classroom*, the results show ratings with means that indicate a predominantly negative level, and in all aspects evaluated, regarding the level of training acquired as future teachers to respond to the educational needs of students in the classroom, both through the organization of the classroom, time management, types of grouping management, methodological strategies, measures and programs for attention to diversity, selection and adaptation of objectives, competences and contents, tasks and activities, and the evaluation of the teaching and learning process. The percentage of the sample that does not feel they have received adequate training in the different aspects related to attention to diversity is 75.2% cumulatively, taking into account the different levels of disagreement with the statements. Only 14.8% of the sample perceived their training as adequate or very adequate.

Statistically significant differences have been obtained between the sample studying at a public university versus a private university, with the former presenting a worse evaluation of the training received ($t=-3.972$; $p<0.001$). Statistically significant differences also appear in the variable of maintain close and regular contact with people in situations of special vulnerability ($t=2.484$; $p=0.014$) and in the group of people who had experience in teaching people in situations of special vulnerability in non-formal contexts ($t=2.009$; $p=0.046$).

Regarding teacher self-efficacy (Tables 2, 4), those with a higher level of self-efficacy rated the training more positively as a total score ($F=1.649$; $p=0.012$), with significant differences between high and low levels ($p=0.016$). Similarly, perceived self-efficacy to optimize their own instruction “instructional strategies” shows significance ($F=2.084$; $p=0.006$), with statistically significant differences between the group with high self-efficacy level compared to those with low ($p=0.002$) and medium level ($p=0.041$), the former feeling more capable of responding to diversity in the classroom.

Regarding the assessment of future teachers regarding the training received in the MDSE and their preparation in the field of diversity, the perceived level is generally low. In Factor 3, *Teacher Training for Diversity* ($M=2.312$; $DT=0.727$), the results show significant differences between the students who take the MDSE in a public or private university, with public university students feeling less prepared in the field of diversity ($t=-3.040$; $p=0.003$). There were also statistically significant differences between groups in terms of the feeling of “self-efficacy” as a total score ($F=1.444$; $p=0.049$), with a

higher level of self-efficacy associated with greater preparation (Table 4).

It is noteworthy that despite the overall results, the best ratings in this factor refer to items on the usefulness of MDSE for “affirming my personal choice toward teaching” ($M=2.98$; $SD=1.097$), “strengthening my interest toward a greater awareness of diversity” ($M=2.75$; $SD=1.149$), and “having a greater awareness of diversity” ($M=2.75$; $SD=1.149$).

The results in Factor 4 ($M=2.333$; $DT=0.530$), *Teaching Practice for Diversity*, which analyzes how the training received in the Master’s Degree allows future teachers to respond to the interests and concerns detected in students with special educational needs in the school context, show a perception of the poor impact of the training received on future inclusive teaching practice, with activities or practical examples developed in classes not improving their knowledge of diversity in general, and not considering themselves sufficiently qualified to face the challenge in the classroom. The item with the highest level of agreement was “A quality diversity care process

TABLE 1 Differences between CFDPAD factors and self-efficacy “Commitment to students” using ANOVA.

CFDPAD factors	Self-efficacy <i>Commitment to students</i>	N	Mean	SD	F	p
F1. Conditioning elements of the diversity attention Process in the classroom	Low	74	3.607	0.299	2.837	<0.001
	Medium	81	3.659	0.264		
	High	63	3.772	0.191		
	Total	218	3.674	0.266		

TABLE 2 Differences between CFDPAD factors and self-efficacy “Instructional strategies” using ANOVA.

CFDPAD factors	Self-efficacy <i>Instructional strategies</i>	N	Mean	SD	F	p
F1. Conditioning elements of the diversity attention process in the classroom	Low	92	3.599	0.307	2.248	0.003
	Medium	55	3.678	0.268		
	High	71	3.768	0.156		
	Total	218	3.674	0.266		
F2. Curricular and organizational response to diversity in the classroom	Low	92	2.144	0.751	2.084	0.006
	Medium	55	2.236	0.666		
	High	71	2.517	0.834		
	Total	218	2.289	0.773		
F4. Teaching practice for diversity.	Low	92	2.284	0.507	2.083	0.006
	Medium	55	2.214	0.469		
	High	71	2.487	0.574		
	Total	218	2.333	0.530		

TABLE 3 Differences between CFDPAD factors and self-efficacy “Classroom management” using ANOVA.

CFDPAD factors	Self-efficacy <i>Classroom management</i>	N	Mean	SD	F	p
F1. Conditioning elements of the diversity attention process in the classroom	Low	76	3.652	0.284	1.896	0.011
	Medium	77	3.650	0.263		
	High	65	3.728	0.242		
	Total	218	3.674	0.266		

TABLE 4 Differences between CFDPAD factors and self-efficacy “Total score” using ANOVA.

CFDPAD factors	Self-efficacy <i>Total score</i>	N	Mean	SD	F	p
F2. Curricular and organizational response to diversity in the classroom	Low	74	3.593	0.317	1.649	0.012
	Medium	71	3.676	0.252		
	High	73	3.754	0.189		
	Total	218	3.674	0.266		
F3. Teacher training for diversity	Low	74	2.255	0.686	1.444	0.049
	Medium	71	2.264	0.639		
	High	73	2.417	0.837		
	Total	218	2.312	0.727		

requires previous experience with students with educational needs” ($M = 3.17$; $SD = 0.750$).

Again, there are statistically significant differences between students who take the Master’s Degree in a public university and those who take it in a private university ($t = -2.743$; $p = 0.007$), with the latter having a more positive rating on the impact of the training on their future teaching practice. There were also significant differences between groups regarding the feeling of self-efficacy “instructional strategies” ($F = 2.083$; $p = 0.006$) (Table 2), with a higher perceived positive impact of the training received for students with a high level of self-efficacy in terms of their ability to improve their teaching skills compared to those with low ($p = 0.015$) or medium levels ($p = 0.004$).

In Factor 5 ($M = 2.695$; $DT = 0.539$), *Teacher Perception of Students with Specific Educational Support Needs in the School Context*, which assesses the attitudes of future teachers toward diversity and responding to educational needs in the classroom in general, the results show relatively positive attitudes resulting from the agreement on inclusive approaches with the perception of added workload. A total of 64.1% of the sample disagrees with the approach of schooling in special education centers, and 81% agrees that “having students with educational needs in the classroom is an added.”

3.2 Teacher training evaluation questionnaire for inclusion CEFI-R

The results in Dimension 1 *Conception of Diversity* (Table 5) show the assessment of the sample regarding the concept of diversity, the place and way in which the schooling of students is considered, the educational policy that forms the basis for these choices, and, in the end, the personal interpretation of inclusive education (European Agency for Development in Special Needs Education, 2011). Nearly half of the sample (48.6%) feels concern again about the increase in workload when having students with specific educational support needs in the classroom, although 78.4% do not consider that a student with these needs interrupts the routine of the classroom or harms the learning of others.

In Dimension 2 *Methodology* (Table 5), regarding the preparation of future teachers for the development of methodological strategies, resources, materials, communication techniques, and evaluation from an inclusive framework, the sample appears polarized, with approximately half of the participants trained and educated in this regard. The use of communication techniques as a tool for inclusion

is the competence in which participants feel most capable ($M = 2.73$; $SD = 0.797$), with 69.3% of the total agreeing with this situation. On the other hand, “adapting my way of evaluating to the individual needs of each of my students” ($M = 2.39$; $SD = 0.868$) and “teaching each of my students differently based on their individual characteristics” ($M = 2.44$; $SD = 0.818$) are the weak points presented by future teachers.

This *Methodology* dimension shows a positive relationship with the variables of having close and regular contact with individuals facing special vulnerability ($t = 2.454$; $p = 0.015$), studying the MDSE at a private university ($t = -3.375$; $p = 0.001$), and having experience in non-formal education for people in situations of vulnerability ($t = 2.609$; $p = 0.010$). Significant statistical differences between groups in the variable of total perceived teacher “self-efficacy” ($F = 2.248$; $p < 0.001$) are also noteworthy, both at this overall level and in each of the components that make it up (Tables 6–9), showing a positive relationship in *post hoc* analysis at its high levels. This group presents a more accurate view of inclusive education and its processes, as well as a feeling of greater preparation for its development through the different elements of inclusive teaching practice. Perceived self-efficacy to optimize one’s own instruction: “Instructional strategies” $F = 2.992$; $p < 0.001$; Perceived self-efficacy to manage the classroom: “Classroom management” $F = 3.058$; $p < 0.001$; Perceived self-efficacy to involve students in learning: “Student engagement” $F = 3.504$; $p < 0.001$.

In Dimension 3 *Supports* (Table 5), which addresses the definition of support, the expected role of support teachers, the possible beneficiaries of this support and where it is provided, the sample shows a greater tendency to position themselves in agreement with the presented inclusive approaches. The highest agreement is given in the items “Joint planning between the support teacher and classroom teacher would facilitate support provision within the classroom,” with 88.2% of the sample in favor of the approach, and “I consider that the place of the support teacher is within the regular classroom with each of the teachers,” with 85.3% in favor. In addition, statistically significant differences are found between groups, studying the MDSE at a public university ($t = 2.187$; $p = 0.003$), with close and regular contact with people in situations of special vulnerability ($t = 2.538$; $p = 0.012$), and with experience in teaching people in situations of special vulnerability in non-formal contexts ($t = 2.420$; $p = 0.017$) positively influencing greater agreement. There are also statistically significant differences between groups in the variable of total teacher “self-efficacy” perception (Table 6), showing a positive relationship in *post-hoc* analysis at high levels compared to low levels ($p = 0.011$), and

TABLE 5 Descriptive statistics CEFI-R dimensions.

CEFI-R dimensions	Range	Min	Max	Mean	SD
D1 Conception of diversity	3.00	1.00	4.00	2.176	0.734
D2 Methodology	3.00	1.00	4.00	2.545	0.665
D3 Supports	3.00	1.00	4.00	3.166	0.584
D4 Community participation	2.67	1.33	4.00	3.623	0.488

TABLE 6 Differences between CEFI-R dimensions and self-efficacy “Total score” using ANOVA.

CEFI-R dimensions	Self-efficacy Total score	N	Mean	SD	F	p
D2. Methodology	Low	74	2.216	0.539	2.248	<0.001
	Medium	71	2.543	0.515		
	High	73	2.879	0.746		
	Total	218	2.545	0.665		
D3. Supports	Low	74	3.060	0.586	3.439	0.034
	Medium	71	3.133	0.571		
	High	73	3.304	0.576		
	Total	218	3.166	0.584		
D4. Community participation	Low	74	3.531	0.570	3.150	0.045
	Medium	71	3.610	0.457		
	High	73	3.730	0.406		
	Total	218	3.623	0.488		

in the perceived self-efficacy for “classroom management” (Table 9) with significant differences between high self-efficacy levels and low ($p=0.03$) and medium ($p=0.016$) levels.

In Dimension 4 (Table 5), *Community Participation*, which collects approaches to collaborative work between the educational center and other community agents, as well as the use of the resources offered by the environment, the results of the sample indicate a high tendency toward agreement on the importance of participation as a key element in the process of attention to diversity in inclusive schools. It is observed how these ideas are more positively shared by the group of women in the sample ($t=2.562$; $p=0.011$) and by those with a greater sense of total teacher “self-efficacy” ($F=3.150$; $p=0.045$); within self-efficacy, it is also positively related to “instructional strategies” ($F=2.101$; $p=0.005$) and for involving students in learning “commitment to students” ($F=2.173$; $p=0.005$) with significant differences in post-hoc analysis between high self-efficacy and low ($p=0.002$) groups.

3.3 Revised scale of feelings, attitudes, and concerns about inclusive education (SACIE-R)

General data about these dimensions were already provided in a previous study in a more comprehensive manner (Arias Pastor et al., 2023). In Factor 1 *Attitudes* ($M=3.1578$; $SD=0.61115$), the sample presents a positive perception of inclusive education and the concept of students who belong to it, showing agreement with the inclusion of students with disabilities and other personal conditions. In addition,

students with difficulties expressing themselves orally and those with attention problems are perceived more positively.

Concerning the examined grouping variables, there are statistically significant differences between the groups in terms of the sex variable, showing a more positive attitude in the female group ($t=2.757$; $p=0.006$).

In Factor 2 ($M=1.4098$; $SD=0.55820$), *Feelings* toward people with disabilities, the results show appropriate feelings with disagreement ratings for the presented approaches related to negative perceptions and emotions between 88.9 and 95.4% of participants. The perceived self-efficacy variable appears to significantly influence optimizing one’s own instruction in “instructional strategies” (Table 10) and in “classroom management” (Table 11) with statistically significant differences between groups, showing more positive feelings in individuals with high levels in these variables.

Last, in Factor 3 ($M=2.9140$; $SD=0.68417$), *Concerns* about having diverse students in the classroom, the findings indicate elevated levels of concern in items associated with teacher preparation to face the challenge of diversity in the classroom, while presenting a lower level of concern in approaches that are more linked to the effects of diversity on teachers themselves. The items that most reflect the level of concern for future teachers are those referring to “I consider it difficult to provide adequate attention to all students in a classroom” ($M=3.43$; $SD=0.732$) and “I am worried about not having the necessary knowledge and skills to teach students with disabilities” ($M=3.30$; $SD=0.867$).

The variables of experience in training in non-formal contexts for people in situations of special vulnerability ($t=-2.080$; $p=0.039$), the area of knowledge of previous studies ($F=2.486$; $p=0.045$), and

TABLE 7 Differences between CEFI-R dimensions and self-efficacy “Instructional strategies” using ANOVA.

CEFI-R dimensions	Self-efficacy <i>Instructional strategies</i>	<i>N</i>	Mean	SD	<i>F</i>	<i>p</i>
D2. Methodology	Low	92	2.308	0.601	2.992	<0.001
	Medium	55	2.494	0.519		
	High	71	2.890	0.706		
	Total	218	2.545	0.665		
D4. Community participation	Low	92	3.568	0.547	2.101	0.005
	Medium	55	3.612	0.452		
	High	71	3.704	0.427		
	Total	218	3.623	0.488		

TABLE 8 Differences between CEFI-R dimensions and self-efficacy “Student engagement” using ANOVA.

CEFI-R dimensions	Self-efficacy <i>Student engagement</i>	<i>N</i>	Mean	SD	<i>F</i>	<i>p</i>
D2. Methodology	Low	74	2.264	0.561	3.504	<0.001
	Medium	81	2.516	0.543		
	High	63	2.911	0.752		
	Total	218	2.545	0.665		
D4. Community participation	Low	74	3.504	0.566	2.173	0.045
	Medium	81	3.625	0.448		
	High	63	3.761	0.403		
	Total	218	3.623	0.488		

motivation to pursue the MDSE ($F=2.958$; $p=0.021$) appear as the variables that directly influence the level of concerns, presenting statistically significant differences between groups. Particularly striking are the differences between groups pursuing the MDSE for “not having a better study option” and for “vocation” ($p=0.035$), and between “the possibility of having a stable job” and “vocation” ($p=0.002$), with individuals pursuing the studies out of vocational interest showing a lower level of concerns and thus lower scores in this factor.

4 Discussion

The main objective of the study is to understand the perception and attitude of future high school teachers toward diversity and their level of acquisition of competences and knowledge in inclusive teacher training, with the Profile of Competences of Inclusive Teaching (European Agency for Development in Special Needs Education, 2012) and its four main values as our reference.

The entire sample showed a favorable attitude toward diversity, recognizing the right to inclusive education for all secondary education students. The future teachers understand that it is the duty of the educational center to attend to all students and, for the most part, they believe that the participation and attitude of the families of students with specific educational support needs, as well as the

support of the school administration, are crucial factors in achieving quality diversity processes.

These data are relevant because the positive attitudes of future teachers are a pivotal factor in the achievement of inclusive education for students with educational needs (Saloviita, 2020). This is related to the effect that attitudes and teacher self-efficacy have on teachers’ intention toward inclusive education practices (Opoku et al., 2021; Diaz-Asto et al., 2022). Furthermore, when examining the model for the establishment of the inclusive teaching profile according to AEDENEEI (Donnelly and Watkins, 2011; European Agency for Development in Special Needs Education, 2012), it becomes evident that, in spite of the importance of knowledge and skills, attitudes are essential; without the right attitudes, inclusion cannot be achieved.

According to the data collected on the perception of inclusive education and the concept of students who fit into it, the majority of the sample agrees with inclusive approaches. However, 23% of the sample disagrees with the mainstreaming of students who use alternative and/or augmentative communication systems. This last finding is, for us, an area for improvement to be included in the MDSE training plans, given that on the one hand, the increasing number of students who benefit from these supports is a reality (for example, the number of students with Autism Spectrum Disorder increased by 8.07% in the 2020–2021 academic year according to the Confederation Autism Spain), and on the other hand, it has been demonstrated that negative attitudes can be transformed into positive attitudes (Garzón

TABLE 9 Differences between CEFIR-R dimensions and self-efficacy "Classroom management" using ANOVA.

CEFIR-R dimensions	Self-efficacy Classroom management	N	Mean	SD	F	p
D2. Methodology	Low	76	2.226	0.514	3.058	<0.001
	Medium	77	2.516	0.564		
	High	65	2.950	0.724		
	Total	218	2.545	0.665		
D3. Supports	Low	76	3.111	0.616	1.860	0.013
	Medium	77	3.087	0.543		
	High	65	3.323	0.570		
	Total	218	3.166	0.584		

TABLE 10 Differences between SACIE-R factors and self-efficacy "Instructional strategies" using ANOVA.

SACIE-R factors	Self-efficacy Instructional strategies	N	Mean	SD	F	p
F2. Feelings	Low	92	1.471	0.557	2.490	0.001
	Medium	55	1.375	0.480		
	High	71	1.356	0.612		
	Total	218	1.409	0.558		

TABLE 11 Differences between SACIE-R factors and self-efficacy "Classroom management" using ANOVA.

SACIE-R factors	Self-efficacy Classroom management	N	Mean	SD	F	p
F2. Feelings	Low	76	1.521	0.579	1.766	0.021
	Medium	77	1.389	0.450		
	High	65	1.302	0.628		
	Total	218	1.409	0.558		

Castro et al., 2016), a process that can be designed and implemented during initial teacher training; attitudes can be educated.

These results, which are ambivalent in certain areas, remind us that teachers face the daily conflict between inclusive values and their beliefs and abilities to implement them (Chiner, 2011), and are also related to the fact that, as Albulhamail et al. (2014) point out, teachers' attitudes are highly correlated with their level of training. In terms of training for inclusive teaching, less than 10% of the sample feels fully prepared for curricular and organizational responses in the classroom, with the lowest ratings related to daily teaching in the classroom: selecting and customizing objectives, competencies, and content, as well as developing measures and programs for diversity support.

These data, along with others presented, are related to the ideas of Boix (2008) who points out that future secondary teachers have a lack of knowledge about classroom diversity and the factors related to their success, such as methodological competences. Developing inclusive practices through instructional methods is, in this and other studies (Rodríguez Macayo et al., 2020; Pérez-Gutiérrez et al., 2021), the weak point of future teachers; they do not feel methodologically competent to work from an inclusive perspective and to respond to the needs of all students. Training deficiencies focus, as previously noted, on daily educational practice with a lack of knowledge and/or competencies on

how to develop learning situations, adapt teaching materials and evaluate students, among others; we must prioritize addressing these needs.

Regarding the conception of specialized support, the majority of the sample considers it important that this support teaching staff be incorporated into the regular classroom with each teacher understanding the need for joint planning between support teacher and regular teacher. These results seem to be related to the desired perception of support as a tool to increase the capacity to respond to diversity (Booth and Ainscow, 2011; Echeita et al., 2013) through joint planning that takes into account all students in the classroom. Additionally, these data are related to the profile of effective teaching and one of the educational practices with the greatest impact on student learning according to Visible Learning. Hattie (2017) concludes that collaborative work and collective efficacy of the teaching staff in a classroom has an effect size of 1.57 on student learning, which is considered a highly influential factor.

Finally, it should be noted that the feelings of the sample toward people with disabilities are positive, but the concerns of future teachers about having different students in the classroom are high. As previously mentioned, these concerns stem from a general sense of inadequacy in training to instruct students with disabilities and

provide proper attention to all students in a classroom. These findings are crucial to consider because, as noted by [Yada and Savolainen \(2017\)](#) and [Yada et al. \(2022\)](#), having positive and appropriate sentiments alone often is not sufficient to establish a firm commitment to implementing inclusive policies. These concerns frequently hinder the ability to manage challenging behaviors in the classroom and effectively teach. Addressing these concerns through initial training, equipping future teachers with the skills they require, can help foster a more positive attitude, higher self-efficacy, and a greater intention to implement inclusive practices in their teaching ([Miesera et al., 2019](#)).

Regarding training and preparation for inclusive teaching and attention to diversity, as mentioned earlier, we find in this research that a high percentage of the sample does not feel sufficiently trained. These data are added to the negative assessment made by the sample about the usefulness of training to respond in the classroom to the interests and concerns of students with educational support needs, also showing the perception of low impact of the training they have received on their future inclusive teaching practice. These data are consistent with others ([Boix, 2008](#)), such as those found by [Cachón Zagalaz et al. \(2015\)](#) who report that teachers have a negative perspective regarding the quality of education received within the MUPES.

It is worth noting that four out of five people surveyed in this study understand that a prerequisite for conducting a high-quality process of diversity attention is previous experience working with students who have educational needs. Perhaps this aspect should be considered in the development of competencies and the design of practices for future secondary school teachers, a process that is qualitatively and quantitatively different from that of teacher training in the stages of Primary and Infant Education. As various studies in our country have pointed out ([García Rodríguez et al., 2011](#)), it is necessary to understand the needs of teachers and provide training through the practicum period for the real context in which they will exercise their role. Additionally, as noted by [Rodríguez Marcos and Esteban Moreno \(2020\)](#), it is important to consider aspects such as carefully selecting schools that have innovative and inclusive teaching, organizational, and guidance practices that do not socialize future teachers in denigrated aspects of traditional teaching, among other proposals. Therefore, it is necessary to review the training systems and the keys of the teaching profession according to new social and cultural scenarios ([Vezub, 2007](#)).

However, what distinguishes the individuals in the sample who exhibit a greater sense of readiness for inclusive teaching in all its facets and who present a more favorable attitude toward 21st century education and diversity care in classrooms? What is the answer to the questions raised initially?

The first data that catches our attention is the fact that the group of women is the one that presents the best attitude, with significant results, toward diversity and inclusive schooling in the different instruments used. Additionally, they demonstrate greater awareness and openness about the significance of the involvement from the educational community and collaborative work among the different agents as a key element in the process of attending to diversity in inclusive schools.

Other notable results are found in the variable of habitual and close contact with persons in situations of vulnerability, which influences both being aware of and showing agreement on the elements that condition a quality education in the process of attending

to diversity in classrooms, as well as the perception of knowledge and capacity for development from the inclusive perspective of methodologies, resources, communication techniques, and evaluation that support said process. Furthermore, this variable presents a direct relationship with greater agreement on approaches aligned with the inclusive concept of supports. This contact with individuals with disabilities is a critical factor, as demonstrated in other studies ([Vera Noriega et al., 2002](#)), in terms of its impact on attitudes and perceptions of inclusion. However, we should also consider Allport's Intergroup Contact Theory, outlined in his work "The Nature of Prejudice" ([Allport, 1977](#)), which suggests that not all contact leads to changes in attitudes. Instead, the nature of the interaction between two groups influences whether the vulnerable group is socially accepted or rejected. Attitudes develop gradually through experiences, so it might be important for future teachers to not only expand their interactions with individuals with various educational needs, but also to be trained to perform their duties with everyone to generate positive experiences in relation to attending to diversity.

It is also worth mentioning the variable "experience in non-formal teaching to people in special vulnerable situations." This variable has the greatest positive impact on the inclusive conception of support, its definition and role in the classroom, as well as on the importance of collaborative work between the specialist teacher and the rest of the faculty. In addition, it is significantly related to the perception of being better prepared to respond to diversity in the classroom, with greater capacity for the development of methodologies, resources, and evaluation and teaching strategies, and with a lower level of concerns about participating in the education of different students in the classroom.

Regarding the impact of teacher self-efficacy on attitudes and training for diversity, the results have shown that it is a decisive factor to consider, and its arbitration is necessary for achieving tasks and goals in teachers, guiding the judgment that each teacher makes about their abilities and competencies to carry out the task of teaching in the current inclusive school. In the sample, teacher self-efficacy has a clear and positive impact on both attitudes and feelings toward people with educational needs; those who feel more self-efficacious in the development of "instructional strategies" to optimize their own instruction and in "classroom management" are those who have more positive feelings toward people with disabilities.

Furthermore, people in the sample with a higher sense of teacher self-efficacy show greater agreement with the elements that are necessary for the development of the diversity attention process in the classroom and also have higher scores in the dimension of knowledge and feeling of capacity for educational response to diversity. Additionally, people with higher levels of self-efficacy are more prepared for the development of methodologies, resources, materials, communication, and evaluation techniques in their day-to-day work in the classroom and show greater openness to participation and collaboration with all educational agents. Self-efficacy to optimize one's own academic instruction, "instructional strategies" and "commitment to students" appear to have the greatest impact on this assessment of community participation and collaborative work, positive attitudes and conception of diversity, and better perception of capacity for curricular and organizational response.

It seems relevant that future teachers perceive themselves as having greater mastery of instructional strategies than classroom management and interpersonal relationships. [Bandura \(1997\)](#) established that the sources of self-efficacy are vicarious experience,

previous experience, physiological and emotional states and verbal persuasion. The complex situations that may arise in the development of the inclusive education process, especially in relational aspects, will require previous experience, which the sample already identifies as relevant and whose importance is evidenced by the results of those who possess it, and for trainee teachers to have inclusive quality references and models in practice as a necessary step in developing higher levels of self-efficacy and personal and job satisfaction.

The results found are consistent with other studies and lead us to think that these future teachers who present high levels of self-efficacy, as noted by Tschannen-Moran and Hoy (2001), will present a higher level of effort in teaching, in achieving their own objectives and the highest expectations for all their students. In addition, they will be open to new approaches and practices and more motivated to experiment and meet the educational needs of all their students (Gibson and Dembo, 1984).

To highlight a final aspect, which requires further analysis with a biggest sample, related to whether students pursued the MDSE program in a public or private university. In terms of assessing the quality and effectiveness of the education received in inclusive curriculum and organizational response, diversity, methodology and resources, communication skills, and evaluation, as well as the anticipated impact of their training on their development as inclusive educators, it's the students who have graduated from a private university's who tend to perceive the training they received as more fitting, valid, and valuable.

Certain limitations identified during the execution of this study, as well as proposals for future research, include considering social desirability as a factor to be taken into account in the interpretation of the data. To address this, it would be beneficial to extend this research by expanding the sample and integrating open-ended questions or topics that can elucidate the perspectives of future teachers, thus yielding more insightful findings to craft practical and more accurate recommendations. Additionally, the sample size represents a limitation, and the objective is to augment it in at the service of a more comprehensive understanding of the subject under investigation. It is deemed essential to raise awareness and engage MDSE students in recognizing the significance of their participation in advancing of the educational system.

5 Conclusion

This research shows that although attitudes toward inclusive education and the students who create it are positive, they are not enough. The concerns that future teachers present and the training gaps to express the value they give to diversity and school in skills and abilities that empower them as inclusive teachers are important.

This situation is partly due to the perception of a limited influence of the training on their forthcoming progress as inclusive teachers, despite their interest in learning about diversity in all its facets. Most of them understand that the MDSE has not adequately trained them in these terms, feeling particularly concerned about issues related to day-to-day teaching, basic aspects of curriculum development and the teaching-learning and evaluation process in a school for all. However, those who perceive better training also have more positive attitudes toward diversity, and feel more competent as future inclusive teachers.

In this study, the importance of variables such as teacher self-efficacy, regular contact with individuals in situations of special vulnerability, and previous experience in non-formal education with them is evident. All of these have demonstrated their power in fostering more positive attitudes and achieving a higher level of training in addressing diversity in general, as well as in the development of methods and inclusive approaches in the classroom. This impact is not only on beliefs but also on the emotional and behavioral aspects of their attitude.

Therefore, programming and facilitating this positive contact, under the guidance of experienced teachers and individuals, seems necessary and key to progress, as inevitably the relationship in the educational act will occur sooner or later in their teaching development. Similarly, it is considered essential to promote formative experiences in non-formal settings and the sense of self-efficacy that always accompanies the foundation of inclusive values.

The practical implications of this study provide valuable insights for decision-making concerning future initiatives within the MDSE. These may include (1) promoting regular channels of communication and evaluation of training requisites, while concurrently addressing the evaluation of in-service teachers' needs for on-going enhancement of initial training; (2) appraising the integration of training requirements related to attention to diversity awareness in the specified domains is crucial, not only in this particular study but also in other relevant research. By addressing training needs in this area, the MSDE can ensure that future educators are equipped with the knowledge, skills, and attitudes necessary to effectively support attention to diversity in the classrooms; (3) providing strategies and guidelines for reducing disruptive behaviors, conflict resolution and improved coexistence from initial training; (4) considering the practice component of the MDSE program, it is important to take into account the selection of schools following criteria suggested by research in the field. This approach guarantees the promotion of inclusion principles and educational innovation; (5) In addition, promoting self-efficacy in practice centers by fostering collaboration and a positive climate that facilitates coordination and mutual aid, and promotes the observation and reflection on the teaching practice; and (6) facilitating future teachers' contact with the students facing the greatest vulnerabilities, is crucial for their development as inclusive educators. This can be achieved through collaborative engagement of Service-Learning Projects (SLPs) in the MDSE subjects, as well as in prior degrees. These actions can provide a way for students to engage in meaningful community experiences that foster both personal growth and the dissemination of knowledge as catalysts for social transformation. The University Network Association for Service-Learning (ApS) offers a valuable and representative platform to facilitate this process, given its extensive experience and established reputation within the university community.

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.

Ethics statement

The Bioethics Committee of the University of Burgos approved the research (Reference UBU 032/2021), respecting all the requirements established in the Declaration of Helsinki of 1975. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

MA-P, SV, JF-S, and JG-B participated in the conception, design of the study, acquiring the data, performing the statistical analysis, writing the manuscript, and revising it. All authors have made a substantial, direct and intellectual contribution to the work, and have approved it for publication.

References

- Ainscow, M., Dyson, A., Goldrick, S., and West, M. (2012). *Developing equitable education systems*; Routledge London, UK. (2012). 35.
- Ainscow, M., Dyson, A., Goldrick, S., and West, M. (2013). Promoviendo la equidad en educación. *Rev. Investig. Educ.* 11, 44–56.
- Alba Pastor, C. *Diseño universal Para el Aprendizaje: Educación Para Todos y Prácticas de Enseñanza Inclusivas*. Morata: Madrid. (2018).
- Albulhamail, S., Al-Sulami, F., Alnouri, M., Mahrous, N., Joharji, D., Albogami, M., et al. (2014). Primary school teacher's knowledge and attitudes towards children with epilepsy. *Seizure European J. Epilepsy* 23, 280–283. doi: 10.1016/j.seizure.2013.12.010
- Allport, G. W. (1977). *La Naturaleza del Prejuicio, 5th Edn*. Editorial Universitaria de Buenos Aires: Buenos Aires, Argentina.
- Arias Pastor, M., Van Vaerenbergh, S., Fernández Solana, J., and González Bernal, J. J. (2023). Secondary education teacher training and emotional intelligence: ingredients for attention to diversity in an inclusive School for all. *Educ. Sci.* 13:519. doi: 10.3390/educsci13050519
- Armando, S., Alfaro, M., Arias, S., and Gamba, Y. (2019). Agenda 2030. Claves Para la transformación sostenible. Madrid: Catarata. 318 pp. *Relac. Int.* 46, 207–210.
- Bandura, A. (1997). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037//0033-295x.84.2.191
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., and Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema* 29, 552–557. doi: 10.7334/psicothema2016.383
- Blotnicky-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
- BOE. (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se Modifica la ley Orgánica 2/2006, de 3 de mayo, de Educación. BOE-A2020-17264. Available at: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17264 (Accessed August 30, 2023).
- Boix, J. L. (2008). Nuevo profesorado de educación secundaria Para nuevas realidades. Formación inicial y periodo umbral. *Rev. D'innovació I Recer. Educ.* 1, 31–50.
- Booth, T., and Ainscow, M. *Index for inclusion. Developing learning and participation in schools. 3rd Edn*. Bristol: CSIE. (2011).
- Booth, T., Simón Rueda, C., Sandoval Mena, M., Echeita, G., and Muñoz Martínez, Y. (2015). Guía Para la Educación Inclusiva. Promoviendo el aprendizaje y la participación en las escuelas: Nueva edición revisada y ampliada. REICE Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación. Available at: https://repositorio.uam.es/bitstream/handle/10486/668285/REICE_13_3_1.pdf?sequence=1 (Accessed August 30, 2023).
- Cachón Zagalaz, J., López Manrique, I., Romero Granados, S., Zagalaz Sánchez, M. L., González, G., and de Mesa, C. (2015). Opinión de docentes y estudiantes del máster de secundaria sobre las aportaciones de este a la formación del profesorado, la calidad docente y los intereses personales. *Magister* 27, 1–10. doi: 10.1016/j.magis.2015.03.001
- Chiner, E. *Las percepciones y actitudes del profesorado hacia la inclusión del alumnado con necesidades educativas especiales Como indicadores del uso de prácticas educativas inclusivas en el aula*. Alicante: Universidad de Alicante. (2011).
- Collado-Sanchis, A., Tárraga-Mínguez, R., Lacruz-Pérez, I., and Sanz-Cervera, P. (2020). Análisis de actitudes y autoeficacia percibida del profesorado ante la educación inclusiva. *Educar* 56, 509–523. doi: 10.5565/rev/educar.1117
- Colmenero Ruiz, M. J., and Pegalajar Palomino, M. C. (2015). Cuestionario Para futuros docentes de Educación Secundaria acerca de las percepciones sobre atención a la diversidad: Construcción y validación del instrumento. *ESE Estud Sobre Educ.* 29, 165–189. doi: 10.15581/004.29.165-189
- Díaz-Asto, Y., Rivera-Arellano, G., and Vega-Gonzales, E. (2022). Sentimientos, actitudes y preocupaciones hacia la inclusión en docentes de Lima Metropolitana. *CienciAmérica* 11, 74–86. doi: 10.33210/ca.v11i1.384
- Donnelly, V., and Watkins, A. (2011). Teacher education for inclusion in Europe. *Prospects* 41, 341–353. doi: 10.1007/s11125-011-9199-1
- Duk, C., and Murillo, F. J. T. (2020). Right to education is the right to an inclusive and equitable education. *Rev. Latinoam. Educ. Inclusiva* 14, 11–13. doi: 10.4067/s0718-73782020000200011
- Durán, D., and Giné, C. (2017). La formación del profesorado Para la educación inclusiva: un proceso de desarrollo profesional y de mejora de los centros Para atender la diversidad. *Revista Latinoamericana de Educación Inclusiva* 1, 153–170.
- Echeita, G., Simón, C., Sandoval, M., and Monarca, H. *Cómo fomentar las redes naturales de apoyo en el Marco de una escuela inclusiva: Propuestas prácticas*. Sevilla: Eduforma. (2013).
- European Agency for Development in Special Needs Education. *Teacher education for inclusion across Europe—Challenges and opportunities*. European Agency for Development in Special Needs Education: Odense, Denmark. (2011).
- European Agency for Development in Special Needs Education. *Profile of inclusive teachers*. European Agency for Development in Special Needs Education: Odense, Denmark. (2012).
- Forlin, C., Earle, C., Loreman, T., and Sharma, U. (2011). The sentiments, attitudes, and concerns about inclusive education revised (SACIE-R) scale for measuring pre-service teachers' perceptions about inclusion. *Except. Educ. Int.* 21, 50–65. doi: 10.5206/eei.v21i3.7682
- García Rodríguez, M. S., Pascual Sevillano, M. A., and Fombona Cadavieco, J. (2011). Las competencias en el prácticum del Máster en Formación del Profesorado de Educación Secundaria Obligatoria, Bachillerato y Formación Profesional. *Magister Rev. Miscelánea Investig.* 24, 109–119.
- Garcón Castro, P., Calvo Álvarez, M. I., and Orgaz Baz, M. B. (2016). Inclusión educativa. Actitudes y estrategias del profesorado. *Rev. Esp. Discapac.* 4, 25–45. doi: 10.5569/2340-5104.04.02.02
- Gibson, S., and Dembo, M. H. (1984). Teacher efficacy: a construct validation. *J. Educ. Psychol.* 76, 569–582. doi: 10.1037/0022-0663.76.4.569

Funding

SV was supported by PID2022-136246NB-I00 grant from Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación.

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.

- González-Gil, F., Martín-Pastor, E., and Orgaz Baz, B. (2017). ¿Están los futuros profesores formados en inclusión?: Validación de un cuestionario de evaluación. *Aula Abierta* 46, 33–40. doi: 10.17811/rife.46.2.2017.33-40
- Hattie, J. *Aprendizaje visible Para Profesores: Maximizando el Impacto en el Aprendizaje*. Parainfo: Madrid, Spain. (2017).
- Herrera-Seda, C. (2018). La Formación Inicial del Profesorado Para una Educación Inclusiva: Desafíos, Oportunidades y Transformaciones. *Revista latinoamericana de educación inclusiva* 12, 17–20. doi: 10.4067/S0718-73782018000200017
- Holdheide, L. R., and Reschly, D. J. *Teacher preparation to deliver inclusive services to students with disabilities. TQ Connection Issue Paper*. (2008). Washington, DC: National Comprehensive Center for Teacher Quality.
- Imbernón, F. (2019). La formación del profesorado de educación secundaria: la eterna pesadilla. *Profesorado. Revista de Currículum y Formación de Profesorado*. 23, 151–163. doi: 10.30827/profesorado.v23i3.9302
- Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., et al. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemp. Educ. Psychol.* 34, 67–76. doi: 10.1016/j.cedpsych.2008.08.001
- Mérida-López, S., and Extremera, N. (2020). When pre-service teachers' lack of occupational commitment is not enough to explain intention to quit: Emotional intelligence matters! *Rev. Psicodidáctica* 25, 52–58. doi: 10.1016/j.psicod.2019.05.001
- Mérida-López, S., Extremera, N., Sánchez-Álvarez, N., Quintana-Orts, C., and Rey, L. (2018). ¿Crees que serás buen docente? Propiedades psicométricas de la adaptación al castellano del teachers' sense of efficacy scale short form en profesores noveles. Póster presentado en el II Congreso Internacional de Psicología del Trabajo y RRHH/VI Congreso Nacional, Valencia 1–3 de marzo 2018.
- Miesera, S., DeVries, J. M., Jungjohann, J., and Gebhardt, M. (2019). Correlation between attitudes, concerns, self-efficacy and teaching intentions in inclusive education evidence from German pre-service teachers using international scales. *J. Res. Spec. Educ. Needs* 19, 103–114. doi: 10.1111/1471-3802.12432
- ONU. (2014) Reunión mundial sobre la educación Para todos: Declaración final de la reunión mundial sobre la EPT de 2014 el Acuerdo de Mascate. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000228122_spa?posInSet=1&queryId=72256f15-e90f-474d-aa88-b23a2efe668f (Accessed August 30, 2023).
- Opoku, M. P., Cuskelly, M., Pedersen, S. J., and Rayner, C. S. (2021). Attitudes and self-efficacy as significant predictors of intention of secondary school teachers towards the implementation of inclusive education in Ghana. *Eur. J. Psychol. Educ.* 36, 673–691. doi: 10.1007/s10212-020-00490-5
- Pegalajar Palomino, M. C., and Colmenero Ruiz, M. J. (2017). Actitudes y formación docente hacia la inclusión en Educación Secundaria Obligatoria. *Rev. Electrónica Invest. Educ* 19, 84–97. doi: 10.24320/redie.2017.19.1.765
- Pérez-Gutiérrez, R., Casado-Muñoz, R., and Rodríguez-Conde, M. J. (2021). Evolución del profesorado de apoyo hacia la educación inclusiva: una perspectiva legislativa autonómica en España. *Rev. Complut Educ.* 32, 285–295. doi: 10.5209/rced.68357
- Prieto Navarro, L. (2002). El análisis de las creencias de autoeficacia: un avance hacia el desarrollo profesional del docente. *Revista de Teología y Ciencias Humanas*. 60, 591–612.
- Rodríguez, A., Caurcel, M. J., and Alai, L. (2019). Medir actitudes profesoras españolas hacia la educación inclusiva Para lograr una escuela Para todos. *Actas Icono* 14, 416–433.
- Rodríguez Fuentes, A., Gallego Ortega, J. L., Navarro Rincón, A., and Caurcel Cara, M. J. (2021). Perspectivas actitudinales de docentes en ejercicio y en formación hacia la inclusión educativa. *Psicoperspectivas Individuo y. Sociedad* 20:1892. doi: 10.5027/psicoperspectivas-Vol20-Issue1-fulltext-1892
- Rodríguez, H. J. (2019). La formación inicial del profesorado Para la inclusión. Un urgente desafío que es necesario atender. *Publicaciones* 49, 211–225. doi: 10.30827/publicaciones.v49i3.11410
- Rodríguez Macayo, E., González Gil, F., Pastor, E., and Vidal Espinoza, R. (2020). Validación De Un cuestionario sobre la actitud docente frente a la educación inclusiva en Chile. *Foro. Educ.* 35, 63–86. doi: 10.29344/07180772.35.2650
- Rodríguez Marcos, A., and Esteban Moreno, R. S. *El Prácticum Factor de Calidad en la Formación del Profesorado de Secundaria y Bachillerato: Teoría y Práctica*. Octaedro: Barcelona, Spain. (2020).
- Ruiz Rodríguez, E. (2019). Diseño universal Para el aprendizaje: estrategias Para un aprendizaje Para todos. *Revista Síndrome de Down* 140, 11–22.
- Sales Ciges, A. (2006). La formación inicial del profesorado ante la diversidad: Una propuesta metodológica para el nuevo Espacio Europeo de Educación Superior. *Rev Interuniv Form Profr.* 20, 201–217.
- Saloviita, T. (2020). Attitudes of teachers towards inclusive education in Finland. *Scand. J. Educ. Res.* 64, 270–282. doi: 10.1080/00313831.2018.1541819
- Sandoval, M., López, M., Miquel, E., Durán, D., Giné, C., and Echeita, G. (2013). Index for Inclusion: Una guía Para la evaluación y mejora de la educación inclusiva. *Contextos educativos* 5, 227–238. doi: 10.18172/con.514
- Sucaticona, G. *Actitud del Docente Hacia La Inclusión en maestros con y sin Alumnos con Necesidades Educativas Especiales de las Instituciones Educativas Estatales de la UGEL. Ph.D. Thesis*. Universidad Ricardo Palma, Santiago de Surco, Perú, (2016).
- Tschannen-Moran, M., and Hoy, A. W. (2001). Teacher efficacy: capturing an elusive construct. *Teach. Teach. Educ.* 17, 783–805. doi: 10.1016/s0742-051x(01)00036-1
- UNESCO. (2005). Guidelines for Inclusion: Ensuring Access to Education for All. Available at: [http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/\[in=epidoc1.in\]/?t2000=021573/\(100\)](http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=021573/(100)) (Accessed August 30, 2023).
- UNESCO. (2020). Equipo del Informe de Seguimiento de la Educación en el Mundo. Inclusión y educación: Todos y todas sin excepción. Informe de seguimiento de la educación en el mundo. Organización de las Naciones Unidas Para la Educación, la Ciencia y la Cultura. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000374817> (Accessed August 30, 2023).
- UNESCO. (2022). UNESCO Transforming education from within current trends in the status and development of teachers. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000383002> (Accessed August 30, 2023).
- Vera Noriega, J. A., Burrueal Valencia, M. A., and Sainz Palafox, M. A. (2002). Actitudes hacia la discapacidad y su influencia en la percepción de inclusión en estudiantes de educación superior. *Rev. Educ. Inclusiva* 15, 45–62.
- Vezub, L. F. (2007). La formación y el desarrollo profesional docente frente a los nuevos desafíos de la escolaridad Profesorado. *Revista de Currículum y Formación de Profesorado* 11, 1–23.
- Yada, A., Leskinen, M., Savolainen, H., and Schwab, S. (2022). Meta-analysis of the relationship between teachers' self-efficacy and attitudes toward inclusive education. *Teach. Teach. Educ.* 109:103521. doi: 10.1016/j.tate.2021.103521
- Yada, A., and Savolainen, H. (2017). Japanese in-service teachers' attitudes toward inclusive education and self-efficacy for inclusive practices. *Teach. Teach. Educ.* 64, 222–229. doi: 10.1016/j.tate.2017.02.005



OPEN ACCESS

EDITED BY

Luisa Losada Puente,
University of A Coruña, Spain

REVIEWED BY

Daniel Bores-García,
Rey Juan Carlos University, Spain
Jordi Colomer,
University of Girona, Spain

*CORRESPONDENCE

Dong Li

✉ lidong58999@163.com

RECEIVED 23 September 2023

ACCEPTED 13 November 2023

PUBLISHED 28 November 2023

CITATION

Zhou T, Wang H and Li D (2023) Focusing on the value of cooperative learning in physical education: a bibliometric analysis. *Front. Psychol.* 14:1300986. doi: 10.3389/fpsyg.2023.1300986

COPYRIGHT

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

Focusing on the value of cooperative learning in physical education: a bibliometric analysis

Tong Zhou¹, Huayi Wang² and Dong Li^{3*}

¹Department of Physical Education, College of Education, Korea University, Seoul, Republic of Korea,

²Department of Education, Kyungil University, Gyeongsan, Republic of Korea, ³Department of International Culture Education, Chodang University, Muan-gun, Republic of Korea

The shift toward cooperative learning has highlighted the growing advantages of individual learning modes during the transition. Nevertheless, a systematic compilation of the precise classification and developmental dynamics of cooperative learning in PE has been absent. This study aimed to organize the existing progress and significance of collaborative learning. The study entailed a meticulous systematic review process, examining 169 articles in this domain with the aid of visualization software. The results of the study indicate that the overall use of cooperative learning in physical education is on the rise and will reach its highest level in 2021; Second, the keywords, major core scholars, journals, countries, and major research topics; the visual knowledge map reveals the major research topics of intrinsic motivation, cooperative learning, motor skills, self-learning, written expression, and pedagogical models. The research primarily centers on primary and secondary education, followed by teacher training and higher education. At the primary and secondary school levels, there is a specific focus on aspects such as motivation, teacher-student relationships, and the group atmosphere. This research also explores sustainable development and training for PE teachers, model integration, and its influence on students' intrinsic motivation; and finally, the future directions of cooperative learning in PE are summarized. This study provides meaningful and valuable information on how cooperative learning models can be used and developed in various teaching and learning environments, physical education teacher education, and overall student development.

KEYWORDS

physical education, cooperative learning, teacher training, teaching method, bibliometric analysis

1 Introduction

Cooperative learning (CL) has been a topic of interest in the field of education, particularly in physical education (PE) (Dyson and Casey, 2016; Casey and Quennerstedt, 2020; Rivera-Pérez et al., 2021). With the advancement of the Fourth Industrial Revolution (4IR) (Oke and Fernandes, 2020) in the education sector, artificial intelligence presents an undeniable challenge. Along with this, there is a more sophisticated social division of work (Janoski, 2015), team collaboration, and collaborative learning (Peduzzi et al., 2020) in teaching and training, also known as cooperative learning. Stanne et al. (1999) found that cooperative learning is more effective in promoting learners' movement skill levels (with an effect size of 0.53) than individual effort or competition (with only 0.36) in a meta-analysis of 64 studies. Moreover, PE has been

seeking education methods suitable for both social development and student needs in its development over the years. From its development in the 1980s to the present, the 21st-century cooperative learning model has been widely applied and promoted in the education field. It contains five key factors: positive interdependence, individual accountability, group processing, active interaction, and social skills (Johnson and Johnson, 2014). Studies have shown that in PE, CL mainly focuses on the interactions between students and between students and teachers (Metzler, 2017).

CL emphasizes the development of students, including physical, cognitive, social, and emotional aspects, which has been fully reflected in PE. According to Casey and Goodyear's (2015) systematic analysis based on a qualitative review of 14 articles, quantitative analysis of 11 articles, and mixed-method analysis of 2 articles, CL can promote students' positivity, listening skills, understanding and encouragement of others, and the ability to build respect, understanding, mutual encouragement, and complete learning tasks together in PE classrooms (Silva et al., 2021a). Most importantly, CL is implemented in a student-centered learning mode, which emphasizes students' subjectivity (Li and Lam, 2013).

On the other hand, CL in PE can enhance students' teamwork abilities (Dunn and Wilson, 1991). For example, Sánchez-Hernández et al. (2018) intervened with CL in soccer games, allowing male and female students to participate together by grouping them into teams. This reduced gender discrimination and provided more opportunities for female students to participate. Peer support and collaboration improved physical fitness and reduced gender bias. Furthermore, CL, as an intervention method, can reduce gender differences in the acquisition of motor (Ortiz et al., 2023). The research also reveals that CL can effectively diminish the impacts of gender disparities and inequalities. It promotes the principles and capabilities of sustainable development education among group members, particularly in addressing fairness and inclusivity concerns in early childhood and primary education (Cañabate et al., 2021). In terms of enhancing students' cognitive and emotional development, CL not only cultivates students' level of sports skills but also enables them to identify their weaknesses and learn better by observing others (Barrett, 2005). Through the process of promoting interaction and communication among students, CL also helps to foster good friendships and establish positive relationships among students (Polvi and Telama, 2000).

From the perspective of teachers, CL shifts some of the power from the teacher to the student-centered approach (Silva et al., 2021b). This shift in focus poses a challenge for both pre-service and in-service teachers, breaking the traditional model of PE instruction (Casey et al., 2009). Under this teaching model, there are many uncertain factors and challenges in the implementation of CL (Casey et al., 2015). Therefore, in the process of sustainable professional development for PE teachers, it is important to consider how to implement CL, as well as classroom responses during the implementation process (Casey et al., 2015). For future PE teachers, pre-employment teacher training is needed to enhance the effectiveness of CL implementation (Silva et al., 2021a).

This study offers a comprehensive overview of the development of CL in PE, with a focus on the physical and mental growth of students and the professional competencies of PE teachers. It aims to provide insights into the current research findings and trends in CL through a bibliometric visual analysis, and it is the first bibliometric visual analysis of CL in PE that complements the literature review. By

utilizing CiteSpace and VOSviewer visualization software, this research analyzes the knowledge map based on relevant research in the field of CL, to provide some valuable information on the use of CL in physical education and on the professional development of PE teachers.

2 Literature review

2.1 CL model

CL is a beneficial teaching model that helps to avoid negative competition among individuals. Individualistic learning constantly faces challenges, and therefore, peer-assisted learning and socialization relationships have gradually gained importance in the learning process (Hartup, 1976). In the process of promoting comprehensive student development, physical, cognitive, social, and emotional learning problems become crucial issues that need to be effectively solved in PE (Metzler, 2017). In the teaching and learning process, the aim is to achieve classroom learning goals, and in an ideal classroom, all students can learn how to collaborate with others and how to compete for fun and enjoyment, striving to achieve common learning goals (Johnson et al., 2014). Therefore, in CL, individuals seek to achieve favorable outcomes for themselves or the whole group while improving their own efficiency and that of others, resulting in a situation of mutual assistance and win-win outcomes. This situation is also a way to promote cooperation and create a better future in the 21st century.

2.2 Application of CL

The development of PE has often emphasized sports, performance, and achievements, while neglecting other factors and the true needs of students (Bailey et al., 2009). Research shows that CL research is more concentrated in secondary education, for primary and secondary schools to carry out CL in the student motivation, the group is divided into and teacher-student interaction can better promote the learning effect of primary and secondary school students (Borres-Garcia et al., 2021). CL can effectively enhance children's or adolescents' participation, socialization with peers, and meet their physical and psychological needs (Jung and Choi, 2016). By encouraging mutual dependence and support among students, they can collectively achieve team honor and academic success (Er and AtaÇ, 2014), leading to improved academic performance (Da Luz, 2015). In such a CL environment, a safe and motivating atmosphere can be created for students. Particularly in PE, the CL is especially important as it helps students avoid being passive observers, which can often occur in PE classes divided into practice and theory units.

Currently, CL has a positive impact on students' behavioral motivation, from elementary to high school and university sports cooperation courses, and affects students' cognitive, physical, and emotional performance (Liu and Lipowski, 2021; Rivera-Pérez et al., 2021). In addition, using CL can improve teachers' satisfaction with the classroom (Gil-Arias et al., 2022), enhance teaching quality, evaluation (Lopez et al., 2008), and promote curriculum reform and optimization under the background of technology and information (Fernández Río, 2017; Kang and Kang, 2019), the multimodal

approaches to math and PE within CL is beneficial in fostering the academic and social attitudes of pre-service teachers (Bassachs et al., 2022), and promote sustainable professional development of PE teachers (Goodyear, 2017). Especially with the development of intelligent systems (Yoda, 2017), networks, and modern scientific and technological advances (Jastrow et al., 2022), CL can be combined ingeniously, adapting to different situations and promoting the comprehensive development of students.

2.3 Value of bibliometrics

The application of bibliometrics in scientific research has been widely recognized, as it provides a useful tool for scholars to gain insights into current research topics and expand their knowledge base (Hood and Wilson, 2001). With the ability to quickly sift through and cluster large volumes of literature, bibliometric tools such as CiteSpace and VOSviewer not only accelerate the production and dissemination of information, but also improve the utilization and visualization of information (Chen et al., 2017). Bibliometric analysis can help researchers by providing a quantitative approach and can review the existing literature in a particular field, helping researchers to systematically understand the current state of the field and future trends (Zhou et al., 2023a). Bibliometrics is widely used in independent fields and also in cross-disciplinary fields, contributing to the development of multidisciplinary intersections (Song and Wang, 2020). In addition, based on the development of modern computer technology, the results of this research field are presented visually with graphical and visual results that can further complement the literature analysis (Donthu et al., 2021). Bibliometrics can be valuable for educational purposes as it offers learners intuitive and visual information about research in a specific field. It assists learners in locating key literature more efficiently and comprehending important information about a particular field or themes (Huang et al., 2020). Additionally, bibliometrics can enhance the quality and speed of scientific research.

Additionally, these tools, in combination with computer program processing capabilities, can assist scholars in organizing, summarizing, and analyzing specific research topics. Through knowledge maps, detailed information about publications, keywords, authors, countries, institutions, journals, and even co-citation relationships can be intuitively presented. As of March 2023, bibliometrics has been widely applied in fields such as environment (Mao et al., 2018), chemistry (Shen et al., 2022), education and intelligent learning (Chen et al., 2021), mental health (Akintunde et al., 2021), and sports medicine (Chen and Shin, 2021). However, there has been limited application of bibliometrics in sports education, and no scholars have conducted a visual analysis of CL.

Therefore, this study aims to conduct a bibliometric analysis of the development of cooperative learning in sports education. The objectives of this study are:

- 1) To determine the extent of publications, primary literature, research topics, keywords, authors, and journals in the field.
- 2) To identify collaborative networks and partnerships among authors and countries.
- 3) To analyze the information presented in the visual knowledge map.

- 4) To comprehend the progress of CL in PE and offer future prospects.

3 Method

3.1 Database search

This section consists of three main steps, namely establishing a database of total records, analyzing the organized data using visualization software to create a knowledge map and evolution analysis. Following the PRISMA (Preferred Reporting Items for Systematic Summaries and Meta-Analyses) flow chart paradigm for literature filtering, the next step was data cleaning (Page et al., 2021). To achieve this goal, four main steps were followed (Figure 1).

The first step involved clarifying the research purpose, which is to introduce the research achievements of cooperative learning in the field of physical education using bibliometric analysis.

The second step involved selecting the literature database. Web of Science (WOS) Core Collection was chosen as the main literature retrieval database for bibliometric analysis due to its high-quality research data, large audience, and long history (Mongeon and Paul-Hus, 2016). The advanced search function of WOSCC was used to conduct a search with the following keywords: TS=((“Cooperative Learning” AND “Physical education”) OR (“Collaborative Learning” AND “Physical Education”)) (Bores-Garcia et al., 2021). To avoid omissions, a time span was not selected, and instead, the time period was set to end in 2023. The complete records and all relevant information related to the articles cited in the database were then exported. CiteSpace and VOSviewer were used as visualization software to conduct bibliometric analysis. These two software packages can complement each other, thus improving the quality of the visualization knowledge map (Zhou, 2023).

The third step involved data cleaning of the obtained literature. Only articles were selected, duplicates were removed, and withdrawn papers were excluded. Additionally, papers that contained irrelevant titles, abstracts, and keywords were also removed. Based on previous studies, VOSviewer version 1.6.16 was used to generate the results. After the final filtering and screening, a total of 167 paper documents related to cooperative learning in physical education were included in the WOS Core Collection database. These papers were published in 69 publications, originating from 36 countries and 185 institutions, and authored by 380 authors. The literature sources were mainly from education, sports, and information technology.

3.2 Data analysis

Bibliometric analysis covers many factors such as structural presentation, dynamic changes, evaluation, prediction, and scientific measurement in the current field, and the methods adopted are more systematic. In addition, both CiteSpace (Chen, 2006) and VOSviewer (Van Eck and Waltman, 2010) were used in the third step as commonly used software tools for creating literature networks (such as keyword cluster analysis, author analysis, and country analysis) and shared citation information among literature (such as shared cited authors and shared cited literature).

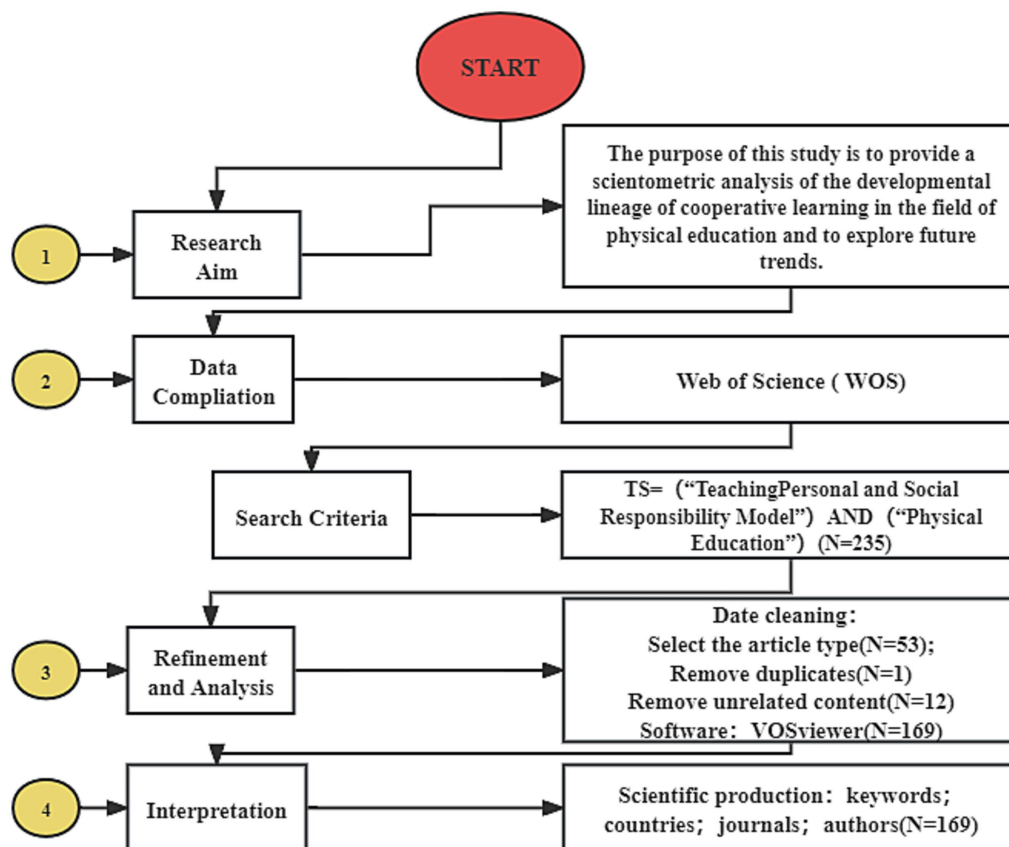


FIGURE 1
Main steps of the study process. Adapted from Leal et al. (2022).

Finally, all the results, including keywords, authors, countries, and publications, were presented in tables and visualization knowledge maps through data analysis.

4 Results

4.1 Dynamic trends in CL

From Figure 2 it can be seen that CL has been applied in the field of education for a long time, but it was not until 1997 that it began to be used in PE. The previous literature search found that it was about social and moral education for high-risk children, that is, moral education in the process of human development (Miller et al., 1997), and it was an empirical study. From 1997 to 2023, the trend showed a slow increase to rapid growth. Spain has the highest number of publications on cooperative learning in physical education, with 71 papers cited up to 404 times, indicating a significant amount of research in this area. The second highest is the UK, with 21 papers published and cited up to 1,842 times, mostly empirical studies that explore current teaching methods. The USA has published 20 papers and cited up to 928 times, also mostly empirical studies. China and France had 13 and 8 publications, respectively (Table 1).

In the research findings, it was observed that among the literature primarily centered on students, studies utilizing Collaborative Learning (CL) were conducted with children ($N=4$, 3.1%), at the

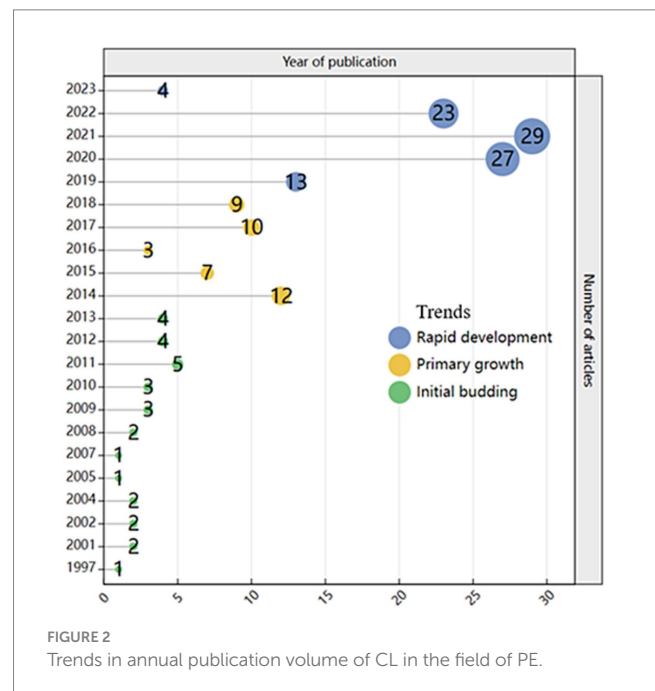
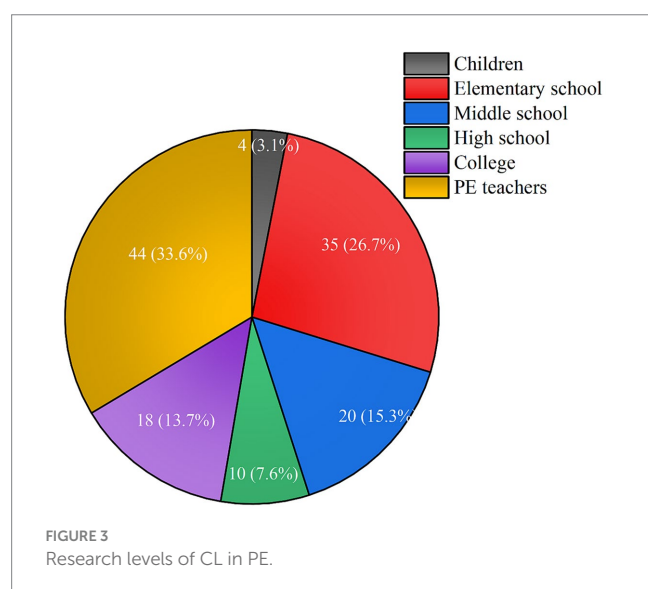


FIGURE 2
Trends in annual publication volume of CL in the field of PE.

elementary school level ($N=35$, 26.7%), middle school level ($N=20$, 15.3%), high school level ($N=10$, 7.6%), and university level ($N=18$, 13.7%). Literature primarily focused on teachers as the subjects of

TABLE 1 Status of publication volume in the 10 most relevant countries.

Rank	Country	Documents	Citations	Total link strength
1	Spain	71	404	5
2	England	21	1,842	12
3	USA	20	928	8
4	People's R China	13	23	1
5	France	8	124	0
6	Ireland	7	462	7
7	New Zealand	6	222	6
8	Sweden	6	98	2
9	Turkey	6	52	0
10	Norway	5	20	1



research accounted for 44 articles, representing 33.6% of the total. In terms of the application of CL in PE, there is a greater concentration of research aimed at primary and secondary education, as well as teacher-centered studies (Figure 3).

4.2 Keyword knowledge graph analysis

Using CiteSpace to perform cluster analysis and centrality calculation of keywords in the current field (Figure 4), we can see that the clustering is mainly centered around “physical education” and “cooperative learning.” The size of the circles represents the frequency of appearance of the keywords, and the outermost circle color represents frequent appearance in recent years. In terms of the application of CL in PE, it is mostly carried out through the relationship between teachers and students. Therefore, the centrality of “teacher” and “student” as keywords reaches 0.03 and 0.25, respectively, and the centrality of “student” is 0.22 higher than that of “teacher,” which is consistent with the goal of CL mainly targeting students for teaching (Li and Lam, 2013). Moreover, as a model, CL is

also constantly exploring the factors that may affect it. Therefore, the centrality of “model” and “educational model” as keywords is 0.04 and 0.05, respectively.

In Figure 5, six cluster themes related to research are identified, which include intrinsic motivation, cooperative learning, motor skills, self-learning, written expression, and educational models. The explosive period of keywords in the application of CL in PE was visually analyzed, and it was found that the keyword “implementation” had a relatively long duration from 2013 to 2019, lasting for 6 years, which was a breakthrough in empirical research on the application of CL in the 21st century. Its impact strength value was 4.08, and it has had a significant amount of research on the roles and influences of students and teachers, which has continued from 2020 to 2023. “Motivation” had a strength value of 2.96 and has positive implications for intrinsic motivation, behavioral motivation, and student behavior in the classroom. “Teacher” had a strength value of 2.7, and “intervention” had a strength value of 1.96. These four main keywords have continued research possibilities in 2023, and they are related to COVID-19, different PE teaching modes (Colao et al., 2020), online CL (Ivone et al., 2020), blended education (Hamzah et al., 2022), etc. in Figure 6. The use of CL in different scenarios is a challenge and discovery that needs to be continuously explored in the future.

4.3 Author knowledge graph analysis

Table 2 provides an analysis of the authors, focusing on the top five authors in terms of publication volume and relevance to the main topics. It was found that three of the authors are affiliated with institutions in Spain, with publication volumes of 19, 14, and 5, and average citation rates of 11.0, 68.9, and 5.0, respectively. Spanish scholars are leading the way in research related to CL in PE, with Goodyear, Victoria A. having published 7 articles with an average citation rate of 74.1. Dyson, Ben from the United States has published 5 articles with an average citation rate of 21.6. This is consistent with the country ranking in Table 1, where Spanish scholars account for about half of the total publications. Spanish scholars have conducted a lot of empirical research in this field. It was found that their research has a significant impact on students’ classroom satisfaction with teacher CL training and attitudes (Martínez Benito and Sánchez Sánchez, 2020). They have also explored to varying degrees the integration of CL with other sports models, such as PE and game-based learning models (Chiva-Bartoll et al., 2018; Evangelio et al., 2021).

4.4 Visualization analysis of literature and journals

Through the visualization analysis of the knowledge graph of literature (Figure 7), the size of the circle represents the frequency of citation of the literature, and the color change indicates the publication year. Ntoumanis (2001) has the highest citation frequency for a single paper, which empirically analyzes the perception of movement ability and the mobilization of students’ intrinsic motivation, emphasizing their importance in PE classes. According to Dyson et al. (2004), who are the second most cited literature, they have provided a deeper understanding of the model of PE. They emphasize that teaching

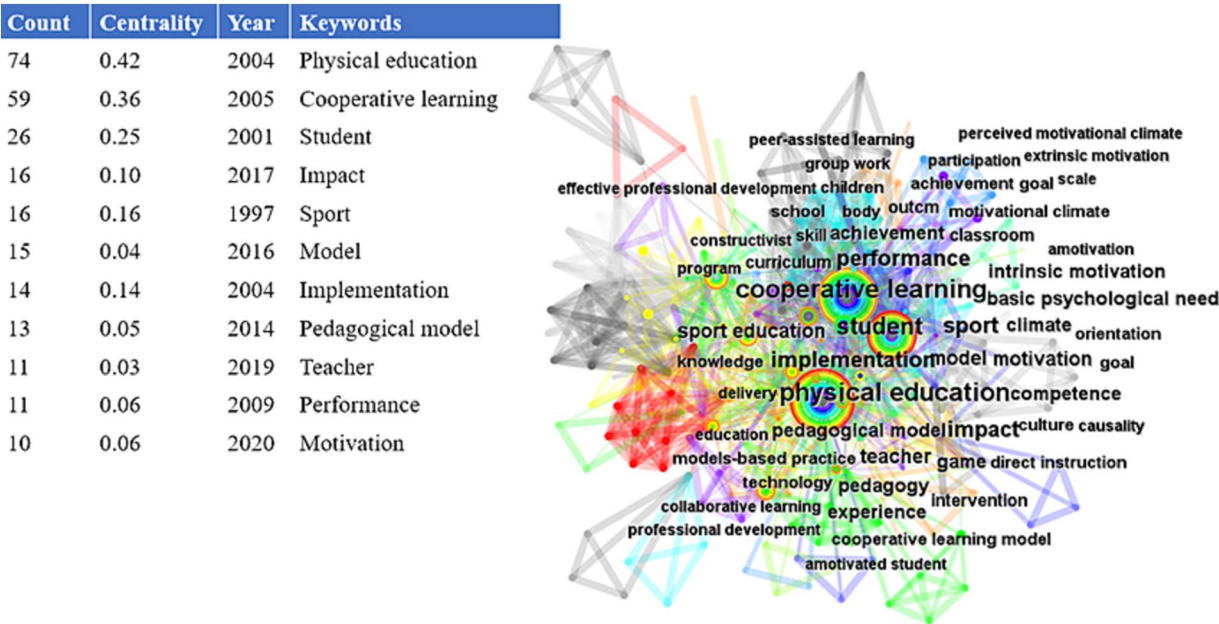


FIGURE 4
Keyword visualization analysis.

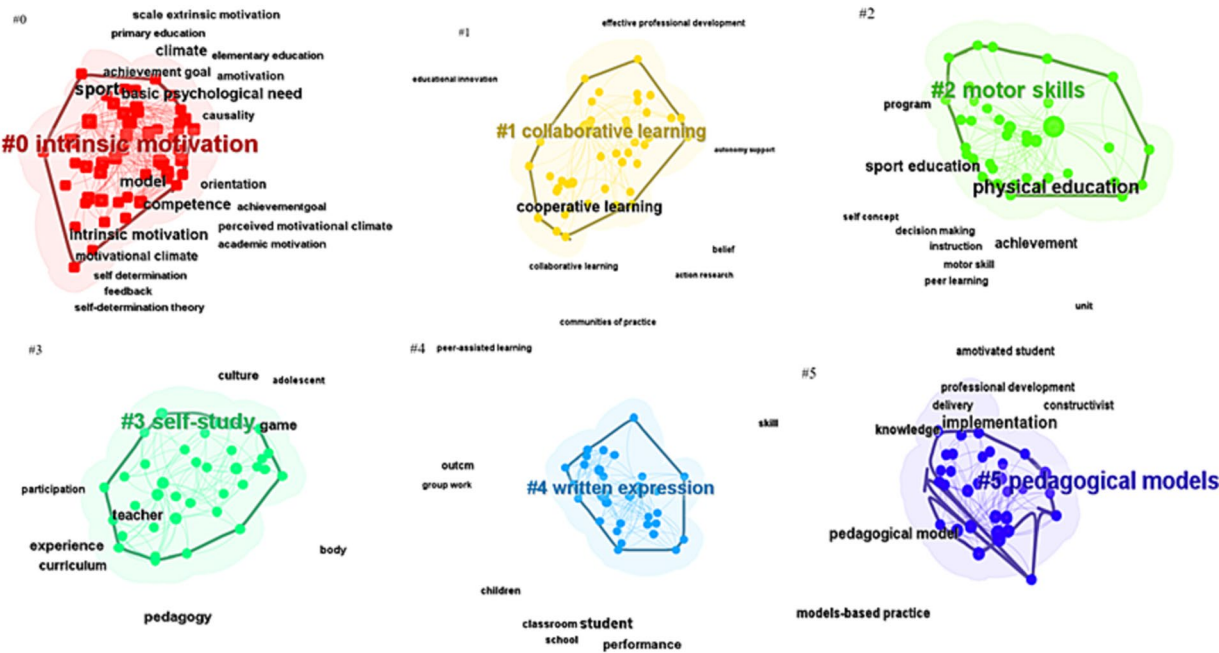


FIGURE 5
Keyword clustering knowledge graph.

should be student-centered in various PE scenarios, rather than being focused on the teacher. This is also a focus of CL, to empower students to take the lead and improve their self-directed learning abilities while engaged in cooperative learning.

As shown in Figure 8, the number of articles on CL in PE is mainly concentrated in *Physical Education and Sport Pedagogy*, with

17 articles published, cited 695 times, and an average citation frequency of 40.9 times, indicating the high quality of the articles. The analysis found that all of them were empirical analyzes, which studied the positive effects of CL on enhancing students' social skills (Silva et al., 2021a), improving their participation in class, and promoting the formation of lifelong values while enhancing their sense of

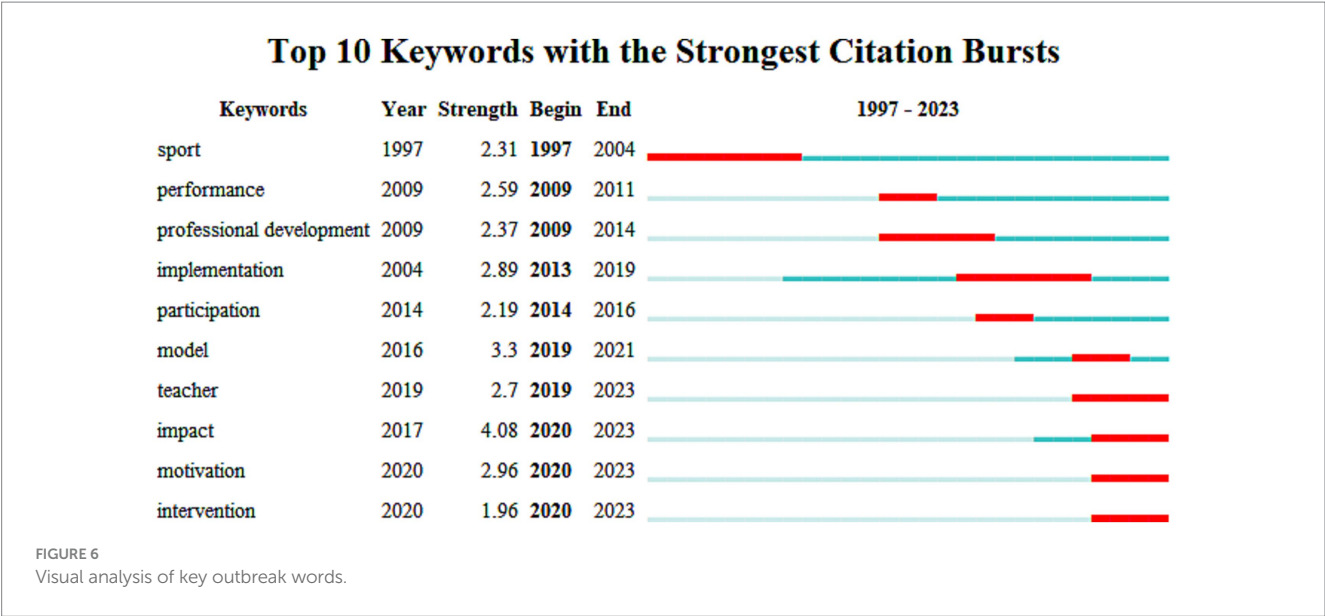


TABLE 2 Analysis of the top 5 authors by number of published articles.

Rank	Country	Author	Documents	Citations	Average citations
1	Spain	Fernandez Rio, Javier	19	210	11.0
2	Spain	Casey, Ashley	14	965	68.9
3	England	Goodyear, Victoria A	7	519	74.1
4	Spain	Rivera Perez, Sergio	5	25	5.0
5	USA	Dyson, ben	5	108	21.6

personal responsibility (Luo et al., 2020). Moreover, it has a positive impact on the use of teaching methods by PE teachers and the sustainable development of their profession (Guzmán and Payá, 2020).

Figure 9 shows that the *Journal of Teaching in Physical Education* is the most cited journal source, with a citation frequency of 454 and a connection strength of 8,673. *Physical Education and Sport Pedagogy* is the second most cited, with a frequency of 264 and a connection strength of 6,097. This is followed by *Sport, Education and Society*, *Quest*, and *European Physical Education*, with citation frequencies of 187, 147, and 143, and connection strengths of 4,192, 3,598, and 3,389, respectively. The connection strength can help analyze the degree of closeness of the journal to this topic.

4.5 Visualization analysis of author and country cooperation

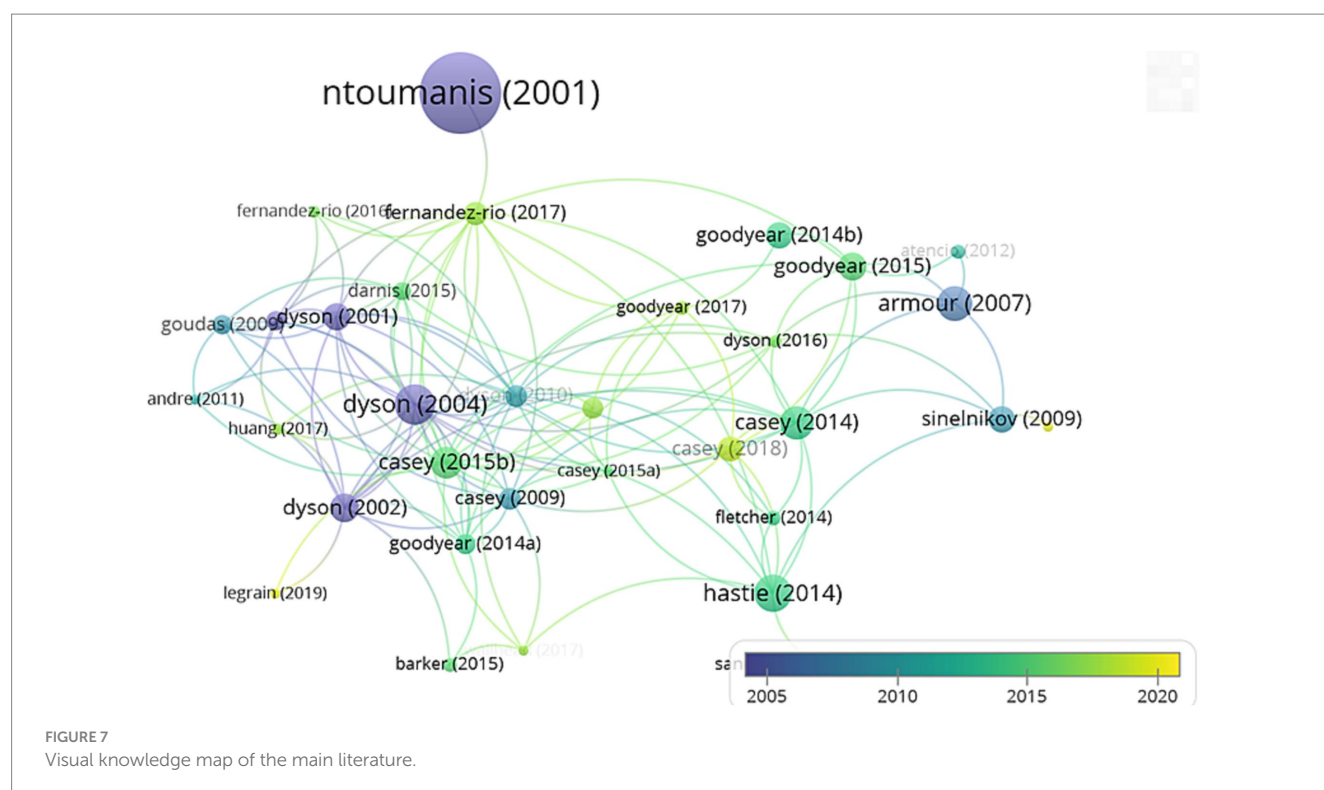
A visualization analysis was conducted on authors who have published 2 or more articles together, as shown in Figure 10. It can be observed that there are roughly 4 clusters with 5 different colors, and the thickness of the lines represents the strength of the connection between authors. The larger the circle, the stronger the author's core leadership ability in collaborative research, and the collaboration is mainly led by scholars from Spain.

Visual analysis of collaboration between countries reveals that there is no clear clustering due to lack of collaboration between

countries. The top-ranked country in terms of publication output, Spain, has closer collaboration with Portugal, the United Kingdom, and the United States, but collaborations with other countries in this field are relatively scarce. Therefore, future research should expand the scope of collaboration among countries.

4.6 Timeline graph

In visualization software, a timeline chart (Figure 11) provides a clearer and more concise understanding of the main research topics in a specific field during a certain time period. Cluster 0 includes key themes such as intrinsic motivation, sports, action, and self-directed learning, which have been continuously researched from 1997 to the present. In sports teaching, student motivation, whether intrinsic or extrinsic (Fernandez-Rio et al., 2017), is influenced by factors such as teachers, methods, and models. Cluster 1 mainly focuses on cooperative learning models (CL), Teaching Games for Understanding (TGfU) (Gil-Arias et al., 2020), Jigsaw and Team Games Tournament (TGT) (Garcia, 2021), PE models, and the development of theories in mixed or individual models. Cluster 2 is mainly about sports skills, PE, achievement, and peer-assisted learning. CL has been found to have a positive impact on students' skill development and correction, academic achievement (Ghaith, 2002), and learning in different projects. Cluster 3 is about self-directed learning, which is related to the individual development of students in CL. In order to help the team



achieve better results, students will make their best efforts to complete their assigned tasks, ultimately resulting in better team performance. Cluster 5 is similar to Cluster 2, both exploring teaching models in the field. Cluster 6 mainly focuses on the positive effects of CL classrooms on students after empirical interventions. The socialization of students, communication and cooperation between individuals, and the promotion of personal responsibility all contribute to the development of their moral level (Yunanda et al., 2018). In addition, new keywords that emerge from the combination of CL with digital information technology and artificial intelligence are worth exploring. In summary, these major clusters are all based on the expansion and exploration of “cooperative learning” and “physical education.”

5 Discussion

The purpose of this study is to use bibliometric methods to analyze the knowledge structure of CL in the field of PE. Through bibliometric analysis, a total of 167 literature references were identified in this field, coming from 69 publications, 185 institutions in 36 countries, and 380 authors. Bibliometrics, as a quantitative tool and analysis method, greatly reduces the subjectivity in information indexing and retrieval, especially when combined with visualization analysis. This can facilitate the identification of potential collaborators for scholars or countries in the field of CL. The application of CL in PE has evolved from simply adapting models to diverse development, demonstrating the superiority of CL in PE.

5.1 Dynamic trends in CL

According to the results analysis, there are four main parts. Firstly, in the field, the first three periods (1997–2013) were in the initial

growth stage, during which CL was just beginning to gain attention. From 2014 to 2018, CL showed significant growth in PE, attracting scholars' attention and being mainly empirical research to verify the learning environment for students (Chiva-Bartoll et al., 2018) and explore the impact of CL on PE (Huang et al., 2017; Sánchez-Hernández et al., 2018). The reason for this growth is mainly related to the popularity of social media and online collaboration tools, which enable multiple people to collaborate on tasks and facilitate interest and research in CL. In 2016, there were only three papers published, which may not be the peak period for CL research in PE, leading to relatively few researchers publishing relevant papers.

From 2019 to 2023, there is a rapid growth period mainly related to online education and distance learning. With the outbreak of COVID-19, many schools or institutions were forced to switch to online education, promoting research and exploration of online learning and cooperation (Yoo et al., 2021). At the same time, the development and breakthrough of emerging scientific technologies and new methods and tools continue to emerge, stimulating more research interests. For example, exploring teaching using different factors and combining computer information and communication technology such as artificial intelligence (Zhou et al., 2023b).

5.2 Information on the value of CL

Perform visual clustering on the second set of keywords, focusing on six clustered topics related to intrinsic motivation, cooperative learning, motor, self-directed learning, influence, and educational models.

In PE teaching environment, students' intrinsic motivation is influenced more by teachers and peers. CL can help students improve their learning enthusiasm, establish a good sports atmosphere, and actively participate in activities (Liu and Lipowski, 2021). CL is also

Source	Documents	Citations	Average citations
Physical education and sport pedagogy	17	695	40.9
RETOS. Nuevas Tendencias en Educación Física, Deporte y Recreación	16	40	2.5
Journal of teaching in physical education	13	773	59.5
European physical education review	11	259	23.5
Sustainability	8	48	6

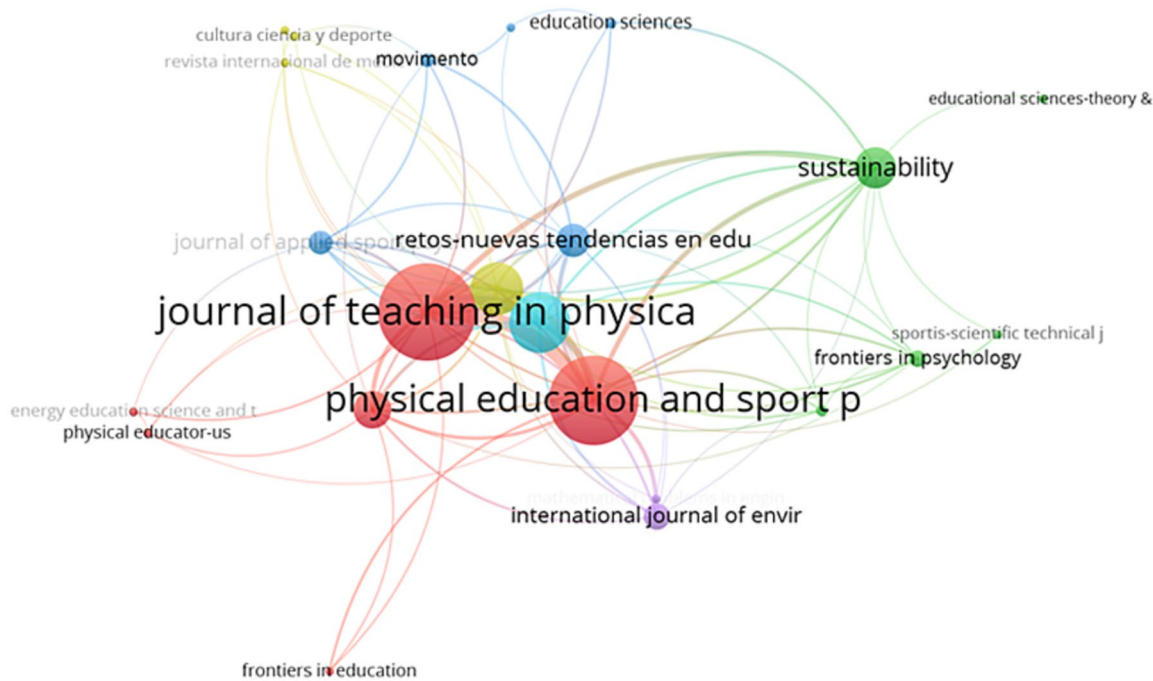


FIGURE 8
Visual knowledge map of major journals.

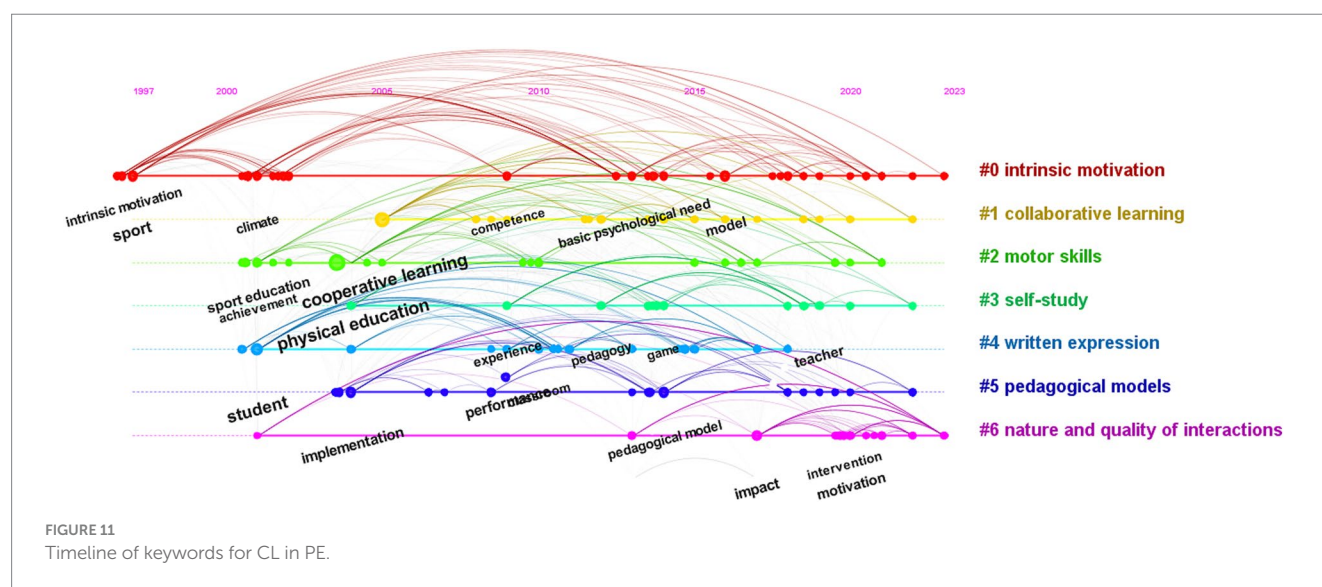
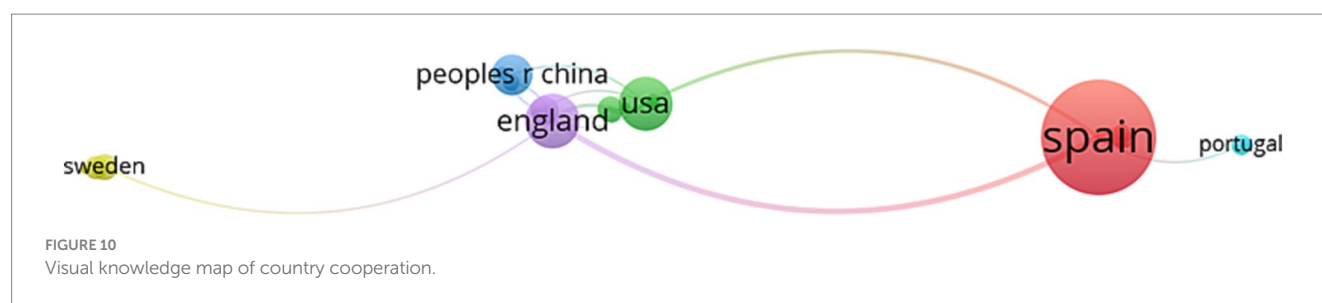


FIGURE 9
Author collaboration visual knowledge graph analysis.

key to building good friendships among peers. Facing current issues such as depression and social isolation among students, strengthening communication and cooperation among students is crucial (Polvi and Telama, 2000). A study investigating the impact of an ongoing CL intervention on student motivation has found that CL can enhance students' intrinsic motivation. Additionally, it also confirmed that

students hold positive perceptions of the classroom atmosphere created by CL (Fernandez-Rio et al., 2017).

From the beginning, CL was intended to promote students' learning of motor (Altınkürk, 2017), and to promote students' physical and cognitive development. In the use of CL for teaching basketball classes, it was observed that students' basketball skills improved, and



students were able to establish positive interpersonal relationships. This study emphasizes the importance of employing suitable grouping methods based on the teaching objectives (Yang et al., 2021). Through peer collaboration, mutual supervision, correction, and learning, cognitive development has also been enhanced, and the keywords are mostly centered around this topic.

Regarding self-directed learning, students are grouped in CL, with each person having different tasks. To complete the tasks, each person needs to independently complete a portion of the task (Janoski, 2015). Through a controlled group experiment involving 96 students in a handball teaching unit, it was confirmed that the mixed model of CL and TGfU significantly impacts the motivational atmosphere among students compared to traditional teaching methods (Chiva-Bartoll et al., 2018). There were also notable differences in students' autonomous participation. This promotes the development of self-directed learning, focusing on influencing self-development, independent learning, and other related topics.

The exploration and influence of teaching models are also mostly explored through empirical research, with the integration and exploration of different PE models such as physical education model, teaching games of understanding model, and teaching personal and social responsibility model all recognized through teaching experiments (Kirk, 2013; Kipp and Bolter, 2020; Anggreni et al., 2022). In the blend of the teaching models for PE and CL, teachers found that after 24 sessions of continuous implementation, they fostered a conducive learning environment, leading to comprehensive student development across various domains (Gil-Arias et al., 2022). However, in the application of the model, it is emphasized that to ensure the

authenticity and effectiveness of data within the implementation of CL, it is essential to integrate CL scenarios to secure the authenticity of the implementation (Casey et al., 2015). In addition, it is clear from the keyword explosion chart and the timeline of keywords that current research is mainly focused on teachers, motivation, influence, as well as current digital communication and artificial intelligence.

Specifically, the research findings indicate that CL is more commonly employed in education focused on elementary and middle school levels with students as the subjects of study. Notably, at the elementary and middle school levels, there is a stronger emphasis on student motivation, creating a motivating atmosphere, and promoting teacher-student interactions, consistent with previous research findings (Bores-Garcia et al., 2021). Approximately 44 studies (33.6%) primarily center around teachers as the subjects of research. In research centered around teachers, the primary focus lies on assessing teachers' pedagogical competence and the utilization of instructional methods. Through a study of multiple cases, it has been observed that the implementation of CL in diverse cultural settings can result in disparate effects, posing a challenge for educators in terms of instructional strategies (Karmina et al., 2021). Hence, to better facilitate the training of PE teachers in implementing CL within intricate settings, it is imperative to address challenges that may emerge during the process. These challenges encompass effective time management, fostering student motivation, and ensuring active participation in physical activities within designated areas. Additionally, research emphasizes the necessity of a strong integration of collaboration within the sports domain (Hortigüela-Alcalá et al., 2020). In addition, studies involving college students primarily

concentrate on whether CL can effectively improve learning outcomes, foster students' intrinsic motivation, increase their level of engagement, and enhance satisfaction. One study has indicated that while CL is successful in boosting social and interpersonal skills, it may not be the most efficient method for enhancing learning outcomes, and it does not show any significant gender differences (Baena-Morales et al., 2020). However, CL does appear to encourage greater participation of female students in sports activities (Goodyear et al., 2014).

On the other hand, CL has shown significant benefits for the psychological and communication development of children ($N=4$, 3.1%). CL can effectively improve children's negative emotions (Jiao et al., 2021). Specifically, CL can aid children, including those with autism, in increasing their interaction frequency with peers in physical education and free play environments, reducing instances of inappropriate interactions (Lee et al., 2022).

In the context of authorship in Spain, visual information analysis shows that Spanish scholars have conducted research mainly on the evaluation of teachers' activities in CL in PE (Bermejo et al., 2020) conducted an empirical study on students' satisfaction, performance, participation, and teacher evaluation diaries in the classroom through CL and found that students submitted very satisfied scores, with scores above 8 out of 10. This also proves that CL as a teaching model should not only be a method. Spain has small-scale contacts with countries such as the United Kingdom, the United States, and China, and in the future, it is necessary to continue to expand contacts between countries and scholars worldwide. Through literature analysis, American scholars have conducted empirical research on PE classes in primary and secondary schools, while China has continued research on PE classes for college students. The United Kingdom has examined CL, aiming to understand the potential and challenges of CL, improve the sustainable development of teachers' professions, avoid structural constraints on teaching, and actively participate in CL cases. The adaptability of CL in different national backgrounds will also be further expanded.

Finally, a visualization analysis was conducted on the journals and literature sources related to CL in PE. It was found that the journals focusing on PE were the main sources of research on CL in this field. The top-ranked journal was *Journal of Teaching in Physical Education*, with 454 citations and a connection strength of 8,673. The second-ranked journal was *Physical Education and Sport Pedagogy*, with 264 citations and a connection strength of 6,097. Following these were *Sport, Education and Society*, *Quest*, and *European Physical Education*.

6 Conclusion and limitation

In this study, a systematic review and visual analysis of 169 documents on CL in PE from the WOS database was conducted. The specific status of CL in PE was examined using bibliometrics and following the specific processes of systematic review and meta-analysis, and the visual analysis software CiteSpace and VOSviewer were used to construct a knowledge map. The individual learning model has important research implications for the development of both learners' and educators' competencies in the process of changing to CL (Liu and Lipowski, 2021). It is important for the development of intrinsic motivation of learners, increasing classroom participation, and improving teachers' ability to use flexible teaching methods.

Furthermore, we observed that the application of CL in PE is primarily concentrated in primary and secondary education and teacher training, with a secondary focus on university-level education. In primary and secondary education, the emphasis is primarily on student motivation, teacher-student interaction, group dynamics, and social learning. In college, the focus shifts toward enhancing students' physical skills, learning efficiency, and participation, particularly among female students. Concerning teacher training, there is a stronger inclination toward fostering teaching methods and technological proficiency among educators. The promotion of the overall development of learners (motor, cognitive, social and affective) and the teaching ability of educators, as well as the adaptation for different teaching environments and the integration of different teaching methods still need to be further improved (Evangeliu et al., 2021). In the future, for the application of CL in PE, it is not only necessary to focus on combining with the current teaching techniques, but also to continue to pay attention to the development of learners' and educators' abilities.

This study also has some limitations, firstly, this study only describes the current dynamic trends of CL in PE from the perspective of systematic review, and does not provide a specific review through quantitative methods; secondly, the data cleaning process of the literature, in order to improve the quality of the articles reviewed, we removed the articles from conferences, posters, books, and other languages, and in the future, we should expand the scope of the literature data; Finally, the diversity of visualization and analysis software in bibliometrics, only two of them were used in this paper for the analysis, every research method has its limitations, and future studies will further conduct a specific examination and categorization of the effects of CL on educational outcomes to make the research information more clear in the future. Through this study, we hope that it can help learners and educators to provide a lot of useful and valuable information for the development and research of CL in PE.

Author contributions

TZ: Writing – original draft, Writing – review & editing. HW: Visualization, Validation. DL: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Funding for this research was covered by the author(s) of the article.

Acknowledgments

We would like to thank the researchers and study participants for their 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

References

- Akintunde, T. Y., Musa, T. H., Musa, H. H., Musa, I. H., Chen, S., Ibrahim, E., et al. (2021). Bibliometric analysis of global scientific literature on effects of COVID-19 pandemic on mental health. *Asian J. Psychiatr.* 63:102753. doi: 10.1016/j.ajp.2021.102753
- Altunkök, M. (2017). The effect of movement education based on cooperative learning method on the development of basic motor skills of primary school 1st grade learners. *J. Balt. Sci. Educ.* 16, 241–249. doi: 10.33225/jbse/17.16.241
- Anggreni, N. P., Repiyasa, I. W., and Dewi, K. A. K. (2022). The effect of jigsaw cooperative learning model on learning outcomes of dribbling and passing using the back of the foot football game in class VII students of SMP. *J. Coach. Educ. Sports* 3, 179–188. doi: 10.31599/jces.v3i2.1549
- Baena-Morales, S., Jerez-Mayorga, D., Fernández-González, F. T., and López-Morales, J. (2020). The use of a cooperative-learning activity with university students: A gender experience. *Sustainability* 12:9292. doi: 10.3390/su12219292
- Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., et al. (2009). The educational benefits claimed for physical education and school sport: an academic review. *Res. Pap. Educ.* 24, 1–27. doi: 10.1080/02671520701809817
- Barrett, T. (2005). Effects of cooperative learning on performance of sixth-grade physical education students. *J. Teach. Phys. Educ.* 24, 88–102. doi: 10.1123/jtpe.24.1.88
- Bassachs, M., Serra, T., Bubnys, R., Cañabate, D., and Colomer, J. (2022). Multimodal approaches to math and physical education within cooperative learning to enhance social attitudes. *Sustainability* 14:16961. doi: 10.3390/su142416961
- Bermejo, J., Pulido, D., Galmés, A., Serra, P., Vidal, J., and Ponseti, F. (2020). Physical education and university: Evaluation of a teaching experience through cooperative learning. *Retos* 2041, 90–97.
- Bores-Garcia, D., Hortigüela-Alcalá, D., Fernandez-Rio, F. J., Gonzalez-Calvo, G., and Barba-Martin, R. (2021). Research on cooperative learning in physical education: Systematic review of the last five years. *Res. Q. Exerc. Sport* 92, 146–155. doi: 10.1080/02701367.2020.1719276
- Cañabate, D., Bubnys, R., Nogué, L., Martínez-Mínguez, L., Nieva, C., and Colomer, J. (2021). Cooperative learning to reduce inequalities: instructional approaches and dimensions. *Sustainability* 13:10234. doi: 10.3390/su131810234
- Casey, A., Dyson, B., and Campbell, A. (2009). Action research in physical education: Focusing beyond myself through cooperative learning. *Educ. Action Res.* 17, 407–423. doi: 10.1080/09650790903093508
- Casey, A., Goodyear, V. A., and Dyson, B. P. (2015). Model fidelity and students' responses to an authenticated unit of cooperative learning. *J. Teach. Phys. Educ.* 34, 642–660. doi: 10.1123/jtpe.2013-0227
- Casey, A., and Quennerstedt, M. (2020). Cooperative learning in physical education encountering Dewey's educational theory. *Eur. Phys. Educ. Rev.* 26, 1023–1037. doi: 10.1177/1356336X20904075
- Casey, A., and Goodyear, V. A. (2015). Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest* 67, 56–72. doi: 10.1080/00336297.2014.984733
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* 57, 359–377. doi: 10.1002/asi.20317
- Chen, W., Liu, W., Geng, Y., Brown, M. T., Gao, C., and Wu, R. (2017). Recent progress on energy research: A bibliometric analysis. *Renew. Sust. Energ. Rev.* 73, 1051–1060. doi: 10.1016/j.rser.2017.02.041
- Chen, B., and Shin, S. (2021). Bibliometric analysis on research trend of accidental falls in older adults by using Citespace—focused on web of science core collection (2010–2020). *Int. J. Environ. Res. Public Health* 18:1663. doi: 10.3390/ijerph18041663
- Chen, X., Zou, D., Xie, H., and Wang, F. L. (2021). Past, present, and future of smart learning: a topic-based bibliometric analysis. *Int. J. Educ. Technol. High. Educ.* 18, 1–29. doi: 10.1186/s41239-020-00239-6
- Chiva-Bartoll, Ó., Salvador-García, C., and Jesús Ruiz-Montero, P. (2018). Teaching games for understanding and cooperative learning: Can their hybridization increase motivational climate among physical education students? *Croat. J. Educ.* 20, 561–584. doi: 10.15516/cje.v20i2.2827
- Colao, A., Piscitelli, P., Pulimeno, M., Colazzo, S., Miani, A., and Giannini, S. (2020). Rethinking the role of the school after COVID-19. *Lancet Public Health* 5:e370. doi: 10.1016/S2468-2667(20)30124-9
- Da Luz, F. S. D. R. (2015). The relationship between teachers and students in the classroom: communicative language teaching approach and cooperative learning strategy to improve learning. In: *BSU Master's Theses and Projects*. Item 22. Available at: <https://vc.bridgew.edu/theses/22>.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* 133, 285–296. doi: 10.1016/j.jbusres.2021.04.070
- Dunn, S. E., and Wilson, R. (1991). Cooperative learning in the physical education classroom. *J. Phys. Educ. Recreat. Dance* 62, 22–28. doi: 10.1080/07303084.1991.10609885
- Dyson, B., and Casey, A. *Cooperative learning in physical education and physical activity: A practical introduction*; Routledge: London, (2016).
- Dyson, B., Griffin, L. L., and Hastie, P. (2004). Sport education, tactical games, and cooperative learning: Theoretical and pedagogical considerations. *Quest* 56, 226–240. doi: 10.1080/00336297.2004.10491823
- Er, S., and AtaÇ, B. A. (2014). Cooperative learning in ELT classes: The attitudes of students towards cooperative learning in ELT classes. *Int. Online J. Educ. Teach.* 1, 109–122. Available at: <http://iojet.org/index.php/IOJET/article/view/28/49>.
- Evangelio, C., Fernández-Rio, J., Peiró-Velert, C., and González-Villora, S. (2021). Sport education, cooperative learning and health-based physical education: another step in pedagogical models' hybridization. *J. Phys. Educ. Recreat. Dance* 92, 24–32. doi: 10.1080/07303084.2021.1977739
- Fernández-Rio, F. J. (2017). "Quality physical education through models-based practice" in *Proceedings of the 8th International Scientific Conference On Kinesiology*, 186–190.
- Fernandez-Rio, J., Sanz, N., Fernandez-Cando, J., and Santos, L. (2017). Impact of a sustained Cooperative Learning intervention on student motivation. *Phys. Educ. Sport Pedagog.* 22, 89–105. doi: 10.1080/17408989.2015.1123238
- Garcia, M. B. (2021). Cooperative learning in computer programming: A quasi-experimental evaluation of jigsaw teaching strategy with novice programmers. *Educ. Inf. Technol.* 26, 4839–4856. doi: 10.1007/s10639-021-10502-6
- Ghaith, G. M. (2002). The relationship between cooperative learning, perception of social support, and academic achievement. *System* 30, 263–273. doi: 10.1016/S0346-251X(02)00014-3
- Gil-Arias, A., Claver, F., Práxedes, A., Villar, F. D., and Harvey, S. (2020). Autonomy support, motivational climate, enjoyment and perceived competence in physical education: Impact of a hybrid teaching games for understanding/sport education unit. *Eur. Phys. Educ. Rev.* 26, 36–53. doi: 10.1177/1356336X18816997
- Gil-Arias, A., Harvey, S., Morante, Ó. M., Claver, F., and Fernández-Rio, J. (2022). Teacher and student's perspectives on their experiences within hybrid sport education-cooperative learning pedagogical model units in elementary physical education. *J. Teach. Phys. Educ.* 1, 1–9. doi: 10.1123/jtpe.2021-0304
- Goodyear, V. A. (2017). Sustained professional development on cooperative learning: Impact on six teachers' practices and students' learning. *Res. Q. Exerc. Sport* 88, 83–94. doi: 10.1080/02701367.2016.1263381
- Goodyear, V. A., Casey, A., and Kirk, D. (2014). Hiding behind the camera: Social learning within the cooperative learning model to engage girls in physical education. *Sport Educ. Soc.* 19, 712–734. doi: 10.1080/13573322.2012.707124
- Guzmán, J. F., and Payá, E. (2020). Direct instruction vs. cooperative learning in physical education: Effects on student learning, behaviors, and subjective experience. *Sustainability* 12:4893. doi: 10.3390/su12124893
- Hamzah, H., Khairiah, K., Tambak, S., Hamzah, M. L., and Purwati, A. A. (2022). Implementation of Jigsaw type cooperative learning method to increase student learning activity in Fiqh learning during COVID-19. *Int. J. Health Sci.*, 4438–4446. doi: 10.53730/ijhs.v6nS1.5914
- Hartup, W. W. *Peer interaction and the behavioral development of the individual child*; Springer Boston, MA, (1976); pp. 203–218.
- Hood, W., and Wilson, C. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics* 52, 291–314. doi: 10.1023/a:1017919924342
- Hortigüela-Alcalá, D., Hernando-Garijo, A., González-Villora, S., Pastor-Vicedo, J. C., and Baena-Extremera, A. (2020). "Cooperative learning does not work for me": analysis of its implementation in future physical education teachers. *Front. Psychol.* 11:1539. doi: 10.3389/fpsyg.2020.01539
- Huang, M.-Y., Tu, H.-Y., Wang, W.-Y., Chen, J.-F., Yu, Y.-T., and Chou, C.-C. (2017). Effects of cooperative learning and concept mapping intervention on critical thinking and basketball skills in elementary school. *Think. Skills Creat.* 23, 207–216. doi: 10.1016/j.tsc.2017.01.002
- Huang, C., Yang, C., Wang, S., Wu, W., Su, J., and Liang, C. (2020). Evolution of topics in education research: A systematic review using bibliometric analysis. *Educ. Rev.* 72, 281–297. doi: 10.1080/00131911.2019.1566212

- Ivone, F. M., Jacobs, G. M., and Renandya, W. A. (2020). Far apart, yet close together: Cooperative learning in online education. *Stud. English Lang. Educ.* 7, 271–289. doi: 10.24815/siele.v7i2.17285
- Janoski, T. (2015). The new division of labor as lean production. *Int. J. Sociol.* 45, 85–94. doi: 10.1080/00207659.2015.1061848
- Jastrow, F., Greve, S., Thumel, M., Diekhoff, H., and Süßenbach, J. (2022). Digital technology in physical education: A systematic review of research from 2009 to 2020. *Ger. J. Exerc. Sport Res.* 52, 504–528. doi: 10.1007/s12662-022-00848-5
- Jiao, W., Xue, Z., and Wang, L. (2021). Improvement of children's negative psychology based on cooperative learning. *Revista de Psicología del Deporte* 30:63. Available at: <http://creativecommons.org/licenses/by-sa/3.0/es/>.
- Johnson, D. W., and Johnson, R. T. (2014). Cooperative learning in 21st century. [Aprendizaje cooperativo en el siglo XXI]. *Ann. Psychol.* 30, 841–851. doi: 10.6018/analesps.30.3.201241
- Johnson, D. W., Johnson, R. T., and Smith, K. A. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *J. Excellence Univ. Teach.* 25, 1–26.
- Jung, H., and Choi, E. (2016). The importance of indirect teaching behaviour and its educational effects in physical education. *Phys. Educ. Sport Pedagog.* 21, 121–136. doi: 10.1080/17408989.2014.923990
- Kang, S., and Kang, S. (2019). Development of curriculum model using ICT content. *Proceedings of the Information Science and Applications: ICISA 2020*, 701–705.
- Karmina, S., Dyson, B., Watson, P. W. S. J., and Philpot, R. (2021). Teacher implementation of cooperative learning in Indonesia: A multiple case study. *Edu. Sci.* 11:218. doi: 10.3390/educsci11050218
- Kipp, L. E., and Bolter, N. D. (2020). Motivational climate, psychological needs, and personal and social responsibility in youth soccer: Comparisons by age group and competitive level. *Psychol. Sport Exerc.* 51:101756. doi: 10.1016/j.psychsport.2020.101756
- Kirk, D. (2013). Educational value and models-based practice in physical education. *Educ. Philos. Theory* 45, 973–986. doi: 10.1080/00131857.2013.785352
- Leal, K. B., Robaina, L. E. D. S., and De Lima, A. D. S. (2022). Coastal impacts of storm surges on a changing climate: a global bibliometric analysis. *Nat. Hazards* 114, 1455–1476. doi: 10.1007/s11069-022-05432-6
- Lee, G. T., He, L., and Xu, S. (2022). Using cooperative physical activities in inclusive settings to enhance social interactions for preschoolers with autism spectrum disorder in China. *J. Posit. Behav. Interv.* 24, 236–249. doi: 10.1177/10983007211035135
- Li, M., and Lam, B. H. (2013). *Cooperative learning*, vol. 1 The Hong Kong Institute of Education, 33.
- Liu, T., and Lipowski, M. (2021). Influence of cooperative learning intervention on the intrinsic motivation of physical education students—a meta-analysis within a limited range. *Int. J. Environ. Res. Public Health* 18:2989. doi: 10.3390/ijerph18062989
- Lopez, V. M., Monjas, R., Manrique, J.-C., Barba, J. J., and Gonzalez, M. (2008). Assessment implications in cooperative physical education approaches: The role of formative and shared assessment in the necessary search for coherence. *Cult. Educ.* 20, 457–477. doi: 10.1174/113564008786542208
- Luo, Y.-J., Lin, M.-L., Hsu, C.-H., Liao, C.-C., and Kao, C.-C. (2020). The effects of team-game-tournaments application towards learning motivation and motor skills in college physical education. *Sustainability* 12:6147. doi: 10.3390/su12156147
- Mao, G., Huang, N., Chen, L., and Wang, H. (2018). Research on biomass energy and environment from the past to the future: A bibliometric analysis. *Sci. Total Environ.* 635, 1081–1090. doi: 10.1016/j.scitotenv.2018.04.173
- Martínez Benito, R., and Sánchez Sánchez, G. (2020). Cooperative learning in physical education: initial challenges and developmental proposals. *Rev. Educ.* 44, 19–29. doi: 10.15517/revedu.v44i1.35617
- Metzler, M. (2017). *Instructional models in physical education*. New York: Taylor & Francis.
- Miller, S. C., Bredemeier, B. J., and Shields, D. L. (1997). Sociomoral education through physical education with at-risk children. *Quest* 49, 114–129. doi: 10.1080/00336297.1997.10484227
- Mongeon, P., and Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* 106, 213–228. doi: 10.1007/s11192-015-1765-5
- Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *Br. J. Educ. Psychol.* 71, 225–242. doi: 10.1348/00709901158497
- Oke, A., and Fernandes, F. A. P. (2020). Innovations in teaching and learning: Exploring the perceptions of the education sector on the 4th industrial revolution (4IR). *J. Open Innov. Technol. Mark. Complex.* 6:31. doi: 10.3390/joitmc6020031
- Ortiz, D. C., Gras, M. E., Pinsach, L., Cachón, J., and Colomer, J. (2023). Promoting cooperative and competitive physical education methodologies for improving the launch's ability and reducing gender differences. *J. Sport Health Res.* 15:15. doi: 10.58727/jshr.94911
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Int. J. Surg.* 88:105906. doi: 10.1016/j.ijso.2021.105906
- Peduzzi, M., Agreli, H. L. F., Silva, J. A. M. D., and Souza, H. S. D. (2020). Teamwork: revisiting the concept and its developments in inter-professional work. *Trabalho, Educação e Saúde* 18:e0024678. doi: 10.1590/1981-7746-sol00246
- Polvi, S., and Telama, R. (2000). The use of cooperative learning as a social enhancer in physical education. *Scand. J. Educ. Res.* 44, 105–115. doi: 10.1080/713696660
- Rivera-Pérez, S., Fernandez-Rio, J., and Iglesias Gallego, D. (2021). Effects of an 8-week cooperative learning intervention on physical education students' task and self-approach goals, and emotional intelligence. *Int. J. Environ. Res. Public Health* 18:61. doi: 10.3390/ijerph18010061
- Sánchez-Hernández, N., Martos-García, D., Soler, S., and Flintoff, A. (2018). Challenging gender relations in PE through cooperative learning and critical reflection. *Sport Educ. Soc.* 23, 812–823. doi: 10.1080/13573322.2018.1487836
- Shen, Y., Mao, S., Chen, F., Zhao, S., Su, W., Fu, L., et al. (2022). Electrochemical detection of Sudan red series azo dyes: Bibliometrics based analysis. *Food Chem. Toxicol.* 163:112960. doi: 10.1016/j.fct.2022.112960
- Silva, R., Farias, C., and Mesquita, I. (2021a). Cooperative learning contribution to student social learning and active role in the class. *Sustainability* 13:8644. doi: 10.3390/su13158644
- Silva, R., Farias, C., and Mesquita, I. (2021b). Challenges faced by preservice and novice teachers in implementing student-centred models: A systematic review. *Eur. Phys. Educ. Rev.* 27, 798–816. doi: 10.1177/1356336X21995216
- Song, P., and Wang, X. (2020). A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years. *Asia Pac. Educ. Rev.* 21, 473–486. doi: 10.1007/s12564-020-09640-2
- Stanne, M. B., Johnson, D. W., and Johnson, R. T. (1999). Does competition enhance or inhibit motor performance: a meta-analysis. *Psychol. Bull.* 125, 133–154. doi: 10.1037/0033-2909.125.1.133
- Van Eck, N., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538. doi: 10.1007/s11192-009-0146-3
- Yang, C., Chen, R., Chen, X., and Lu, K.-H. (2021). The efficiency of cooperative learning in physical education on the learning of action skills and learning motivation. *Front. Psychol.* 12:717528. doi: 10.3389/fpsyg.2021.717528
- Yoda, I. The development of cooperative learning model based on local wisdom of Bali for physical education, sport and health subject in junior high school. In: *Proceedings of the IOP Conference Series: Materials Science and Engineering*, (2017); p. 012166.
- Yoo, J.-I., Han, J.-K., Youn, H.-S., and Jung, J.-H. (2021). Comparison of health awareness in South Korean middle school students according to type of online physical education classes during the COVID-19 pandemic. *Int. J. Environ. Res. Public Health* 18:7937. doi: 10.3390/ijerph18157937
- Yunanda, H., Advinda, L., and Sumarmin, R. (2018). Effects of cooperative learning model type games teams tournament (TGT) and entry behavior student to learning competence class XI IPA senior high school 1 Lengayang. *Int. J. Progress. Sci. Technol.* 24, 1333–1348. doi: 10.1016/j.tate.2007.10.003
- Zhou, T. (2023). Bibliometric analysis and visualization of online education in sports. *Cogent Soc. Sci.* 9:2167625. doi: 10.1080/23311886.2023.2167625
- Zhou, T., Su, Y., Huang, Y., and Liu, K. (2023a). “Development review: 2012–2022 physical training visualization” in *Proceedings of the 2023 4th International Conference on Education, Knowledge and Information Management (ICEKIM 2023)*, 1665–1675.
- Zhou, T., Wu, X., Wang, Y., Wang, Y., and Zhang, S. (2023b). Application of artificial intelligence in physical education: a systematic review. *Educ. Inf. Technol.*, 1–18. doi: 10.1007/s10639-023-12128-2



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Aneta Grochowska,
University of Applied Sciences in Tarnow,
Poland

Susana Henriques,
Universidade Aberta (UAb), Portugal

*CORRESPONDENCE

Liang Liu
✉ T23092100329@cityu.mo

RECEIVED 23 October 2023

ACCEPTED 16 November 2023

PUBLISHED 04 December 2023

CITATION

Deng X, Liu L and Liu X (2023) Research on service quality of professional education based on the PZB service quality model.
Front. Educ. 8:1325256.
doi: 10.3389/feduc.2023.1325256

COPYRIGHT

© 2023 Deng, Liu and Liu. 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.

Research on service quality of professional education based on the PZB service quality model

Xuejun Deng¹, Liang Liu^{2*} and Xiaojun Liu¹

¹School of Business, Shenzhen Institute of Technology, Shenzhen, Guangdong, China, ²Faculty of International Tourism and Management, City University of Macau, Macau, Taipa, China

With the arrival of Internet technology and the information era, China's higher education has entered a new stage of development. In this context, new learning methods represented by "Internet+" have gradually emerged and rapidly spread, and traditional classroom teaching is facing a huge impact. At the same time, more and more students choose to learn independently through online courses due to resource constraints, time and space. This new learning method has posed a great challenge to the traditional education concept. Therefore, how to improve the quality of college education services has become one of the urgent problems to be solved at present. In this paper, from the perspective of service quality of college education, it is of practical significance to apply the theory of hybrid service quality to college hybrid education to explore its service quality improvement strategy. In this paper, the application of the PZB service quality model in professional education teaching is studied and discussed by using the methods of literature, mathematical statistics, and logical analysis. In this paper, the structure of the service quality evaluation system, influencing factors and results analysis, and the summary of improvement suggestions are deeply analyzed.

KEYWORDS

quality of teaching services, higher school education, PZB service quality, Delphi method, hierarchical analysis method

1 Introduction

At present, the concept of service quality has been more in-depth and comprehensive understanding. From the initial focus on customer satisfaction and service remediation, a more complete and mature theoretical system has been gradually formed. Among them, service quality in the field of education is mainly reflected in the construction of relevant indicators to measure whether the educational services provided by schools to students meet the standards and how well they meet the needs, and on this basis, analyze the factors that affect the work, and then propose corresponding strategies and methods. However, there are still some shortcomings in the research on service quality, especially in the special form of hybrid education. Therefore, in this paper, we will take a university's hybrid education service quality as an example and conduct further exploratory and confirmatory analysis research.

2 Literature review

2.1 PZB service

In this study, PZB service quality refers to comprehensive, effective, and accurate guidance by the teachers of professional education in the learning process of students, so as to achieve the purpose of guiding and improving the physical and mental quality of athletes and cultivating good habits. In the process of sports, the teaching contents and methods, means and forms of activities as well as relevant policies and regulations are organically combined (Simcic Brønn, 2012). PZB enables teachers and students to realize their self-worth and personal needs better and faster. It mainly includes the following aspects: (1) Provide high-quality curriculum resources with teaching objectives, so that students can get better development in professional education courses; (2) Establishing a sound and rational curriculum and teaching materials with strong regularity and relevance; (3) Regularly hold various activities to attract students to participate in physical exercise; (4) Regularly hold various activities so that students can get physical and mental health in sports and improve their interest in physical exercise (Streetman and Goodman, 2022).

2.2 Characteristics of service quality in professional education

The study of service quality is the focus of foreign scholars in recent years on the quality of education, students' physical and mental health, social and economic development, etc. This paper analyzes the problems in university professional education and proposes corresponding countermeasures. In this regard, the following characteristics of service quality are proposed:

- 1 From the student's point of view, we understand the current situation of the service quality of professional education teaching in schools. At the same time, teachers' use of new technologies to teach is also a professional teaching method to improve student performance (Liu et al., 2023). Through the survey and research, the relevant factors affecting the quality of work of professional education teachers and students' satisfaction with them are analyzed, and suggestions are made to effectively improve the physical fitness level of our college students, improve the curriculum and enhance the logistic support.
- 2 Students' perceptions of excellent and poor service quality, the composition and influencing factors of professional education service quality in colleges and universities were analyzed. To improve the service quality, it is necessary to strengthen the main position of students, improve the logistics guarantee system. We found that a large proportion of students do not engage in physical activity regularly, and the proportion of physical exercisers who exercise less than 4 h a day account for less than 50% (Gil-Madrona et al., 2019a). Therefore, we must focus on cultivating students' interest in physical exercise and pay attention to the cultivation of students' interest in sports and exercise, and improve the level of teachers' professional skills, etc. These measures are used to improve the current development of professional education (Shang et al., 2014).

- 3 Quality is assessed not only in terms of the outcome of the service, but also in terms of the process of the service. In the teaching process, it has been proved that the teaching dimension and the respect dimension are the two most important dimensions of teaching authority (Gil-Madrona et al., 2020). and the level of teaching quality is improved through the analysis of students, teachers and sports facilities. Meanwhile, Parsuraman, Zeithaml, and Berry proposed the service quality gap model (Figure 1). This model is a customer-oriented, service quality-as-a-whole model that evaluates and judges' students' teaching activities, thus achieving the goal of improving communication between teachers and students and the sports department as well as improving the management of professional education in universities. This model is a way to evaluate the teaching process using effective tools based on factors such as students, teachers and sports facilities, thereby improving the level of quality of professional education services in colleges and universities.

3 Methods of evaluating the quality of professional education teaching services in colleges and universities

The research on the evaluation method of professional education teaching service quality mainly refers to the use of scientific rationality principle and mathematical statistics and other theories combined with practice to build the satisfaction measurement system of college students, teachers and parents. Through the satisfaction survey of college students, teachers and parents, the factors that affect the quality of teaching and learning can be identified, and based on this, corresponding countermeasures can be proposed. In order to evaluate the teaching accurately and effectively, an evaluation model is established as shown in Figure 2. The main function of the model is: based on the satisfaction of teachers and students, through the analysis of its evaluation model, the problems in the teaching process can be found, and there is a need to strengthen the management and improve the quality of service in the practice of sports programs (Ritzer et al., 2021).

3.1 Delphi method

3.1.1 Concept of Delphi method

In the actual teaching process, the use of the Delphi method can effectively improve the level of communication and interaction between teachers and students and the level of service quality. It uses a certain number of investigators to explore and communicate the research questions to be studied, and then feedback to the relevant experts, who will then revise, amend their opinions and finally get the conclusions, in the use of this technique will generally have a process steps: firstly, determine the number of options proposed by the researcher, and then evaluate the options, and according to the feedback to amend the finalized questionnaire. Secondly, the results of the effectiveness evaluation should be accepted to a different degree for each feasible person selected, and then the results are given back to the students, so that teachers can improve the quality of their

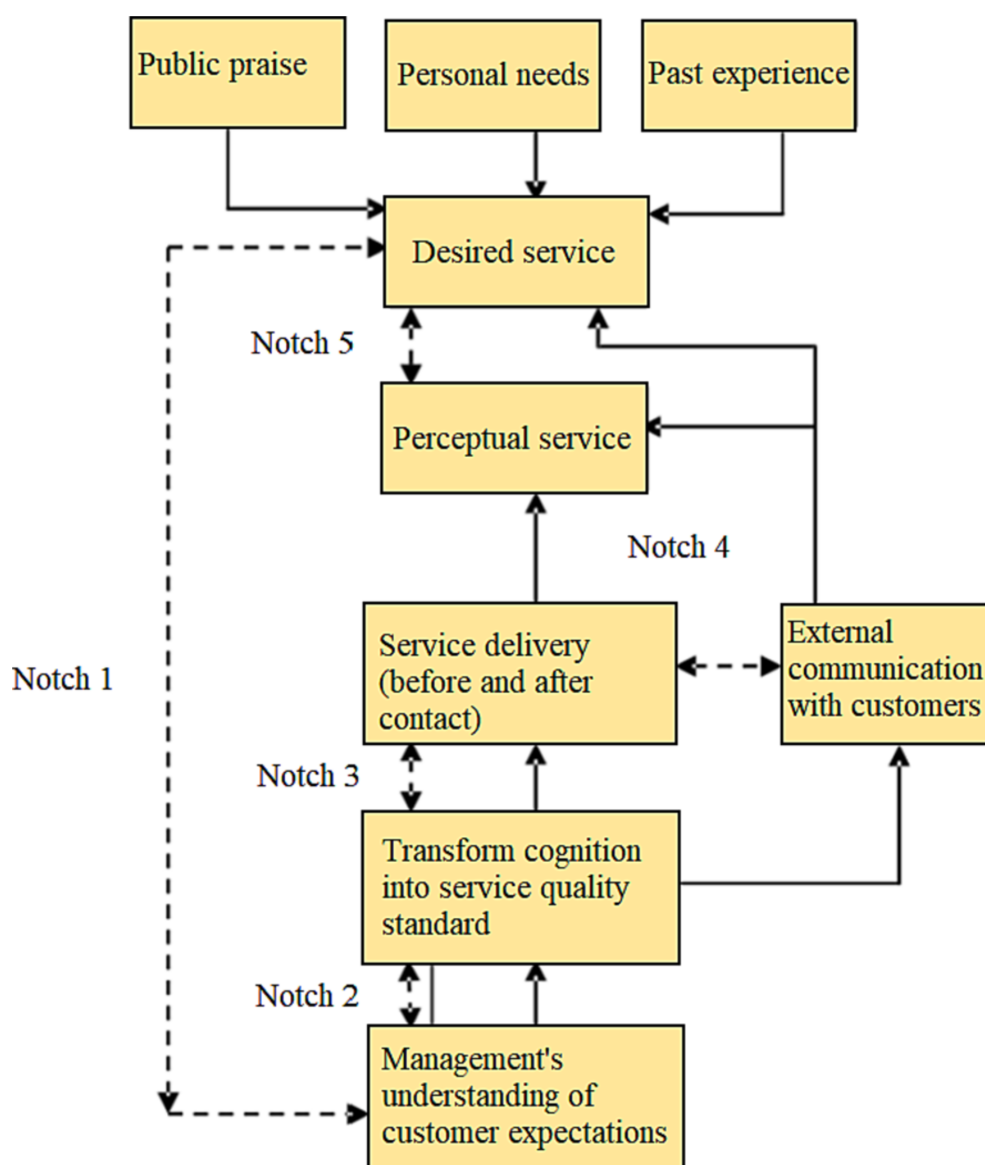


FIGURE 1
PZB service quality gap model.

services by conducting teaching activities according to different evaluation criteria (Zubkova and Tagirova, 2020). Finally, it is suggested that each scholar should adjust and improve the content of each factor in a targeted way.

3.1.2 Delphi method for expert research

The Delphi method is a qualitative analysis method, the basic principle of which is to organize and summarize the information in the minds of experts, statistics and calculations, and transform it into simple and easy-to-understand regular characteristics, and then into a simple and easy to understand mathematical model, so that students have a general understanding of things, so that they can grasp the regular knowledge more easily to understand and remember. In practical applications, anonymous surveys are used. This random sampling of questions allows more teachers to

participate in the survey to understand the teaching situation and related work dynamics, and also to get their ideas and needs, so that they can take the initiative to improve the teaching work, thus increasing students' interest in learning, so that their motivation can be better mobilized (Miśniakiewicz and Krnáčová, 2021). The process of expert survey using the Delphi methods shown in Figure 3.

The comprehensive analysis of student satisfaction in the process of evaluating the quality of sports services is a very important and practical and meaningful area of research. By building on the PZB-type structural model, the required information on the relationship between college teachers and students, teaching contents and methods, and mutual communication between teachers and students was constructed, and the judgment matrix of its literature information resource evaluation is shown in Table 1.

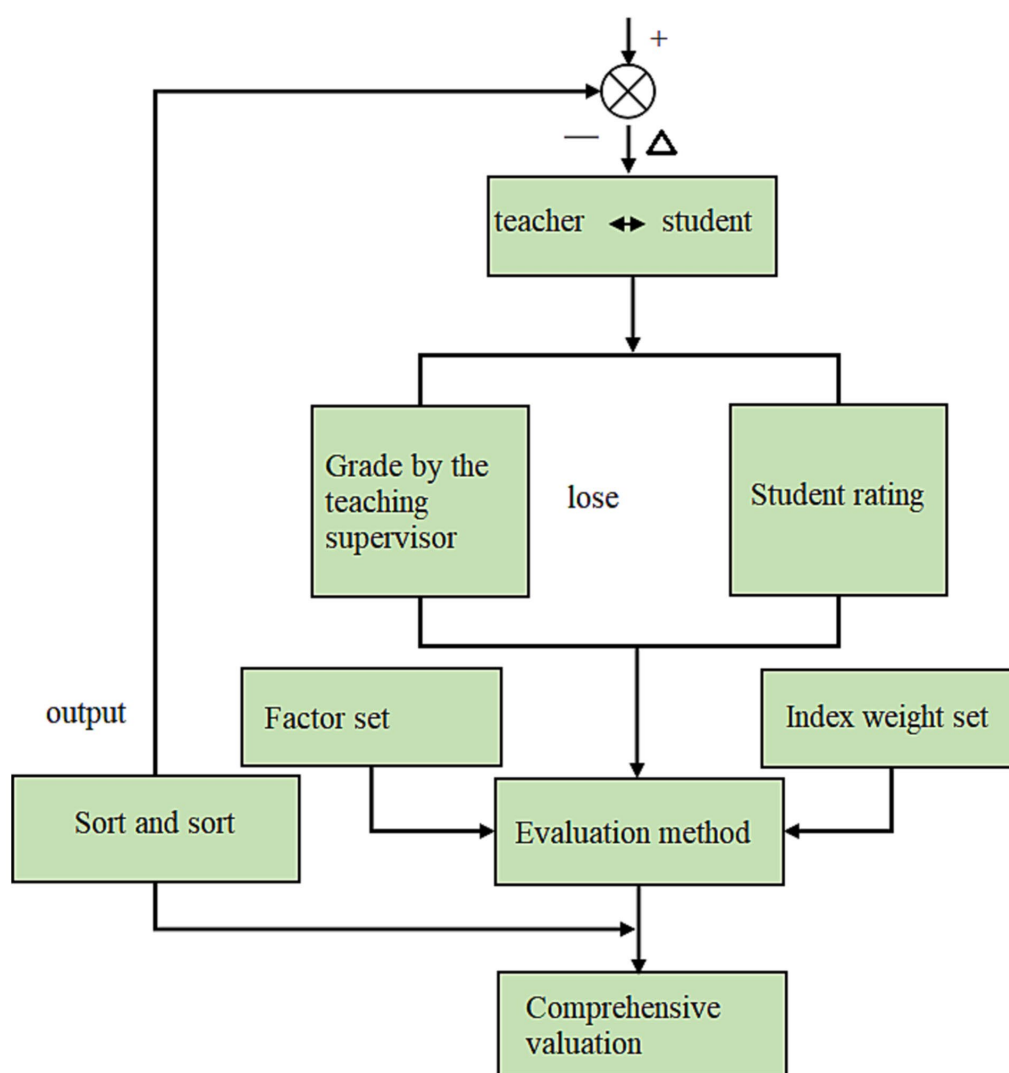


FIGURE 2
Teaching evaluation model.

3.2 Hierarchical analysis method

3.2.1 Hierarchical analysis

The hierarchical analysis is a systematic research method that combines qualitative and quantitative aspects, translates the problem into a specific value, and gives reasonable suggestions based on the actual situation. When used in the application process, it is necessary to first determine what kind of relationship exists between the factors, and secondly to consider several aspects affecting students' learning effectiveness, teaching quality and improving the performance of professional education teachers, so as to achieve a comprehensive and systematic analysis of students and to come up with reasonable suggestions. This can provide a theoretical basis for professional education teaching reform, which is also an important way to solve the current problems in the quality of professional education services (Parida and Baksi, 2011). By using this method, it can accurately reflect the current situation and development trend of school professional education service quality:

first, ensure that the course content provided by the school conforms to the regulations of the Ministry of Education of China; second, strengthen the interaction between teachers and students and improve students' participation; third, enhance the teaching enthusiasm of professional education teachers by establishing a perfect evaluation mechanism; fourth, strengthen the construction and management of the course assessment system (Long et al., 2020). Based on this, the specific steps of the derived hierarchical analysis method are shown in Figure 4.

3.2.2 Calculation of index weights by using hierarchical analysis method

In the evaluation index of teaching service quality, the analysis of students' satisfaction is mainly done by establishing a hierarchical structure model to determine the influence degree weights of each factor and determining the influence degree of each factor on students' satisfaction by fuzzy comprehensive evaluation method, so as to come up with the strategy to improve

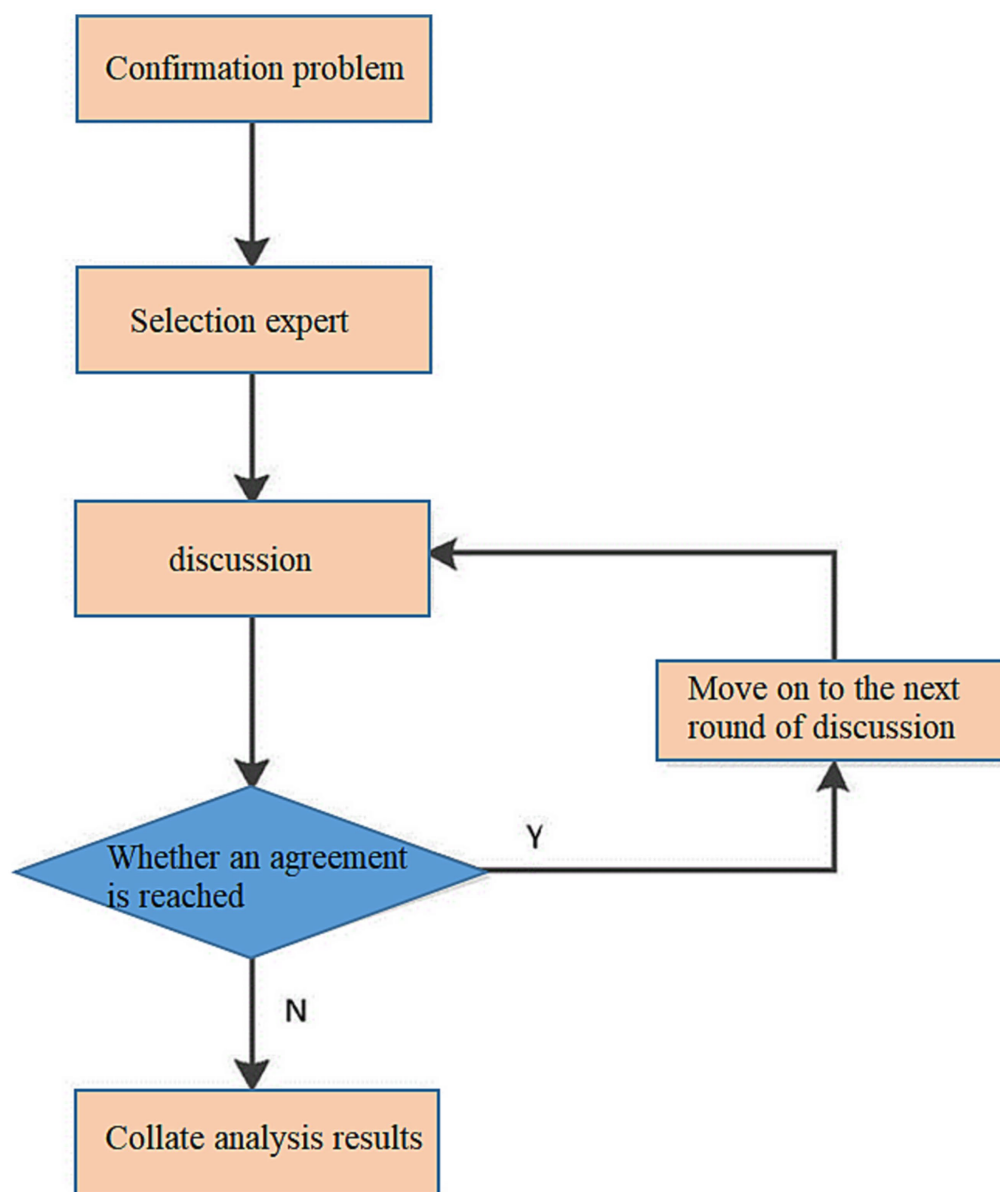


FIGURE 3
Flow of expert survey using Delphi method.

TABLE 1 Judgment matrix of literature information resources evaluation.

A	A1	A2	A3	A4
A1	1	1/4	1/4	2
A2	4	1	1	3
A3	4	1	1	3
A4	1/2	1/3	1/3	1

the management level of professional education teaching service quality. In order to determine the weights, this paper uses hierarchical analysis to identify the main factors affecting the quality of professional education teaching services in colleges and universities and quantifies each dimension to derive the final

results. By using the fuzzy comprehensive evaluation model to study the current management status and development trend of colleges and universities in most regions of China, the steps of its teaching evaluation are:

- ① Set up a set of factors, which is a collection of indicators composed of evaluation indicators, as shown in formula (1).

$$U = [u_1, u_2, \dots, u_n] \quad (1)$$

- ② Set up an evaluation set, which is a set that indicates the degree of merit of the evaluation objectives, as shown in formula (2).

$$C = [c_1, c_2, \dots, c_m] \quad (2)$$

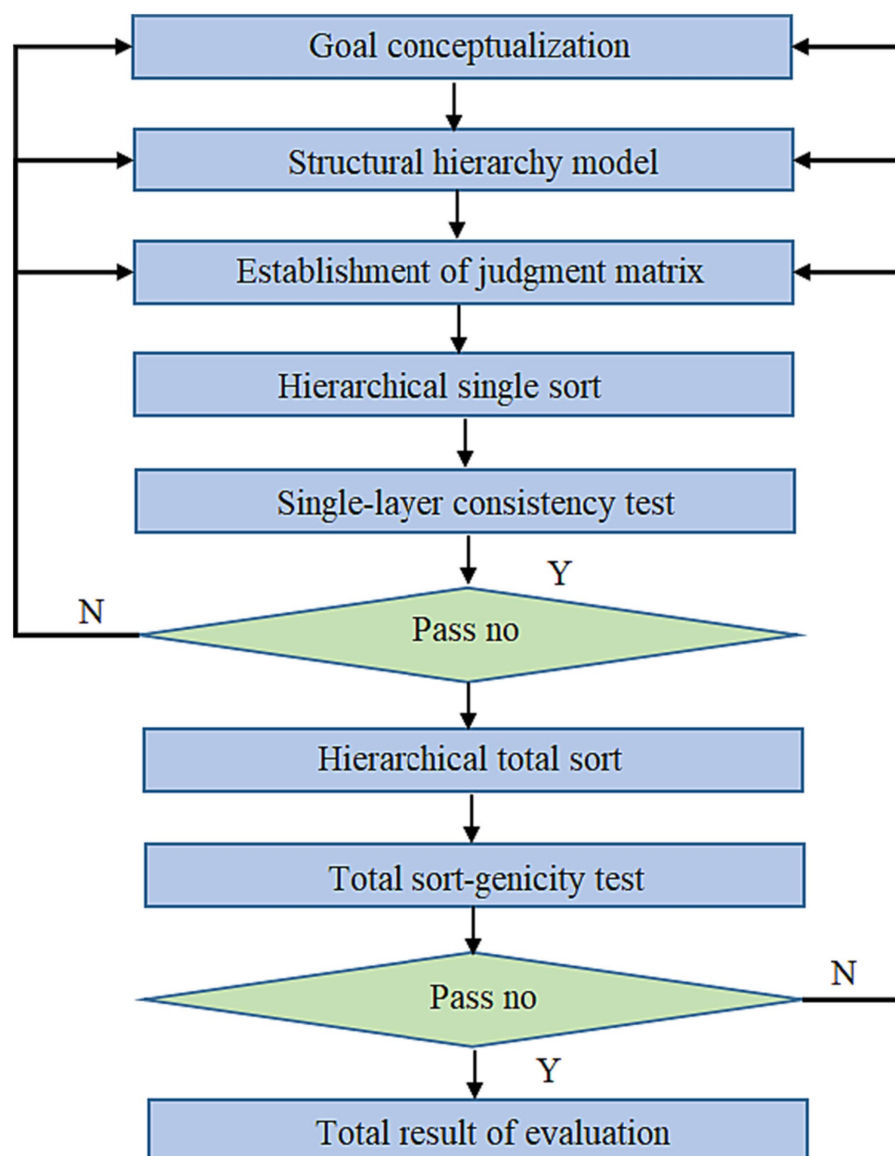


FIGURE 4
Specific steps of the hierarchical analysis method.

- ③ The weight matrix of indicators is obtained from the hierarchical analysis (AHP), as shown in formula (3).

$$W = [w_1, w_2, \dots, w_m] \quad (3)$$

where, we are the weight value of each index. Through the investigation and analysis of the current situation of teaching service quality of professional education teachers, it is found that the current professional education teaching service in China has problems such as weak awareness of student participation, poor interaction between teachers and students, and lack of scientific and reasonable evaluation mechanism (Fallah Jelodar, 2016). In response to these problems, this paper puts forward relevant suggestions: improve the importance of school leaders, improve the management system and assessment standards; strengthen the construction of logistic support, and

establish effective incentive and restraint mechanisms and rewards and punishments, so as to promote teachers and students to actively learn new knowledge and skills, which can lay a solid foundation for better development in the future.

- ④ Construct a single-factor evaluation matrix R , where $R_i = (r_{i1}, r_{i2}, \dots, r_{in})$ is the result of the evaluation of the factor as shown in formula (4), which has an impact on the satisfaction of students. The questionnaire and interview results found that professional education teachers have problems such as irregularity and unreasonableness in the teaching process, and the corresponding countermeasures are proposed for these phenomena: optimize the curriculum structure, strengthen the communication mechanism with parents to establish a good teacher-student relationship, improve the logistics guarantee system to improve the level of logistics services, so as to ensure the quality of

professional education work in colleges and universities, improve student satisfaction, and actively improve the quality of logistics services.

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \quad (4)$$

- ⑤ The weighted average type of comprehensive judgment model [as shown in formula (5)] was used to evaluate and analyze the quality of professional education teaching services in colleges and universities, and it was found that the management mechanism of teaching service quality has not yet been formed, and the enthusiasm of students' participation needs to be improved. In the survey, it was learned that the factors affecting the efficiency and effectiveness of professional education teachers in colleges and universities are multifaceted and multidimensional, mainly including incomplete teaching resources and equipment, poor professional quality of professional education teachers, and inadequate management of students by school leaders. In response to the above problems, corresponding countermeasures were proposed, such as strengthening the construction of hardware facilities and improving the equipment of professional education, improving the efficiency and quality level of logistics support, establishing an effective feedback mechanism and reforming the assessment system (Biese et al., 2021).

$$M = (\oplus) = \sum_{i=1}^n (W_i \ r_{mn}) \quad (5)$$

Calculating the comprehensive judgment matrix [as shown in formula (6)] is most used in fuzzy mathematics in three models: hierarchical analysis, fuzzy comprehensive judgment model, and gray system method. At present, the evaluation of professional education teaching service quality is mainly based on qualitative indicators. The hierarchical analysis method is the most widely used in education evaluation, which can make full use of the information provided by the factor fuzzy matrix to quantify the evaluation indicators of professional education teaching service quality, so that students can have an intuitive and comprehensive understanding of their learning effect and teachers' working efficiency, thus improving the teaching service level (Nowoświat and Leszczyńska, 2016; Goodson et al., 2021).

$$A = W \quad R = [w_1, w_2, \dots, w_m] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \quad (6)$$

- ⑥ The gray correlation degree was calculated to determine the correlation order, analyze the influencing factors, and propose countermeasures to improve the quality of professional education teaching services in colleges and universities in order to improve students' participation in professional education

classes. The correlation degree is calculated as shown in formulas (7)–(9).

$$\eta_{ij}(k) = \frac{\min_i \min_k \Delta_i(k) + \ell \max_i \max_k \Delta_i(k)}{\Delta_i(k) + \ell \max_i \max_k \Delta_i(k)}, \ell \in (0,1) \quad (7)$$

$$\Delta_{ij}(k) = |A'_j(k) - A'_i(k)| \quad (8)$$

$$\eta_{ij} = \frac{1}{k} \sum_{i=1}^k \eta_{ij}(k), k = 1, 2, \dots, n \quad (9)$$

3.3 Computational analysis of teaching quality

The calculation and analysis of teaching quality refer to the dynamic adjustment of the relationship between teachers and students and between teachers and students through actual perception, feedback and monitoring in the process of students' learning of professional education courses. Before the teaching practice is carried out, it is necessary to understand whether the content taught meets the requirements of this stage, and if not, it is necessary to correct and reformulate the relevant system in time to ensure the effective operation of the classroom; on the contrary, when problems arise, the teaching should be stopped immediately or measures should be taken to solve them, so as to ensure that the teaching quality is continuously improved and improved, and finally improve the comprehensive quality and ability of students. The above-mentioned method and model are used to evaluate the quality of classroom teaching in a certain teaching evaluation in our college, for example.

- ① Set up the factor set: $U = [\text{us/teaching content, us/teaching method, us/teaching attitude, us/teaching basic skills, us/teaching effect}]$.
- ② Set up the evaluation set and assign the value: $C = [\text{very good, good, better, fair, poor}]$.
- ③ By the AHP method, on the basis of consulting experts, the corresponding judgment matrix is established by the 1–9 scale method to determine the weights of factors at the same level, and then the judgment matrix is shown in the formula (10).

$$A = \begin{bmatrix} 1 & 2 & 3 & 2 & 2 \\ 1/2 & 1 & 2 & 2 & 2 \\ 1/3 & 1/2 & 1 & 1/2 & 1/2 \\ 1/2 & 1/2 & 2 & 1 & 1/2 \\ 1/2 & 1/2 & 2 & 2 & 1 \end{bmatrix} \quad (10)$$

- ④ The one-factor evaluation matrix R obtained by the expert panel for the three teachers' ratings is shown in formulas (11)–(13).

$$R_1 = \begin{bmatrix} 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.2 & 0.3 & 0 \\ 0.2 & 0.5 & 0.2 & 0.1 & 0 \\ 0 & 0.2 & 0.4 & 0.3 & 0.1 \\ 0 & 0.3 & 0.5 & 0.1 & 0.1 \end{bmatrix} \quad (11)$$

$$R_2 = \begin{bmatrix} 0.2 & 0.4 & 0.3 & 0.1 & 0 \\ 0.2 & 0.3 & 0.4 & 0.1 & 0 \\ 0.5 & 0.3 & 0.1 & 0.1 & 0 \\ 0.1 & 0.4 & 0.2 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.2 & 0.2 & 0 \end{bmatrix} \quad (12)$$

$$R_3 = \begin{bmatrix} 0.1 & 0.3 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.3 & 0.2 & 0.2 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0.1 & 0.3 & 0.2 & 0.3 & 0.1 \\ 0.3 & 0.2 & 0.4 & 0.1 & 0 \end{bmatrix} \quad (13)$$

A weighted average type of comprehensive evaluation model was used to analyze the current situation of the quality of professional education teaching services in Hunan universities using the questionnaire survey method. Problems affecting students' low participation and lack of teachers' active communication ability were found, and the evaluation results were obtained by using the synthetic operation of the fuzzy matrix, as shown in the formula (14).

$$A_1 = W$$

$$R_1 = [0.34361, 0.24012, 0.09638, 0.13791, 0.18198] \quad (14)$$

$$\begin{bmatrix} 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.2 & 0.3 & 0 \\ 0.5 & 0.5 & 0.1 & 0.1 & 0 \\ 0 & 0.2 & 0.4 & 0.3 & 0.1 \\ 0 & 0.3 & 0.5 & 0.1 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.101661 \\ 0.271124 \\ 0.350898 \\ 0.209967 \\ 0.06635 \end{bmatrix}$$

4 Design of hybrid education quality evaluation system based on the PZB service quality model

4.1 System architecture design

The hybrid education model is a new teaching method based on computer technology and communication network. It is based on computer information technology and uses various advanced means to build a systematic and standardized service quality evaluation system for colleges and universities. By monitoring and managing all aspects of the service process, the modelist improves students' learning efficiency and satisfaction, enhances students' learning motivation and realizes the sharing of higher-quality educational resources. From the

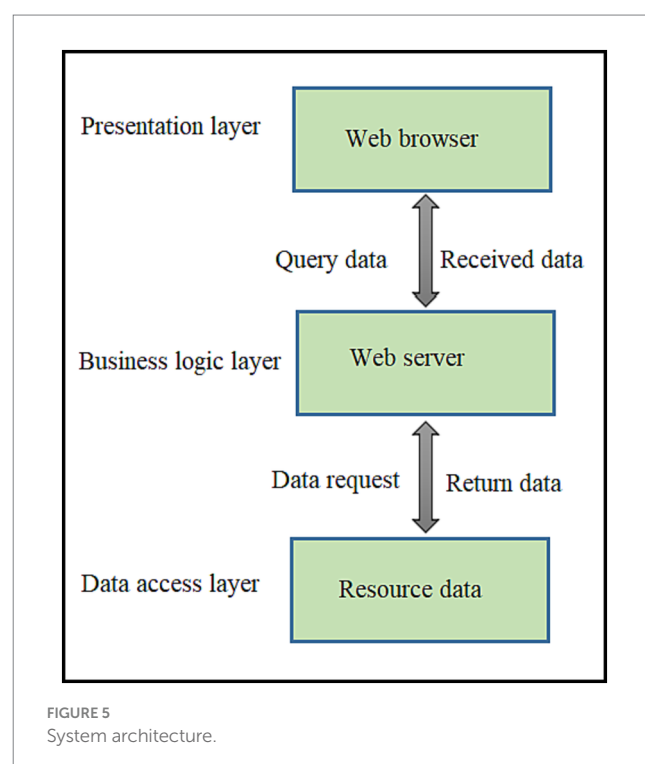
theoretical point of view: it is important to establish a perfect scientific and reasonable evaluation index system to improve students' comprehensive quality, and from the practical point of view: a scientific and reasonable evaluation model is established by taking multiple dimensions such as teachers' teaching effect and school running efficiency as evaluation criteria. The system architecture is shown in Figure 5.

4.2 Teaching quality evaluation system function module

Based on the PZB service quality evaluation system, the quality of hybrid education is scientifically and reasonably evaluated, mainly based on the number of problems encountered by students in the learning process, the length of time and the learning effect and other indicators to determine the teaching content. The evaluation criteria and methods are used to determine whether the teachers and students have achieved the expected goals in the implementation of the course. Based on this paper, the teaching quality evaluation system is studied, and its functional modules are shown in Figure 6.

4.3 Sampling strategies

This study adopts the "purposive sampling" of non-probability sampling in quantitative study research sampling methods, that is, according to the research purpose, the research objects that can provide the maximum amount of information for the research question are selected (Borrego et al., 2009).



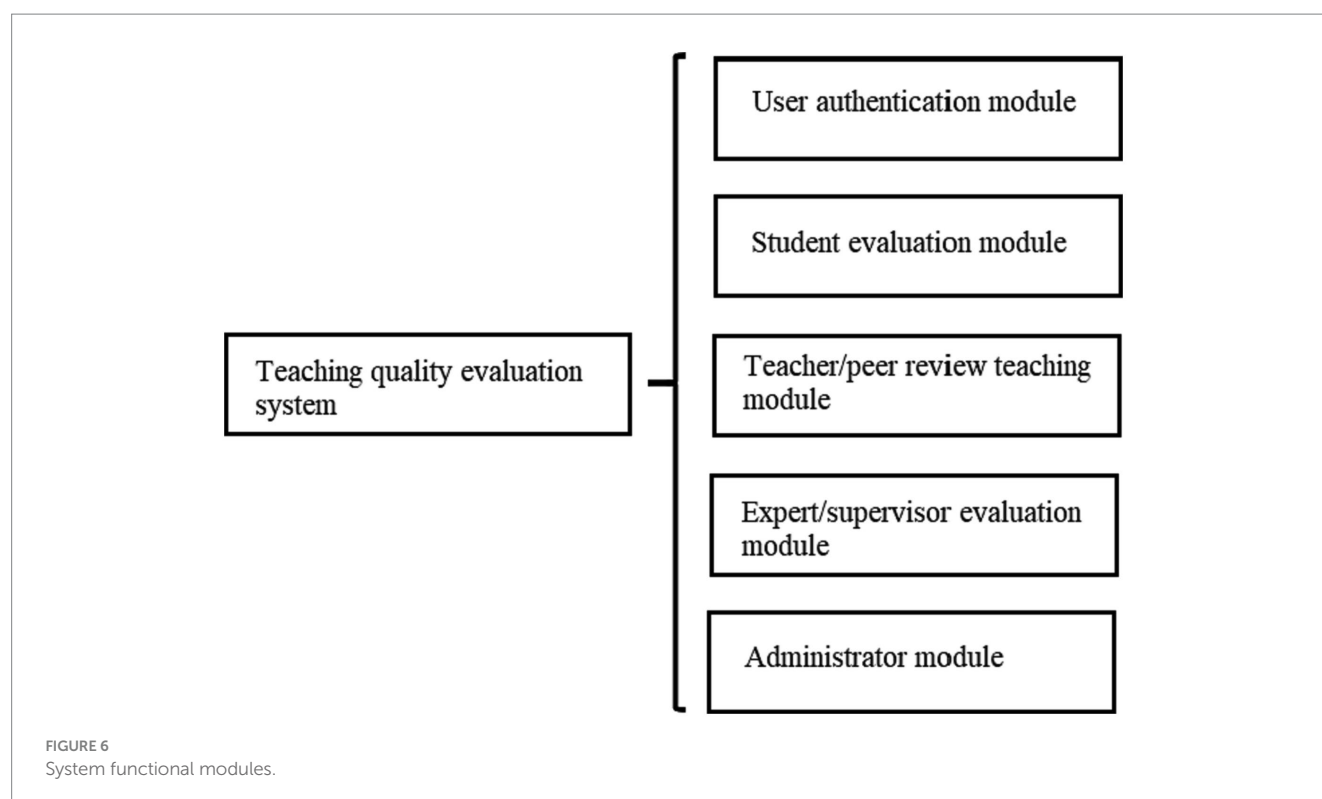


TABLE 2 Analysis table of college students' overall expectation of professional education teaching service quality ($N = 770$).

Layer	Average	Standard deviation	Sort
Visible view	4.15	0.54	1
Reaction view	4.10	0.65	2
Assurance view	4.02	0.47	3
View of human nature	3.84	0.56	4
Whole	4.02	0.55	

4.4 Data collection

For this study, we collected data at Shenzhen Institute of Technology. We used purposeful sampling to select 800 students for a questionnaire survey. After extracting the data and excluding questionnaires with consistent answers and incomplete answers, there were 770 valid data.

5 Research results and discussion

5.1 Analysis of college students' expectations of PE teaching service quality

5.1.1 Overall expectation analysis

Table 2 shows that the overall expectation of college students for the service quality of professional education is 4.02, which

indicates that students have high expectations for the teaching quality of professional education courses. Through the survey of professional education teaching quality, we found that most students are generally satisfied with the service quality of the existing courses in colleges and universities, which shows that at present our professional education is constantly developing and improving. On the other hand, this also reflects the current public participation in sports and fitness awareness is not strong enough, the school management is not in place and other problems. Therefore, it is of great significance to improve the physical condition of college students and promote the healthy growth of the whole people.

5.1.2 Analysis of expectations of students from different backgrounds on professional education teaching service quality

Through the analysis of teachers' and students' expectations in the questionnaire, there are differences in the perceived degree of professional education teaching service quality among students from different backgrounds. Under the traditional exam-oriented education, most schools focus on theoretical knowledge learning and neglect practical ability cultivation, therefore, many colleges and universities limit the teaching content to books for explanation and demonstration and other ways to improve classroom efficiency. Or, in order to cope with the examination only to arrange some courses with a large amount of class time, high difficulty, and difficulty to grasp and complete the task, they ignore the impact of factors such as practical hands-on operation and development of creative thinking ability of students on the quality of teaching, which makes students in professional education teaching cannot really master the knowledge learned, thus affecting the quality of school education.

Under the PZB service quality model, universities investigate and study their own conditions and actual situation and propose corresponding countermeasures.

5.1.3 Analysis of college students' expectations in different levels of professional education teaching service quality

By analyzing the different expectations of college students regarding the quality of professional education teaching services, we found that most students are satisfied with the teachers and school logistics, but some are dissatisfied. Firstly, there is a lack of communication between teachers and students, as teachers and students lack mutual understanding and timely and effective resolution of conflicts; secondly, the school has not established a perfect evaluation system to motivate and supervise the staff to improve their own business level, and again it is the teachers' low expectation of professional education service quality, low quality and even vicious circle, which leads to students in a passive position in college education. The last thing is that because of the low quality of professional education teaching, leads to the failure of the logistics work in colleges and universities, thus making the quality of school education low.

5.2 Analysis of students' actual feelings toward professional education teaching service quality

5.2.1 Analysis of overall feeling

As can be seen from [Table 3](#), the overall average feeling of college students about the quality of professional education teaching service is 3.11, which shows a rising trend year by year, which indicates that there are big problems in the quality of professional education teaching service in colleges and universities. This paper mainly adopts the questionnaire survey method to study the factors influencing college students' participation in professional education teaching quality.

5.2.2 Analysis of the actual feelings of students from different backgrounds about the quality of professional education teaching services

In the teaching process, professional education teachers should make reasonable plans for the course contents according to the actual situation of students. For example, students with a good foundation, strong self-control and high motivation can choose some easy-to-learn sports. For different groups of students, different forms and levels of classroom activities should be adopted to stimulate their interests and enthusiasm, while modern science and technology can be used to assist the teaching process, the use of multimedia technology to assist the professional education teacher to explain and demonstrate, so that students can intuitively perceive the relevant content.

5.2.3 Analysis of students' actual feelings at different levels

From the students' point of view, they have a high perception of the quality of teaching services and are able to grasp the problems they have in their daily activities. At the same time, the survey found that most

TABLE 3 Analysis of students' overall feelings about the quality of professional education services in college ($N = 770$).

Layer	Average	Standard deviation	Sort
Visible view	3.16	0.75	3
Reaction view	2.91	0.72	4
Assurance view	3.17	0.69	2
View of human nature	3.20	0.80	1
Whole	3.08	0.63	

TABLE 4 Analysis of the overall discrepancy in the quality of professional education services ($N = 770$).

Layer	Average	Standard deviation	Sort
Visible view	−0.94	0.93	2
Reaction view	−1.26	0.94	1
Assurance view	−0.85	0.80	3
View of human nature	−0.64	0.98	4
Whole	−0.93	0.78	

students hope that teachers can provide more targeted, practical and up-to-date courses, which requires teachers to continuously improve their own service consciousness and professional skills in the teaching process, which can enable students to contribute more actively and enthusiastically to the professional education course learning.

5.3 Analysis of the gap between the actual feelings and expectations of college students on the quality of professional education teaching services

5.3.1 Analysis of the overall discrepancy

As can be seen from [Table 4](#), the overall actual feeling of students on the quality of professional education teaching services is less than the expected value, and the satisfaction of students with the quality of teaching services is lower than the expected value. Through the survey, it is found that the main factor affecting the work enthusiasm and effect of professional education teachers in colleges and universities is their own quality. Therefore, in order to improve the image of professional education in society and enhance the overall level of the school, it is necessary to strengthen the construction and improvement of the teaching management system, so that teachers and students can really enjoy the services of modern higher education with high quality and efficiency, thus improving the quality of teaching ([Bezruk and Skorik, 2015](#)).

5.3.2 Analysis of the overall disparity of students with different variable backgrounds

The variability in the quality of professional education services is mainly for different groups of students, and there are

large differences between individual students. For those students who have good grades in school and have a good learning base, they receive more higher-quality and higher-level quality fitness sports programs, while for those who are not willing to participate in sports activities, there are negative effects and reduced positive emotions generated. Those who participate in sports or participate in other programs may also reduce the efficiency of sports due to lack of exercise physical fitness and other reasons. The most popular leisure activities among young people during the week are the use of technology and the neglect of physical exercise (Gil-Madrona et al., 2019b), which requires teachers to improve the quality of teaching services according to different groups of students, through the effective quality of teaching services to improve the communication skills between professional education teachers and students and the enthusiasm for sports participation.

6 Conclusion

The PZB service quality model is a service quality concept that aims to achieve service quality by satisfying customer needs and through information exchange and feedback. Its core is customer satisfaction, i.e., “making customers satisfied.” Adopting the PZB service model in the field of professional education teaching service can effectively improve teaching quality and also promote professional education teaching reform, which is conducive to the school and teachers to jointly improve the management ability and teaching level in all aspects of professional education classroom. Based on the service quality theory, this study constructs a mixed education service quality model for colleges and universities in China based on relevant literature and practical experience at home and abroad. On this basis, the model is verified through questionnaires, and the current situation and problems of mixed education service quality in colleges and universities are analyzed from three aspects of “expectation,” “feeling,” and “difference between feeling and expectation.” We also analyzed the current situation and problems of hybrid education service quality in three aspects: “expectation,” “feeling,” and “gap between feeling and expectation.” The innovations of this paper are as follows: firstly, there is no research on service quality in the field of mixed education in academia, and most of the existing studies are focused on higher education; secondly, although foreign countries have started to pay attention to the research on service quality, due to the different cultural background, the foreign service quality models cannot be directly applied to the actual situation in China, so we need to establish a service quality evaluation system that meets Chinese characteristics.

7 Limitations

Vocational education is a huge education system, and the cultural backgrounds and teaching strategies of each country and region are different. In addition, the data collected in this study come from the same Chinese school. The conclusions drawn may

not be generalizable to countries outside Asia. education system, and I hope that future foreign researchers can also do the same research to explore whether the education system in other different cultural backgrounds can get consistent conclusions and promote it.

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 humans were approved by Shenzhen Institute of Technology, Shenzhen, Guangdong Province, 518116, China. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

XD: Conceptualization, Data curation, Funding acquisition, Resources, Writing – original draft. LL: Supervision, Writing – review & editing, Methodology, Visualization, Validation. XL: Conceptualization, Data curation, Writing – original draft, Formal Analysis, Investigation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This paper was supported by Achievements of Key Research Base of Humanities and Social Sciences in Shenzhen: Skilled Society Research Center of Shenzhen Institute of Technology.

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

- Bezruk, V., and Skorik, Y. (2015). Multicriterial choice of telecommunications means using hierarchical analysis method. *Informat Telecommun Sci* 2, 18–24. doi: 10.20535/2411-2976.22015.18-24
- Biese, K. M., Winans, M., Fenton, A. N., Hernandez, M., Schaefer, D. A., and Bell, D. R. (2021). High school sport specialization and injury in collegiate Club-sport athletes. *J. Athl. Train.* 56, 1271–1277. doi: 10.4085/1062-6050-0021.21
- Borrego, M., Douglas, E. P., and Amelink, C. T. (2009). Quantitative, qualitative, and mixed research methods in engineering education. *J. Eng. Educ.* 98, 53–66. doi: 10.1002/j.2168-9830.2009.tb01005.x
- Fallah Jelodar, M. (2016). Prioritization of the factors affecting bank efficiency using combined data envelopment analysis and analytical hierarchy process methods. *Aust. J. Optim.* 2016:817. doi: 10.1155/2016/5259817
- Gil-Madrona, P., Aguilar-Jurado, M. Á., Honrubia-Montesinos, C., and López-Sánchez, G. F. (2019b). Physical activity and health habits of 17-to 25-year-old young people during their free time. *Sustainability* 11:6577. doi: 10.3390/su11236577
- Gil-Madrona, P., Hinojosa, L. M. M., Pérez-Segura, J. J., Sáez-Sánchez, M. B., and Poblete, G. Z. (2020). Scale of pedagogical authority meanings in the classroom (ESAPA) for Ibero-America built on the opinions of teaching students. *Teach. Teach. Educ.* 93:103079. doi: 10.1016/j.tate.2020.103079
- Gil-Madrona, P., Prieto-Ayuso, A., Silva, S. A. D. S., Serra-Olivares, J., Aguilar Jurado, M. Á., and Díaz-Suárez, A. (2019a). Hábitos y comportamientos relacionados con la salud en adolescentes durante su tiempo libre. *Anales de Psicología* 35, 140–147. doi: 10.6018/analesps.35.1.301611
- Goodson, A., Franklin, C. A., and Bouffard, L. A. (2021). Male peer support and sexual assault: the relation between high-profile, high school sports participation and sexually predatory behaviour. *J. Sex. Aggress.* 27, 64–80. doi: 10.1080/13552600.2020.1733111
- Liu, L., Liu, Y., Lai, I. K. W., and Liu, X., (2023) Research on the practice of VR entering the classroom-taking City University students of Macau as an example. In *International conference on blended learning* (pp. 191–201). Cham: Springer Nature Switzerland.
- Long, N., Linh, D., Anh, T., Lan, N., Thang, N., Trang, V., et al. (2020). Prioritizing the level of negative emotional coping strategies of cancer patients' family members by using extended hierarchical analysis method. *Decis Sci Lett* 9, 456–476. doi: 10.5267/j.dsl.2020.2.003
- Mišniakiewicz, M., and Krnáčová, P. (2021). Academic tutoring as a form of modern cooperation with the student and an element of improving the quality of teaching services. *Market Sci Res Organ* 39, 47–66. doi: 10.2478/minib-2021-0003
- Nowoświat, A., and Leszczyńska, M. (2016). Application of hierarchical analysis method to design the structural partitions with different material and structural solutions for window glazing. *Architect Civil Engineer Environ* 9, 95–104. doi: 10.21307/acee-2016-038
- Parida, B. B., and Baksi, A. K., (2011). Application of PZB service quality model in identifying service quality gap—a study on State Bank of India. *Siddhant-A Journal of Decision Making* 11, 1–10.
- Ritzer, E. E., Yang, J., Kistamgari, S., Collins, C. L., and Smith, G. A. (2021). An epidemiologic comparison of acute and overuse injuries in high school sports. *Inj. Epidemiol.* 8:51. doi: 10.1186/s40621-021-00344-8
- Shang, X., Xing, X., and Gao, P. (2014). Hierarchical analysis method and model based data of controlling parts of the body for clothing. *J. Fiber Bioengineer Informat* 7, 479–494. doi: 10.3993/jfbi.12201402
- Simcic Brønn, P. (2012). Adapting the PZB service quality model to reputation risk analysis and the implications for CSR communication. *J. Commun. Manag.* 16, 77–94. doi: 10.1108/13632541211197978
- Streetman, A. E., and Goodman, B. D. (2022). Does high school sports participation predict physical activity in college-age women? *Int. J. Exerc. Sci.* 11:2. doi: 10.1249/01.mss.0000876628.87592.2a
- Zubkova, T., and Tagirova, L. (2020). Using the hierarchical analysis method for decision making in a business activity. *J. Phys* 1553:012008. doi: 10.1088/1742-6596/1553/1/012008



OPEN ACCESS

EDITED BY
Pietro Cipresso,
University of Turin, Italy

REVIEWED BY
Enrique Ortega-Toro,
University of Murcia, Spain

*CORRESPONDENCE
Juan Ángel Simón-Piqueras
✉ juanangel.simon@uclm.es
Juan Carlos Padierna-Cardona
✉ juan.padierna@udea.edu.co

RECEIVED 09 October 2023
ACCEPTED 12 December 2023
PUBLISHED 11 January 2024

CITATION
Méndez-Hinojosa LM, Gil-Madrona P,
Zamorano-García D, Simón-Piqueras JÁ,
Padierna-Cardona JC, Flores-Ferro E,
González-Villalobos MF, Pavlovic R,
Maureira-Cid F, Gómez-Santos N,
Morales-Calvo S and Cárdenas-Rodríguez M
(2024) Adaptation of the SAGEFS scale on
attitudes toward gender equality in football
in the international school context.
Front. Psychol. 14:1265021.
doi: 10.3389/fpsyg.2023.1265021

COPYRIGHT
© 2024 Méndez-Hinojosa, Gil-Madrona,
Zamorano-García, Simón-Piqueras,
Padierna-Cardona, Flores-Ferro,
González-Villalobos, Pavlovic, Maureira-Cid,
Gómez-Santos, Morales-Calvo and
Cárdenas-Rodríguez. 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.

Adaptation of the SAGEFS scale on attitudes toward gender equality in football in the international school context

Luz Marina Méndez-Hinojosa¹, Pedro Gil-Madrona²,
David Zamorano-García², Juan Ángel Simón-Piqueras^{2*},
Juan Carlos Padierna-Cardona^{3*}, Elizabeth Flores-Ferro⁴,
Martín Francisco González-Villalobos⁵, Ratko Pavlovic⁶,
Fernando Maureira-Cid⁷, Natalia Gómez-Santos²,
Sonia Morales-Calvo⁸ and Magaly Cárdenas-Rodríguez¹

¹Departamento de psicología, Facultad de psicología, Universidad de Nuevo León, Monterrey, México, ²Faculty of Education of Albacete, University of Castilla-La Mancha, Albacete, Spain, ³Facultad de Educación Física, Recreación y Deporte, Institución Universitaria Politécnica Colombiano Jaime Isaza Cadavid, Medellín, Antioquia, Colombia, ⁴University of Bernardo O'Higgins, Santiago, Chile, ⁵University Health Sciences Center, University of Guadalajara, Guadalajara, Mexico, ⁶Faculty of Physical Education and Sports, University of East Sarajevo, East Sarajevo, Bosnia and Herzegovina, ⁷Departamento de Educación Física, Deporte y Recreación, Universidad Metropolitana de Ciencias de la Educación, Chile, Santiago, Chile, ⁸Faculty of Social Sciences and Information Technologies of Talavera de la Reina, University of Castilla-La Mancha, Talavera de la Reina, Spain

The objective of the present study is to evaluate the psychometric properties of the Scale of Attitudes toward Gender Equality in Football in the Context of Schools (SAGEFS) in the international context. This sample was formed by $N = 6,101$ students. The study was conducted by applying the SAGEFS. The model of the three factors correlated in the complete sample and for each country was correlated using AFC. The structural model was appraised by employing eight indices: the relative Chi-squared index; the goodness of fit index and its adjusted formula; the normal fit index; the comparative fit index; the standardized Quadratic Mean; and the Quadratic Mean Error of Approximation. To conclude, the results evidence the presence of psychometric properties that are indispensable for the measurement of attitudes toward gender equality in the context of schools.

KEYWORDS

instrument development, soccer, children's sports, gender, physical education, equality sport

Introduction

Mixed-gender education currently pre-dominates in most countries. However, coeducation is not that simple and requires an educational approach. In this respect, it is also possible to add aspects related to the socio-psychological dimension and several other variables, such as egalitarian beliefs, sexist beliefs, and the perception of sexist relationships. Coeducational education is not only the transmission of content that the student must learn, it also affects other constructs and attitudes along with the social structures involved (Valdivia-Moral et al., 2015).

In the case of teaching staff, Physical Education (PE) teachers do not provide real alternatives to combat the gender prejudices and inequalities present in sexist relationships to spread equality at school. This has been explained in studies carried out in several continents (Castillo et al., 2012; Valdivia-Moral et al., 2012, 2015, 2018).

For example, Soler (2009) and Sánchez-Hernández et al. (2018) indicate that soccer is one of the subjects taught with a higher gender culture load. This gender load remains dormant as regards PE classes about aspects such as space, materials, kit, language employed, feedback, evaluation, group formation, etc (Soler, 2009; Valdivia-Moral et al., 2018). It is, therefore, important that this material and content such as soccer are employed to work on egalitarian beliefs, which can be understood as the set of ideological principles of a person or group that promote the treatment of men and women, of the activities that they carry out, and of their achievements in an egalitarian way.

For this reason, it is important to teach pupils about gender equality in different knowledge areas (Brown, 2018). As Joy and Vasanathan (2022) explained, feminism fosters the erasing of the social discrepancy between what the patriarchy expects from women and reality. Due to this, education is the key to understanding the reality around us and the different social needs that might appear due to gender inequality (Joy and Vasanathan, 2022). The same was reported by Gaytán and Basso (2022), whose study included gender stereotypes and inequalities regarding not only gender but also ethnicity and sexuality. Men and women encounter many differences in sports because of the existence of stereotypes that create barriers and problems. All these problems are affected by a toxic masculinity that builds false expectations, which are reflected in violence against all collectives and practices, such as sports (Gaytán and Basso, 2022).

Upon analyzing the role of women in soccer, Woodhouse et al. (2019) and Ramírez and Restrepo (2018) recently showed that the participation of female players is much lower than that of their male counterparts, despite the increase in female players taking part in this sport (FIFA Activity Report, 2018). Although advances have been made regarding equality between men and women in soccer, many gender stereotypes and sexist beliefs continue to exist in this discipline. This is, in large part, because there are still many countries and cultures in which women do not have the same rights as men. At this point, it is important to highlight that those sexist beliefs are practices or attitudes that foment the different treatment of people according to their gender. In the case of soccer, this leads to a situation in which more importance is placed on the achievements of men, such that women's soccer has less impact. The study by Olmedilla et al. (2021) indicates that male youth soccer players manage the stress of competition and the evaluation of their performance better, and youth soccer players have higher scores in team cohesion than the boys.

As explained by Vannini and Fornssler (2011), there is also a problem related to women in sports: over-sexualization. Women in sports have long been sexualized because of their clothes or gestures and, depending on the sports they play, they are more sexualized than other women. For this reason, it is important to empower women to play sports, educate themselves, and be active no matter what (Mercer-Mapstone and Mercer, 2018).

Both egalitarian and sexist beliefs play a highly important role in physical activity, specifically in the case of soccer, since they promote situations in which girls do not participate in physical activities and do not, therefore, obtain the benefits that have been reported at a psycho-socio level (Bustamante-Sánchez and Del Coso, 2020). This has a direct consequence on the way they perceive their relationships with members of the opposite gender. In the case of girls, this perception is determined by the influence of social construction on body identity and the role that education, sports and health have on it (Walseth et al., 2017). In the case of boys, their perception of sexist relationships in this sport tends to be influenced by the idea that soccer is an activity that reinforces masculinity. Another key aspect to highlight is that although women have been excluded from practicing sports until very recent times, they have managed, in recent years (Salido-Fernandez, 2020), to challenge gender stereotypes to obtain a greater presence in a masculinized environment in both the elite (Sainz de Baranda, 2014) and the amateur field.

There is currently a lack of studies that analyze the role played by girls in soccer, along with their relationship with playing the game and with their male companions. The studies carried out by Ramírez and Restrepo (2018) in Latin America and Woodhouse et al. (2019) in Europe regarding the role played by women in soccer, therefore, stand out in this respect. Nevertheless, the studies by Sánchez-Hernández et al. (2018) and Soler (2009) in Spain carry out a more detailed analysis of what occurs in PE classes. This consequently shows that it is necessary to develop studies in different countries in order to be able to compare different cultures.

Upon focusing on the creation of instruments with which to evaluate equality in physical activity, it is possible to find tools such as that developed by Valdivia-Moral et al. (2015), which is intended for use within the teaching profession, or that of García et al. (2010), which evaluates students' attitudes toward equality in PE. The questionnaire concerning gender beliefs and stereotypes regarding physical activity and sport (GBSFAS) developed by Granda et al. (2018) is, meanwhile, more focused on the sphere of sport. After reviewing what the scientific literature presents in relation to this matter in question, the conclusion was reached that there are no instruments with which to evaluate attitudes related to gender equality in people of school age. Moreover, it is important to bear in mind that gender inequality and social exclusion might start at high school due to stereotypes based on gender, such as the ability to practice sports, to compete, and to be part of a sports team (Mercer-Mapstone and Mercer, 2018).

Keeping the current situation in mind, along with the state of the aforementioned variables, the objective of the present study is to adapt the Scale of Attitudes toward Gender Equality in Football in the School Context (SAGEFS) to the context of Spain, Ukraine, Morocco, Iran, Bosnia, Mexico, Colombia, Chile, and Brazil.

Materials and methods

Participants

The study was developed using a non-probabilistic sample. This sample comprised $N = 6,101$ students, 49.5% of whom were male and 50.5% of whom were female. The participants were an average

of 13.31 years of age, CI 95% [13.25, 13.34], with the youngest being 9 and the oldest being 20. The mean age of the male students was 13.17 (SD = 2.323), while that of the female students was 13.45 (SD = 2.369). Statistically significant differences were found upon comparing the ages of the male and the female students: $t = -4.714$, $p < 0.001$. The mean regarding the perception of playing soccer well was 5.31 (SD = 2.70). The mean for the male students was 6.20 (SD = 2.521), while that for the female students was (SD = 2.585). Statistically significant differences were found upon comparing the perceptions of male and female players playing soccer well: $t = 26.848$, $p < 0.001$.

With respect to the students' nationalities, 15.3% were from Spain, 1 5.8% were from Ukraine, 6.2% were from Morocco, 17% were from Iran, 14.8% were from Bosnia, 9.3% were from Mexico, 7.9% were from Colombia, 17.7% were from Chile, and 6.1% were from Brazil. Statistically significant differences were found upon comparing the average age of all the nationalities: $F = 297.911$, $p < 0.001$.

The frequencies and descriptive statistics of the socio-demographic and anthropometric variables are shown in [Table 1](#).

Instruments

The study was carried out by applying the SAGEFS, whose design and evaluation of psychometric properties were created for Spain. This scale includes 20 positive items with 5 Likert-type response options, ranging from “totally disagree” (1) to “totally agree” (5). The scale has three dimensions: (1) egalitarian beliefs, (2) sexist beliefs, and (3) perception of sexist relationships. The product of the adaptations and psychometric analyses in the nine countries in which the scale was validated was the Short Scale of Attitudes toward Gender Equality in Football in the Context of Schools (SSAGEFS), which contained 11 items [egalitarian beliefs (4 items), sexist beliefs (4 items), and perception of sexist relationships (3 items)].

Perceived motor competence was measured using the Spanish version of the AMPET [Achievement Motivation in Physical Education test ([Pérez et al., 2015](#))]. This factor was composed of seven positive items, with five response options ranging from ‘totally disagree’ to ‘totally agree’ on a Likert scale of 1 to 5.

The internal consistency index obtained after its application was $\alpha = 0.797$.

Procedure

To carry out this study, the coordinators contacted renowned researchers in various countries to inform them about the study and discover their interest in participating in it. In this way, the application of the scale to the context of each country was programmed, counting on those researchers who showed interest in participating in the study. For this, translators were hired, who translated the scale into the language of each country where it would be applied using the back translation method, which is a translation verification tool involving translating text into the desired language and then back into its original language. For this activity, the translators were asked to use a synonym or idiom if they identified any word that was not familiar in the context where the instrument would be applied. It is necessary to mention at this time that this article only intends to analyze the construct validity and reliability of the instrument for its adaptation to the international context; since the content validity was done in a previous study ([Gil-Madrona et al., 2022](#)) and the purpose of the present study was to adapt the SAGEFS scale regarding attitudes toward gender equality to the international school context. However, in accordance with [Ortega et al. \(2008\)](#), who, as part of the content analysis, established the judgment of experts and the understanding of the items by pilot sample, the procedure was applied to a sample of five subjects in each country; this sample had the same characteristics as the final sample. Each researcher participating in the study (one from each country) told the subjects to read the items carefully, answer them, and indicate if they had questions, doubts, or any comments about the writing or understanding of the instrument (qualitative degree of understanding of the test). All information was recorded in detail. Next, they were asked to write “1” for each item if the item could be understood within a single reading and 0 otherwise (quantitative degree of understanding). Based on the subjects' comments, corrections were made to the items. Regarding the quantitative values and the intercoder reliabilities, where the

TABLE 1 Frequencies and descriptive statistics of sociodemographic and anthropometric variables.

		All		Age		Weight		Height	
		N	%	M	SD	M	SD	M	SD
Country	Spain	933	15.3	10.83	0.881	42.60	9.451	150.37	9.148
	Ukraine	351	5.8	13.40	2.149	52.00	13.256	162.69	13.165
	Morocco	376	6.2	13.06	2.128	50.89	13.018	159.32	12.075
	Iran	1036	17	12.67	1.959	52.00	0.000	155.00	0.000
	Bosnia	900	14.8	14.00	2.581	57.31	14.434	168.21	12.280
	Mexico	566	9.3	14.81	2.144	55.55	13.173	159.67	14.081
	Colombia	482	7.9	13.42	2.145	49.79	15.483	150.71	15.676
	Chile	1082	17.7	14.40	1.871	55.09	10.747	160.27	11.149
	Brazil	375	6.1	14.19	2.091	53.58	11.708	159.44	10.578

N, number of cases; M, arithmetic mean; SD, standard deviation.

intercoder reliability was 0.85 or more, the comments made by the subjects were taken into account.

The schools that stated their intention to participate were given a training course via Skype in order to provide a detailed explanation of the objective of the research, what the study would consist of, the instructions and protocols required to utilize the questionnaires in each of the educational institutions, the databases to which they would have to upload the data eventually obtained, and informed consent forms that would have to be filled out by the families and the schools of those participating in the research. The aforementioned information was also sent in written format to those responsible in each of the participating countries.

Before handling the questionnaire, written permission to do so was obtained from the academic authorities corresponding to each of the schools in the countries studied. The students were invited to take part in the study on a voluntary basis, and a parent or tutor signed the corresponding informed consent form for underage participants. Then, the students were provided with an explanation of the object of the research, along with the relevant instructions, and were given the questionnaires. They did not receive any type of economic, material, or academic recompense for their participation. All the data were obtained in accordance with the ethical guidelines laid down in the Helsinki Declaration.

Data analysis

The model of the three factors correlated in the complete sample and for each country was correlated using CFA. The structural model was appraised by employing eight indices: the relative chi-squared index (χ^2/df), the goodness of fit index (GFI) and its adjusted formula (AGFI), the normal fit index (NFI), the comparative fit index (CFI), the standardized quadratic mean (SRMR), and the quadratic mean error of approximation (RMSEA). Values of $\chi^2/df \leq 2$; GFI, NFI, and CFI ≥ 0.95 ; AGFI ≥ 0.90 ; and SRMR and RMSEA ≤ 0.05 were considered to show a good fit, while values of $\chi^2/df \leq 3$; GFI, NFI, NNFI, and CFI ≥ 0.90 ;

AGFI ≥ 0.85 ; SRMR ≤ 0.10 ; and RMSEA ≤ 0.08 were considered to indicate an acceptable fit. The equivalence of the goodness of fit was verified by testing the difference between the chi-squared and relative differential chi-squared statistics ($\Delta\chi^2/\Delta df$), and the difference between CFI, NFI, and NNFI. Values of $p > 0.05$ for $H_0: \Delta\chi^2 = 0$, $\Delta\chi^2/\Delta df < 2$, ΔCFI , ΔNFI , and $\Delta NNFI \leq 0.01$ were considered to show the equivalence of the goodness of fit (Byrne, 2016).

The internal consistency of the three-correlated-factor model was verified by employing McDonald's coefficient ($\omega > 0.70$; Raykov and Hancock, 2005). The normality fit of the description of the distribution of the SAGEFS was contrasted by employing the Kolmogorov-Smirnov test with correction for Lilliefors significance correction. The multivariate normality of the structural model was determined by employing the Mardia coefficient, which should be ≤ 70 (Rodríguez and Ruiz, 2008).

The means per gender were compared by employing the Student's *t* statistic for independent samples. This parametric test was utilized despite the lack of a normal distribution owing to the fact that it is robust when this supposition is not fulfilled when the sample size is large (Poncet et al., 2016). The effect size (ES), as regards the difference in the means, was estimated by employing the Hedges *g* statistic, whereby values higher than 0.20 are considered to be a small ES, those of 0.40 are medium ES, and those of 0.80 are large (Hedges, 1981).

The concurrent validity was verified using Pearson's correlation coefficient. Since the normality was not fulfilled, the confidence intervals were calculated using the repetitive sample with the simulation of 2,000 random samples created using the percentile method (Bishara and Hittner, 2015).

Results

In the nine participating countries, the application of the scale to a pilot sample indicated that its application was viable.

TABLE 2 Goodness of fit indices for each sample and total sample.

Indices	Total	Spain	Ukraine	Morocco	Iran	Bosnia	Mexico	Colombia	Chile	Brazil
χ^2	641.106	135.249	105.451	84.007	—	169.763	108.623	50.816	115.149	74.100
<i>df</i>	41	41	41	41	—	41	41	41	41	41
<i>p</i>	< 0.001	<0.001	< 0.001	<0.001	—	< −0.001	<0.001	< 0.001	<0.001	< 0.001
χ^2/df	15.637	3.299	2.572	2.049	—	4.141	2.649	1.239	2.809	1.807
GFI	0.981	0.974	0.948	0.959	—	0.965	0.966	0.981	0.980	0.965
AGFI	0.970	0.958	0.916	0.935	—	0.944	0.946	0.970	0.969	0.944
NFI	0.968	0.896	0.823	0.875	—	0.918	0.889	0.945	0.946	0.947
NNFI	0.960	0.898	0.840	0.906	—	0.914	0.902	0.985	0.952	0.967
CFI	0.970	0.924	0.881	0.930	—	0.936	0.927	0.989	0.964	0.976
RMSEA (CL 90%)	0.048 (0.045,0.052)	0.050 (0.041,0.059)	0.067 (0.051,0.083)	0.053 (0.037,0.069)	—	0.059 (0.050,0.068)	0.054 (0.042,0.066)	0.022 (0.000,0.040)	0.041 (0.032,0.050)	0.046 (0.029,0.063)
SRMR	0.034	0.043	0.059	0.046	—	0.041	0.045	0.033	0.032	0.045

Indices of fit: χ^2 , Minimum value of the discrepancy function optimized for maximum likelihood; *df*, degrees of freedom of the statistic χ^2 ; *p*, probability of the statistic χ^2 for $H_0: \chi^2 = 0$ having a tail (goodness of fit); χ^2/df , relative chi-squared; GFI, goodness of fit index; AGFI, corrected goodness of fit index; NFI, normalized fit index; NNFI, non-normalized fit index; CFI, comparative fit index and RMSEA (90% CL); quadratic mean approximation error (estimation per interval with a confidence level of 90%) and SRMR, standardized mean quadratic error.

Validity and reliability of the three-correlated-factor model

The three-factor model was specified using its independent measurement residues. Table 2 shows the fit of the three-correlated-factor model to the complete sample and by country. With regard to the general sample, the fit was good for seven of the eight indices employed (GFI, AGFI, CFI, NFI, NNFI, RMSEA, and SRMR), although the relative chi-squared value was high ($\chi^2/\text{gl} = 15.637$). However, this index rarely obtains acceptable values for large samples (Barrett, 2007). Moreover, the factorial weights varied between 0.43 and 0.78 and were, therefore, greater than the minimum required value of 0.40 (Williams et al., 2010). The mean λ was .634, which was close to the desired value of 0.70 (Hair et al., 2014). Please note that it was necessary to eliminate nine items in order to improve the indicators of fit. The model obtained for the SAGEFS is shown in Figure 1.

With regard to the CFA per country, the goodness of fit values for Spain were good for five of the eight indices (χ^2/gl , GFI, AGFI, RMSEA, and SRMR), while the value for CFI was acceptable and the values for NFI and NNFI were close to those desired. In the case of Ukraine, the values were good for five of the eight indices (χ^2/gl , GFI, AGFI, RMSEA, and SRMR), while the values for CFI, NFI, and NNFI were close to those desired. In the Moroccan sample, the values were good for four of the eight indices (χ^2/gl , GFI, AGFI, RMSEA, and SRMR), while the NNFI and CFI values were acceptable and that for NFI was close to the desired value. With regard to the Bosnian sample, the values were good for four of the eight indices (GFI, AGFI, RMSEA, and SRMR) and acceptable for the χ^2/gl , NFI, NNFI, and CFI. In the Mexican sample, the values were good for five of the eight indices (χ^2/gl , GFI, AGFI, RMSEA,

and SRMR), the NNFI and CFI values were acceptable, and the NFI was close to the desired value. All eight goodness of fit indices were good for the samples obtained from Colombia, Chile, and Brazil.

As mentioned, the internal consistency was checked through the McDonald coefficient ($\omega > 0.70$; Raykov and Hancock, 2005), obtaining acceptable values for the Egalitarian Beliefs factor ($\omega \geq 0.772$) and the Sexist Beliefs factor ($\omega \geq 0.70$) but unacceptable values for the Perception of Sexist Relationships factor ($\omega < 0.70$).

Distribution of the factors of the SAGEFS

The distributions of EB, SB, and PSR, both in general and by gender, are shown in Table 3. The three factors have positive asymmetry or a long tail toward the right. None of them have a normal distribution (Table 4). Table 3, meanwhile, shows the descriptive statistics of the three factors of the SAGEFS separated into samples per country. In this particular case, the sample obtained for Iran was not considered for the PSR because data regarding that factor were not obtained.

TABLE 3 Descriptive statistics of the factors of the SAGEFS per country.

		<i>M</i>	<i>SD</i>	CL 95%		<i>Min</i>	<i>Max</i>
				<i>LI</i>	<i>LS</i>		
EB	Bosnia	20.88	4.415	20.59	21.17	6	30
	Ukraine	19.06	4.382	18.60	19.52	6	30
	Spain	24.89	4.702	24.58	25.19	6	30
	Morocco	24.36	4.531	23.90	24.82	9	30
	Mexico	23.01	4.685	22.63	23.40	6	30
	Colombia	24.09	4.467	23.69	24.49	6	30
	Chile	24.35	4.610	24.08	24.63	6	30
	Brazil	23.58	4.861	23.09	24.08	6	30
	Iran	16.76	4.416	16.49	17.03	6	30
SB	Bosnia	32.86	7.976	32.34	33.38	10	50
	Ukraine	30.28	6.859	29.56	31.00	11	48
	Spain	43.78	6.768	43.34	44.21	10	50
	Morocco	42.70	6.659	42.02	43.37	14	50
	Mexico	39.39	7.284	38.79	40.00	15	50
	Colombia	40.28	6.278	39.72	40.84	13	50
	Chile	41.77	6.492	41.38	42.16	10	50
	Brazil	41.80	6.304	41.16	42.44	14	50
	Iran	29.81	4.777	29.52	30.10	11	46
PSR	Bosnia	12.94	3.219	12.73	13.15	4	20
	Ukraine	12.70	2.643	12.43	12.98	6	20
	Spain	9.58	2.631	9.41	9.75	4	20
	Morocco	12.16	4.078	11.75	12.58	4	20
	Mexico	11.61	3.297	11.34	11.88	4	20
	Colombia	11.12	3.679	10.79	11.45	4	20
	Chile	11.81	3.559	11.60	12.02	4	20
	Brazil	13.05	3.250	12.72	13.38	4	20

M (CI 95%), arithmetic mean as regards accurate estimation and per interval with a confidence level of 95%; SD, standard deviation of sample; Min, minimum value; Max, maximum value; EB, egalitarian beliefs; SB, sexist beliefs; PSR, perception of sexist relationships.

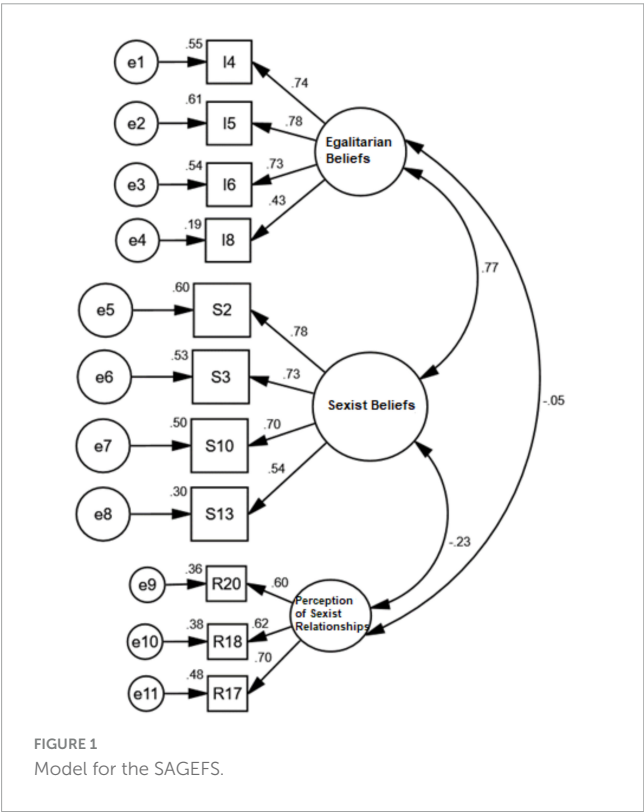


TABLE 4 Descriptive statistics of the factors of the SAGEFS at a general level and by gender.

Statistic	EB	SB	PSR	MEB	FEB	MSB	FSB	MPSR	FPSR
<i>M</i>	22.14	37.79	11.75	21.49	22.77	36.27	39.28	11.16	12.20
	CL95%	CL95%	CL95%	CL95%	CL95%	CL95%	CL95%	CL95%	CL95%
	(22.00, 22.27)	(37.57, 38.00)	(11.66, 11.84)	(21.30, 21.67)	(22.58, 22.96)	(35.94, 36.57)	(38.97, 39.59)	(11.03, 11.29)	(12.07, 12.34)

M (CI 95%), arithmetic mean as regards its accurate estimation and per interval with a confidence level of 95%, *SD*, standard deviation of sample; *Min*, minimum value; *Max*, maximum value; *KSL*, Kolmogorov-Smirnov test with the correction for Lilliefors significance; EB, egalitarian beliefs; SB, sexist beliefs; PSR, perception of sexist relationships; MEB, male egalitarian beliefs; MSB, male sexist beliefs; MPSR, male perception of sexist relationships; FEB, female egalitarian beliefs; FSB, female sexist beliefs; FPSR, female perception of sexist relationships.

TABLE 5 Intergroup comparison depending on gender.

	Groups	Descriptive			Levene		Student			<i>g</i>
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	<i>t</i>	<i>df</i>	<i>p</i>	
EB	Male	3019	21.49	5.38	0.004	0.948	−9.32	6099	0.000	0.24
	Female	3082	22.77	5.37						
SB	Male	3019	36.27	8.44	0.196	0.658	−14.04	6099	0.000	0.36
	Female	3082	39.28	8.35						
PSR	Male	2479	11.16	3.25	29.58	0.000	−10.76	5040.82	0.000	0.30
	Female	2586	12.20	3.63						

N, number of cases; *M*, arithmetic mean; *SD*, standard deviation. Levene Test: *F*; contrast statistic, *p*; probability of a tail. Student's *t*-test for independent samples: *t*, contrast statistic; *df*, degrees of freedom when *df* is < 6099 and *p*, probability of two tails; *g*, Hedges' effect size statistic. EB, egalitarian beliefs; SB, sexist beliefs; PSR, perception of sexist relationships.

Comparison of means by gender with respect to the three factors of the SAGEFS

Upon comparing the mean regarding the NSCS of the male and female students, a significant difference of $t = -9.32$, $p < 0.001$ was found. The mean for the female students was greater. The effect size of gender on the NSCS was small, $g = 0.24$. The mean of the female students was significantly greater than that of the male students regarding the NSCS, $t = -14.04$, $p < 0.001$, with a small effect size of gender on the NSCS, $g = 0.36$. The NR was significantly greater for the female students than for the male students, $t = -10.76$, $p < 0.001$. The effect size of gender on the NR was small, $g = 0.30$ (Table 5).

Concurrent construct validity regarding perceived motor competence

The correlations of the absolute value with the three external validity criteria varied from 0.111 to 0.324. The MC accordingly correlated to a greater extent with the SB factor than with NSCI and PSR. The correlations among MC, EB, and PSR were, therefore, very similar but of the opposite sign (Table 6).

Discussion

The present study focuses on the attitudes toward gender equality in soccer. The aim of this study was to adapt the Scale of Attitudes toward Gender Equality in Football in the School Context (SAGEFS) to the context of Spain, Ukraine, Morocco, Iran, Bosnia, Mexico, Colombia, Chile, and Brazil. The results shown

above made it possible to obtain a valid and reliable instrument that permits the analysis of intercultural differences.

The lack of studies regarding gender equality in the context of schools was mentioned by Martínez-Galindo (2006) a decade and a half ago, and the interest in this subject has continued to grow at a social level. In the case of soccer, studies on gender equality are even more scarce, particularly if we consider that it is an aspect of Physical Education that is closely associated with male gender or boys (Valdivia-Moral et al., 2018). All of that stated above, therefore, provides significance to this study, since it shows the definition of a valid measurement instrument for something as important as soccer. Factors such as skills, beliefs, and previous experience regarding gender in Physical Education, and how they are promoted, have also been taken into account (Dragutinovic and Mitrovic, 2019; Reina et al., 2019). The scale presented herein sheds light on three dimensions, which can be described as follows.

Egalitarian beliefs

This dimension is composed of a series of indicators that evidence the student's favorable attitude toward girls playing soccer in an egalitarian manner. It signifies that the same importance

TABLE 6 Correlations between SAGEFS and external validity criterion with CL 95%.

External validity criterion	SAGEFS		
	EB	SB	PSR
MC	−0.111*	−0.324*	0.124*
	(−0.135, −0.086)	(−0.348, −0.300)	(0.095, 0.153)

* $p < 0.001$. EB, egalitarian beliefs; SB, sexist beliefs; PSR, perception of sexist relationships; MC, perceived motor competence.

is placed on matches played by girls and boys, that the same amount of time and resources should be invested, and that women's soccer is considered to be a benefit to society in terms of reducing sexist beliefs. A negative view in this dimension would show an unfavorable attitude of the students towards the practice of football in an equal manner. This would imply that students consider that more resources should be invested in boys than in girls. On the other hand, they would consider that women's football has little impact on society.

Perception of sexist relationships

This indicates the degree to which the student perceives girls to be accepted in soccer, in the form of physical and psychological aggression.

The dimensions of this questionnaire were influenced by studies that promote attitudes toward equality Azorín (2017) and Garaigordobil and Durá (2005). With regard to the perception of sexist relationships, Azorín (2017) and Gil-Madrona et al. (2014), along with Gil-Madrona et al. (2017), have placed importance on the fact that boys should be more permissive in terms of allowing their female classmates to play this sport with them.

With regard to the evaluation of the psychometric properties of the instrument, the CFA provides evidence of its construct validity by obtaining a three-factor model with its residues of independent measures regarding both the sample as a whole and the samples obtained from each participating country. The fitness indicators were good and the original factorial structure was confirmed with appropriate factorial weights (λ 's > 0.40). It was, however, necessary to eliminate nine items in order to obtain better indicators.

In this respect, the CFA is similar to instruments like the CACEF created by Valdivia-Moral et al. (2015), which studies the coeducational beliefs of Physical Education teachers. The instrument in question was, however, validated within a population at Spanish schools and deals with different aspects of Physical Education. Other less recent instruments in the same line are those developed by Del-Castillo (2009) in Spain and Papaioannou (1998) in Greece.

The reliability of McDonald's omega was acceptable for the Egalitarian Belief ($\omega \geq 0.772$) and Sexist Belief ($\omega \geq 0.70$) factors but not for the Perception of Sexist Relations factor. This may be related to the idiosyncratic characteristics of each country regarding the relationship between the gender when playing soccer. As stated by Borges and Branco (2019), this is owing to the fact that each country's popular culture is molded into the different social relationships that appear in the context of schools. In this study, we, therefore, stress that the approach employed in this work brings together different cultures and religions in which the relationships between genders are different.

The analyses conducted led to the creation of a reduced version of the SAGEFS, which examines the attitudes toward gender in the context of schools by interpreting egalitarian and sexist beliefs, along with the perception of sexist relationships.

In conclusion, the results of the analysis of the different samples obtained for the SAGEFS evidence the presence of psychometric properties that are indispensable for the measurement of attitudes toward gender equality in the context of schools.

It is, however, necessary to carry out more psychometric studies in varied populations, in addition to analyzing the factorial invariance for country and gender. It would also be possible to carry out similar studies focused on the contents of Physical Education that are associated with gender, such as dance.

Data availability statement

The datasets presented in this article are not readily available because all requests to access the datasets should be directed to pedro.gil@uclm.es.

Ethics statement

The studies involving humans were approved by the Comisión de convivencia y ética escolar del CRA "Calar del Mundo". Castilla-La Mancha. Universidad de Castilla-La Mancha. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

LM-H: Methodology, Investigation, Writing—original draft, Formal analysis, Validation. PG-M: Writing—original draft, Writing—review and editing, Methodology, Investigation, Project administration, Supervision. DZ-G: Conceptualization, Writing—original draft, Validation, Formal analysis, Methodology, Writing—review and editing. JS-P: Writing—original draft, Formal analysis, Methodology, Validation, Conceptualization, Data curation, Writing—review and editing. JP-C: Writing—review and editing, Data curation, Formal analysis, Investigation, Validation. EF-F: Writing—review and editing, Data curation, Investigation, Project administration, Supervision. MG-V: Writing—review and editing, Data curation, Investigation, Conceptualization, Formal analysis. RP: Data curation, Investigation, Writing—review and editing, Project administration, Supervision. FM-C: Conceptualization, Data curation, Investigation, Project administration, Writing—review and editing. NG-S: Supervision, Conceptualization, Data curation, Investigation, Writing—review and editing. SM-C: Project administration, Data curation, Investigation, Supervision, Writing—review and editing. MC-R: Formal analysis, Investigation, Supervision, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Part of this research was funded by the Castilla-La Mancha University. Investigation Project: "Physical Activity and Health, aid for carrying out applied research projects, within the framework of the UCLM research plan, 85% co-financed by the European

Regional Development Fund.” Organic number (2022-GRIN-34290).

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

- Azorín, C. M. (2017). Actitudes hacia la igualdad de género en una muestra de estudiantes de Murcia. *Rev. Complut. Educ.* 28, 45–60. doi: 10.5209/rev_RCED.2017.v28.n1.48715
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Pers. Individ. Differ.* 42, 815–824. doi: 10.1016/j.paid.2006.09.018
- Bishara, A. J., and Hittner, J. B. (2015). Reducing bias and error in the correlation coefficient due to nonnormality. *Educ. Psychol. Meas.* 75, 785–804. doi: 10.1177/0013164414557639
- Borges, T., and Branco, A. (2019). The production of school violence from a cultural psychology perspective. *Estud. Psicol.* 40, 107–133. doi: 10.1080/02109395.2019.1569814
- Brown, L. E. C. (2018). “Post-colonial feminism, black feminism and sport,” in *The Palgrave Handbook of Feminism and Sport, Leisure and Physical Education*, eds L. Mansfield, B. Wheaton, B. Watson, and J. Caudwell (London: Palgrave Macmillan), 479–495. doi: 10.1057/978-1-137-53318-0_30
- Bustamante-Sánchez, A., and Del Coso, J. (2020). Physical activity and depression in young women. *J. Sport Health Res.* 12, 1–17.
- Byrne, B. M. (2016). *Structural Equation Modeling with AMOS Basic Concepts, Applications, and Programming*, 3rd Edn. New York, NY: Routledge. doi: 10.4324/9781315757421
- Castillo, O. D., Romero, S., González, T., and Campos, M. C. (2012). Gender equity in physical education: The use of information. *Sex Roles* 67, 108–121. doi: 10.1007/s11199-011-0103-5
- Del-Castillo, O. (2009). *Evaluación de los Factores Psicosociales y Didácticos Relacionados con la Equidad de Género en Educación Física*. Sevilla: Universidad de Sevilla.
- Dragutinovic, K., and Mitrovic, M. (2019). Differences in the attitudes of fourth grade pupils of the elementary school according to the teaching of physical education in terms of gender. *J. Anthropol. Sport Phys. Educ.* 3, 29–34. doi: 10.26773/jaspe.190106
- FIFA Activity Report (2018). *FIFA Club World Cup Saudi Arabia 2023™*. Available online at: <http://www.fifa.com> (accessed March 20, 2021).
- Garaigordobil, M., and Durá, A. (2005). Neosexismo en adolescentes de 14 a 17 años: Relaciones con autoconcepto-autoestima, personalidad, psicopatología, problemas de conducta y habilidades sociales. *Clin. Salud* 17, 127–149.
- García, R., Rebollo, M. A., Buzón, O., González, R., Barragán, R., and Ruiz, E. (2010). Actitudes del alumnado hacia la igualdad de género. *Rev. Investig. Educ.* 28, 217–232.
- Gaytán, M. S., and Basso, M. L. (2022). The political economy of puto: Soccer, masculinities, and neoliberal transformation in Mexico. *Frontiers* 43:28–61. doi: 10.1353/fro.2022.0011
- Gil-Madrona, P., Cachón-Zagalaz, J., Díaz-Suárez, A., Valdivia-Moral, P., and Zagalaz-Sánchez, M. L. (2014). Las niñas también quieren jugar: La participación conjunta de niños y niñas en actividades físicas no organizadas en el contexto escolar. *Movimiento* 20, 103–124.
- Gil-Madrona, P., Marina Méndez-Hinojosa, L., Cárdenas-Rodríguez, M., and Delgado, M. (2022). Attitudes towards Gender equality in Soccer in the school context in Spain (EAIGFU): Design and evaluation of its psychometric properties. *Rev. Psicol. Dep.* 31.
- Gil-Madrona, P., Valdivia-Moral, P., González-Villora, S., and Zagalaz-Sánchez, M. L. (2017). Percepciones y comportamientos de discriminación sexual en la práctica de ejercicio físico entre los hombres y mujeres preadolescentes en el tiempo de ocio. *Rev. Psicol. Deporte* 26, 81–86.
- Grandá, J., Alemany, I., and Aguilar, N. (2018). Gender and its relationship with the practice of physical activity and sport. *Apunts Educ. Fis. Deportes* 132, 123–141. doi: 10.5672/apunts.2014-0983.es.(2018/2).132.09
- Hair, J., Black, W., Babin, B., and Anderson, R. (2014). *Multivariate Data Analysis*. London: Pearson Education Limited.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *J. Educ. Stat.* 6, 107–128. doi: 10.2307/1164588
- Joy, A., and Vasanthan, R. (2022). The interconnection between feminism and education for women empowerment in laura bates' girl up. *J. Lang. Linguist. Stud.* 17, 2884–2890.
- Martínez-Galindo, C. (2006). *Motivación, Coeducación y Disciplina en Estudiantes de Educación Física*. Murcia: Universidad de Murcia.
- Mercer-Mapstone, L., and Mercer, G. (2018). A dialogue between partnership and feminism: Deconstructing power and exclusion in higher education. *Teach. High. Educ.* 23, 137–143. doi: 10.1080/13562517.2017.1391198
- Olmedilla, A., Cánovas, M., Olmedilla-Caballero, B., and Ortega, E. (2021). Características psicológicas relevantes para el rendimiento deportivo: Diferencias de género en fútbol juvenil. *Cuadernos Psicol. Dep.* 21, 127–137.
- Ortega, F. B., Ruiz, J. R., Castillo, M. J., and Sjöström, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. *Int. J. Obesity* 32, 1–11.
- Papaioannou, A. (1998). Goal perspectives, reasons for being disciplined, and self-reported discipline in physical education lessons. *J. Teach. Phys. Educ.* 17, 421–441. doi: 10.1123/jtpe.17.4.421
- Pérez, L. M. R., Collado, N. R., Sanz, J. L. G., Nieto, M. P., and Coll, M. V. G. (2015). GRAMI-2: Desarrollo de un test para evaluar la coordinación motriz global en la educación primaria. *Rev. Iberoamericana Psicol. Ejercicio Dep.* 10, 103–111.
- Poncet, A., Courvoisier, D. S., Combescure, C., and Perneger, T. V. (2016). Normality and sample size do not matter for the selection of an appropriate statistical test for two-group comparisons. *Methodology* 12, 61–71. doi: 10.1027/1614-2241/a000110
- Ramírez, E. Y., and Restrepo, J. A. (2018). El rol de la mujer: Una perspectiva sociocultural en el fenómeno del fútbol. *Ánfora* 25, 109–126. doi: 10.30854/anf.v25.n44.2018.401
- Raykov, T., and Hancock, G. R. (2005). Examining change in maximal reliability for multiple-component measuring instruments. *Br. J. Math. Stat. Psychol.* 58, 65–82. doi: 10.1348/000711005X38753
- Reina, R., Hutzler, Y., Iniguez-Santiago, M. C., and Moreno-Murcia, J. A. (2019). Student attitudes toward inclusion in physical education: The impact of ability beliefs, gender, and previous experiences. *Adapt. Phys. Activity Q.* 36, 132–149. doi: 10.1123/apaq.2017-0146
- Rodríguez, M. N., and Ruiz, M. A. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial. *Psicológica* 29, 205–227.
- Sainz de Baranda, C. (2014). Las mujeres en la prensa deportiva: Dos perfiles. *Cuadernos Psicol. Deporte* 14, 91–102.
- Salido-Fernandez, J. (2020). Olimpismo, género y comunicación: Una aproximación al deporte femenino y a su representación en los medios deportivos. *Comun. Género* 3, 173–182.
- Sánchez-Hernández, N., Martos-García, D., Soler, S., and Flintoff, A. (2018). Challenging gender relations in PE through cooperative learning and critical reflection. *Sport Educ. Soc.* 23, 1–12. doi: 10.1080/13573322.2018.1487836
- Soler, S. (2009). Los procesos de reproducción, resistencia y cambio de las relaciones tradicionales de género en la Educación Física: El caso del fútbol. *Cult. Educ.* 21, 31–42.
- Valdivia-Moral, P., López-López, M., Lara-Sánchez, A., and Zagalaz-Sánchez, M. L. (2012). Concepto de Coeducación en el Profesorado de Educación Física y metodología utilizada para su trabajo. *Movimiento* 18, 197–217. doi: 10.22456/1982-8918.29593

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.

- Valdivia-Moral, P., Molero, D., Díaz-Suarez, A., Cofre, C., and Zagalaz-Sánchez, M. (2018). Coeducational methodology used by physical education teachers and students' perception of it. *Sustainability* 10:2312. doi: 10.3390/su10072312
- Valdivia-Moral, P. A., Molero, D., Campoy, T. J., and Zagalaz, M. L. (2015). Coeducational ideas of physical education teachers: Psychometric properties of a scale. *Rev. Int. Med. Cienc. Act. Fís. Deporte* 15, 269–288 doi: 10.15366/rimcafd2015.58.005
- Vannini, A., and Fornssler, B. (2011). Girl, interrupted: Semenya's body, gender verification testing and public discourse. *Cult. Stud. Crit. Methodol.* 11, 243–257.
- Walseth, K., Aartun, I., and Engelsrud, G. (2017). Girls' bodily activities in physical education How current fitness and sport discourses influence girls' identity construction. *Sport Educ. Soc.* 22, 442–459. doi: 10.1080/13573322.2015.1050370
- Williams, B., Brown, T., and Onsmann, A. (2010). Exploratory factor analysis: A five-step guide for novices. *Aust. J. Paramed.* 8, 1–13. doi: 10.33151/ajp.8.3.93
- Woodhouse, D., Fielding-Lloyd, B., and Sequerra, R. (2019). Big Brother's little sister: The ideological construction of women's super league. *Sport Soc.* 22, 1–20. doi: 10.1080/17430437.2018.1548612



OPEN ACCESS

EDITED BY
Martin Musalek,
Charles University, Czechia

REVIEWED BY
Boris Banjevic,
University of Montenegro, Montenegro

*CORRESPONDENCE
Pedro Gil-Madrona
✉ Pedro.Gil@uclm.es

RECEIVED 27 September 2023
ACCEPTED 02 February 2024
PUBLISHED 14 February 2024

CITATION
Gil-Madrona P, Losada-Puente L, Mendiri P,
Sá C, Silva IP and Saraiva L (2024) Is it possible
to identify physical-motor profiles of
preschool children on their association with
selected biosocial factors?
Front. Psychol. 15:1302402.
doi: 10.3389/fpsyg.2024.1302402

COPYRIGHT
© 2024 Gil-Madrona, Losada-Puente,
Mendiri, Sá, Silva and Saraiva. 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.

Is it possible to identify physical-motor profiles of preschool children on their association with selected biosocial factors?

Pedro Gil-Madrona^{1*}, Luisa Losada-Puente², Paula Mendiri²,
César Sá^{3,4}, Inês P. Silva³ and Linda Saraiva^{3,4}

¹Department of Didactics on Physical, Artistic and Music Education, Faculty of Education, Albacete, University of Castilla La Mancha, Ciudad Real, Spain, ²Department of Specific Didactics and Methods of Research and Diagnosis in Education, Faculty of Educational Sciences, University of Coruña, A Coruña, Spain, ³Escola Superior de Educação de Viana do Castelo, Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal, ⁴Center for Research and Innovation in Education, Porto, Portugal

Biosocial factors play a crucial role in the physical-motor development (PMD) of children during the preschool age. The present study aims to identify physical-motor profiles throughout preschool age (3–6 years) and explore associations between profiles and selected biosocial factors such as age, sex, prematurity, weight, height, BMI, and participation in extracurricular physical activities. Data from 412 typically developing children (46.6% girls and 53.4% boys), aged 35–71 months ($M = 51.21$, $SD = 10.47$) was collected using the Psychomotor Activities Checklist and specifically the scale of Psycho-Motor Aspects. Cluster analysis made it possible to define four different childhood PMD profiles. High PMD; High PMD except left laterality; medium-low PMD; and low PMD. High PMD profile includes older children, with anthropometric measurements closer to the WHO recommendations, fewer preterm children, and greater participation in extracurricular physical activities. Low PMD profile includes younger children, with weight slightly above and height slightly below the WHO recommendations and low participation in extracurricular physical activities. This study allows us to identify specific trends that may be decisive for the motor development of children throughout preschool age, highlighting selected biological variables and participation in extracurricular physical activities.

KEYWORDS

physical-motor development, preschool children, age, sex, prematurity, body mass index, extracurricular physical activities

1 Introduction

All healthy children have the potential to develop and learn motor skills during early childhood. Nevertheless, research shows that children may have different developmental pathways to achieve motor proficiency (Hadders-Algra, 2018; Boonzaaijer et al., 2021). Understanding and explaining this inter-individual variability is still challenging in contemporary research.

In the past, the theoretical and empirical approach to this topic has been portrayed by the “nature versus nurture” dichotomy, suggesting that genetics (nature) or environmental factors (nurture) were the primary drivers of motor development. Researchers abandoned this approach and recognized that acquiring a broad motor repertoire during early childhood is determined by the complex interaction between biological factors and environmental conditions in which a child lives and grows (Krebs, 2009; Gabbard, 2014). Among various contemporary approaches, Bronfenbrenner’s Bioecological Theory (Bronfenbrenner, 2005) provides a useful theoretical and conceptual framework for better understanding the biosocial influences on children’s motor development.

In the last decade, several reviews and systematic reviews support that a wide range of biosocial factors influences children’s motor development (Lubans et al., 2010; Barnett et al., 2012; Holfelder and Schott, 2014; Livonen and Sääkslahti, 2014), that can contribute to the variability observed in children’s acquisition and refinement of motor skills. Heterochrony in children’s motor development can be explained by the rate of growth and maturation (Drenowatz and Greier, 2019; Werneck et al., 2020; Pereira et al., 2021). For instance, growth hormone has been found to impact the central nervous system, including the brain. It plays a role in developing neural structures and may influence cognitive and motor functions. Variations in the timing of hormonal changes during childhood can contribute to differences in cognitive and motor development among children (Webb et al., 2012).

Variations in children’s motor development can also be explained by age, gender, physical activity, and preschool-based programs (Livonen and Sääkslahti, 2014). The systematic review by Barnett et al. (2012) on correlates of motor competence in typical development children and young people (3–18 years), using an ecological approach, reinforces that increasing age was the most consistent correlate of all aspects of motor competence. The weight status (healthy), sex (male), and socioeconomic background (higher) were consistent correlates for specific motor skills. In the same study, the authors also emphasize the positive correlation between physical activity and composite and motor coordination skills, despite the relationship that is not consistent for object control and locomotion skills.

Regarding the age effect, research showed that age was positively associated with locomotor, object control, and stability skills (Barnett et al., 2012; Valentini et al., 2016; Bolger et al., 2018; Li et al., 2023); however, the magnitude of its effect tends to decrease or even disappear with age, given that children reach the expected motor proficiency in some motor skills. On this issue, it is essential to point out that age reflects the child’s biological and neurological maturity and the accumulated effects of stimulation and environmental factors. It is also reasonable that the interaction of maturation and environmental effects may help to explain the increased variability in motor development with advancing age.

The relationship between sex and motor skills in preschool children has been studied (e.g., Saraiva et al., 2013; Valentini et al., 2016; Navarro-Patón et al., 2021). For instance, Saraiva et al. (2013) concluded that in preschool age, there were significant differences between boys’ and girls’ performance in grasping, visual-motor integration, and object manipulation subtests. Girls had superior performance in grasping and visual-motor integration skills, while boys performed better in object manipulation skills. This different motor profile between boys and girls is consistent with most studies

found in the literature (Saraiva et al., 2013; Valentini et al., 2016; Navarro-Patón et al., 2021; Escolano-Pérez et al., 2022), and it has been attributed to environmental and educational influences but also to biological factors such as advanced neurological development favoring girls, and to some morphological characteristics favoring boys (Ikeda and Aoyagi, 2008; Gabbard, 2011). However, it is essential to note that the magnitude of the sex effect varies according to age and motor skill specificity. In the study by Saraiva et al. (2013), it was found that boys’ advantage in object manipulation skills becomes progressively greater throughout preschool age, which suggests a progressive modulation effect of motor experiences across childhood. These differences between boys and girls in motor skills at preschool age can be explained by support and opportunities for structured and unstructured physical activity at home, school, and the community. On this issue, Valentini et al. (2016) warn that opportunities for practicing and learning motor skills must be structured and developmentally appropriate for the children, providing an adequate challenge to improve motor proficiency. Previous systematic reviews have shown that motor intervention programs implemented throughout childhood are crucial to improving motor competence and increasing participation in physical activity (Hardy et al., 2014; Holfelder and Schott, 2014; Wick et al., 2017; Coppens et al., 2021; Jaakkola et al., 2021).

Researchers have also examined the relationship between motor competence and potential health benefits in children and adolescents (Lubans et al., 2010; Jones et al., 2020; Xin et al., 2020). From this investigation emerges the idea that motor competence is positively associated with physical activity and is inversely correlated with weight and BMI status. However, the strength of these relationships seems to vary across age and specificity of motor skills. For example, overweight and obese children tend to have poorer performance than their non-overweight peers, and these differences appear to be more apparent for locomotor skills than object manipulation skills (Okely et al., 2004; Cliff et al., 2012). Likewise, underweight children tend to have less motor coordination and muscular power than normal-weight children, but underweight and normal-weight youth generally perform better than overweight or obese youth (Lopes et al., 2018; Verbecque et al., 2022).

Finally, it is also important to mention that distal biological factors such as gestational age and birth weight can also explain motor variability. Lower birth weight and gestational age are strongly related to poorer motor outcomes in the first years of development (e.g., Oudgenoeg-Paz et al., 2017). Studies have shown that premature birth (birth before 37 weeks of gestation) can have various effects on a child’s development, including motor skills, muscle strength, and physical activity levels than term-born children (FitzGerald et al., 2021; Karageorgi et al., 2021). This adverse effect of preterm birth on a child’s motor performance may persist into adulthood (Moreira et al., 2014; Allotey et al., 2018), depending on the extent of prematurity and associated complications.

Based on the assumption that the influence of biosocial factors on motor proficiency throughout preschool age varies according to age and motor specificity, the present study sought to identify physical-motor profiles throughout preschool age (3–6 years) and explore associations between profiles and selected biosocial factors (age, sex, prematurity, weight, height, BMI, and participation in extracurricular physical activities).

TABLE 1 Tasks evaluated and Cronbach's alpha (95% confidence interval) for each motor domain of the PSAS scale.

Dimension	Item/task	α (original)	α (present study)
Laterality (LAT)	Grasps objects with the left hand. Grasps objects with the right hand. Hits objects with the left leg. Hits objects with the right leg. Throws objects with the left hand. Throws objects with the right hand.	0.57 (0.56–0.60)	Original LAT = 0.83 (0.81–0.82) Right LAT = 0.88 (0.89–0.80) Left LAT = 0.91 (0.84–0.92)
Dynamic Coordination	Is able to roll on a surface. Jumps with both feet together. Jumps with one foot. Is able to move sideways. Is able to walk backward. Runs freely without difficulty.	0.77 (0.75–0.81)	0.89 (0.85–0.90)
Balance	Maintains balance by walking in a straight line. Maintains balance by walking along a curved line. Maintains balance by walking on a trestle. Maintains balance by walking on a bench. Is able to maintain a balanced posture.	0.87 (0.86–0.92)	0.86 (0.77–0.82)
Motor Execution	Is able to use the materials correctly. Is able to jump over obstacles. Is able to circle around obstacles.	0.65 (0.63–0.67)	0.86 (0.78–0.81)
Tonic Postural Control	Moves following the indicated rhythm. Runs and is able to stop at a signal. Lies down flat on the back.	0.60 (0.58–0.63)	0.88 (0.84–0.88)

2 Materials and methods

2.1 Sample

A convenience sample of 412 typically developing children aged between 35 and 71 months ($M = 51.21$, $SD = 10.47$) was recruited from 19 children's classrooms from different kindergartens in the province of Albacete. Children included in the study met the following criteria: age between 36 and 71 months, absence of any known intellectual, physical, or emotional disabilities, as well as without special educational needs, as proven by records of special education teams. The sample has been non-probabilistic, the sampling technique has been used on purpose, as voluntary collaboration has been requested in all schools in Albacete province, and we have finally worked with the schools that agreed to participate.

In total, 46.60% of the sample were girls, and 53.40% were boys. The average weight ranged from 10.90 to 28 kg ($M = 17.10$, $SD = 2.63$), and height ranged from 82 to 126 cm ($M = 103.72$, $SD = 7.46$). The body mass index (BMI) ranged from 12.18 to 21.43 ($M = 15.86$, $SD = 1.43$). Most children were born at full term (i.e., not premature: $n = 323$, 78.4%). There were 67 missing values. Only 21.8% of the children played extracurricular physical activities ($n = 90$), mainly football, and to a lesser extent, swimming, tennis or paddle, martial arts, rolling, or dancing, versus a majority who did not practice ($n = 322$, 78.16%).

The parents or legal guardians of the preschool children were informed about testing procedures, the guarantee of anonymity and confidentiality of data, and corresponding written consent was obtained.

2.2 Measures

2.2.1 Checklist of psychomotor activities- psychomotor aspects scale

In this study, the Psychomotor Activities Checklist (CPA; [Romero Martínez et al., 2018](#)) was used to measure the physical-motor development of children, and specifically the scale of Psycho-Motor Aspects (PSAS), composed of five dimensions: Laterality (LAT, six items), Dynamic Coordination (DC, six items), Tonic-Postural Control (CTP, three items), Motor Execution (ME, three items), and Balance (BAL, five items). LAT is the dominance of one side of the brain in controlling activities or functions, measured by tasks such as "hitting objects with the left leg." DC is the ability to apply skills, knowledge, and situational awareness, measured by tasks such as "jumping with one foot." ME is composed of overt and volitional movement, for example, "lies down with his back straight." BAL is a state of equilibrium or equipoise, measured with actions such as "maintains balance by walking on a curve."

The tasks in each motor domain can be found in [Table 1](#) alongside the reliability values, contrasting the original scale with that of the present study. It should be noted that in the present study, the dimension of laterality was explored in two sub-dimensions, given that the literature points out that its establishment, including handedness, typically occurs during early childhood, around the age of 3 to 5 years ([Scharoun and Bryden, 2014](#); [Bondi et al., 2020](#)).

Furthermore, from the analyses carried out, it could be observed that, although Cronbach's Alpha values were acceptable when separating the items relating to left and right laterality, an excellent index was obtained in both cases.

Participants were evaluated by their teachers using a 5-point Likert scale from 1 (never) to 5 (always), depending on their ability to perform the proposed task on each item. Further details on the CPS's psychometric properties (validity and reliability) may be consulted in [Romero Martínez et al. \(2018\)](#).

2.2.2 Anthropometric measures

Statures and weight were measured using standard procedures ([Lohman et al., 1988](#)), and the Body mass index was calculated as weight (BMI = kg/m²).

2.2.3 Birth and extracurricular physical activities data

A questionnaire was applied to parents to collect the child's birth date, gestational age, and extracurricular physical activity data. The prematurity definition for this study is considered in the reports mentioned above (children born before week 37).

2.3 Procedures

In the first moment, different kindergartens and respective teachers in the province of Albacete, Spain, were contacted. An explanatory document for the research's purposes was presented to the school's directors, and the parents were invited to participate. Informed consent was requested from parents who agreed to participate. This moment occurred between December 2022 and February 2023.

In the second moment, teachers were training to use the CPA instrument, more specifically the PSAS, and posteriorly, the teachers evaluated the children with the help of a research team member. This work occurred in children's natural context, that is, in their schools and usual work classrooms, and in the presence of reference adults with whom they are familiar (teachers, internship students, and researchers). The researchers and teachers observed the children at three moments of the school day: in the classroom, in the gym, and during recess. The parents were asked to answer a questionnaire about their children's birth date, gestational age, and extracurricular physical activity.

The third moment, the last, was dedicated to joining all data and realizing all the analysis.

This research was first submitted for evaluation by the Ethics Committee of the University of Castilla La Mancha, which issued a favorable report. Subsequently, approval was sought and obtained from the families for the children to participate in this study.

2.4 Data analysis

Firstly, descriptive statistics (means, standard deviations, skewness, and kurtosis) and Pearson correlations were calculated among all measures. Pearson's *r*-value varies between [−1 and 1]. Values closer to 1 indicate a higher degree of relationship between variables ([David and Sutton, 2004](#)).

Secondly, hierarchical agglomerative cluster analysis was used to identify groupings of students in the data set according to their scores on the right and left laterality, balance coordination, motor performance, and tonic. Previously, data with missing values on any

observation items were removed, and those with missing values only on the socio-demographic data were kept. For carrying out the analyses, a series of steps have been carried out, whereby the cases start as individual clusters and, step by step, the most similar clusters are joined to give rise to a cluster containing all cases until obtaining a cluster containing all cases ([Mojena, 1977](#); [Clarworthy et al., 2005](#)). The analyses were performed with Machine Learning in Jasp 0.17.1.0, using neighborhood-based clustering (K-means clustering or generalized Lloyd Algorithm, GLA) method in which the squared Euclidean distance is used to measure dissimilarity between a data point and its cluster representative, considering the cohesion degree of the neighborhood of an object and the coupling degree between neighborhoods of objects ([Cao et al., 2009](#); [Lai and Huang, 2011](#); [Schubert, 2022](#)). This method allows us to reduce the computational complexity of Ward's method ([Lai and Huang, 2011](#)).

We used the stopping rules to determine the final cluster solution ([Mojena, 1977](#); [Clarworthy et al., 2005](#); [Baidari and Patil, 2020](#)) by examining the agglomeration schedule and the Elbow criterion ([Schubert, 2022](#)). An inconsistent increase in the dissimilarity measure suggests that the groups that came together at that stage were quite different and that the grouping should cease at an earlier stage. To prevent the researchers' bias associated with the subjectivity of this method, we also incorporate formal rules and equations to determine the number of clusters in a sample ([Clarworthy et al., 2005](#)). Therefore, we have relied on a combination of fit indices and methods and the interpretability of the solutions based on theory ([Xu and Núñez, 2023](#)). Thus, the pseudo-F statistic ([Calinski and Harabasz, 1974](#), as cited in [Baidari and Patil, 2020](#)) has been included. Large values of the Calinski–Harabasz pseudo-F index indicate distinct clustering, whereas small values indicate less clearly defined cluster structure, and the Silhouette index ([Kaufman and Rousseeuw, 1990](#)), a coefficient (range [−1, 1]) that reflects the closeness of a point within the cluster and in the neighboring clusters. We also considered the maximum diameter, the minimum separation, and the Dunn Index ([Clarworthy et al., 2005](#)). Compact and well-separated clusters in the data set result from the small diameter of the clusters and the large distance between the clusters; thus, the Dunn index is maximized. A final criterion was to consider adequate and feasible only those profiles that are composed of more than 5% of the sample reflecting enough profile extraction ([Xu and Núñez, 2023](#)).

After identifying the optimal number of profiles, additional analyses were conducted to test the validity of the classification. The analysis of variance (ANOVA) test was carried out to test whether there were statistically significant differences among clusters on all measures and to examine whether the clusters differed in age, weight, height, and BMI. The effect size (η^2) was calculated and interpreted according to ([Cohen's 1988](#), as cited in [Xu and Núñez, 2023](#)), that is $\eta^2 = 0.01$ ($d = 0.20$) a small effect size; $\eta^2 = 0.059$ ($d = 0.50$), a medium effect size; and $\eta^2 = 0.138$ ($d = 0.80$), a large effect size. Tukey's *post-hoc* test was used for pairwise differentiation. The chi-square test was performed to test the differences related to other biosocial qualitative variables such as sex, prematurity, and participation in extracurricular physical activities. Cramer's V was calculated as a correction that is applied to the Chi-square coefficient. The closer this index is to 1, the greater the association between variables ([David and Sutton, 2004](#)). All values were considered significant for a value of *p* of less than 0.05 (95% of confidence level).

3 Results

3.1 Descriptives

Table 2 presents means, standard deviations, skew, kurtosis, and Pearson correlations among all measures. All of them were found to be significantly positively correlated.

3.2 Hierarchical clustering

The first agglomerative hierarchical cluster analysis optimized according to the BIC index provided a solution consisting of nine clusters. Figure 1 shows through the Elbow Method Plot the percentage of variance explained by each cluster, gathered in the first four clusters. The change in the inertia value (“elbow”) can be observed in the line represented by WSS (sum of squared errors) that breaks off in cluster 4.

Furthermore, the Silhouette, CHI, and Dunn Index are presented in Table 3. Despite having an acceptable 9-clusters solution, additional refinements were carried out to test the improvements offered by lower clustering, given the aim of extracting a more compact solution. Specifically, an additional test was carried out with five, four, and three clusters (Table 3).

Table 3 shows the index extracted for different solutions. It was observed that R2 decreased as the number of clusters dropped, while the Silhouette Index improved considerably accordingly, reaching acceptable values from 5-clusters solution upwards (>0.25). Similarly, the CHI Index reported more satisfactory results in solutions composed of a smaller number of clusters, thus suggesting a better-defined structure in these latter solutions. As for the Dunn index, the overall results are less satisfactory since the diameter of the clusters increases—it is desirable for them to be more compact—although the separation between them is maintained from five clusters onwards. The lack of improvement in the reduction to three clusters is identified here, so the 4-cluster solution was chosen.

Figure 2 presents two graphs of the t-distributed stochastic neighbor embedding (t-SNE) plots. These plots show the clustering between cases according to their distance in the 2D map. The comparison between the plots for nine clusters (left) and four clusters (right) shows the differences in the island groupings of the datasets. Complementarily, the cluster mean plot allows us to clearly identify the four groupings that are extracted in each cluster: High PMD (cluster 1); High PMD, except left laterality (cluster 2); medium-low PMD (cluster 3); and low PMD (cluster 4).

3.3 Profile characterization

Table 4 presents the mean values and deviation of the participant’s scores in each cluster. An analysis of variance is included, showing statistically significant differences between clusters in all the measures used, with a large effect size ($\eta^2 > 0.14$).

Post-hoc tests (Table 5) reported the presence of statistically significant differences between all contrast groups in all variables, except between the High PMD group (cluster 1) and High PMD, except left laterality (cluster 2) in right laterality, and between the latter concerning low PMD (cluster 4) in left laterality.

Description of the four clusters (profiles) regarding biosocial variables are presented in Table 6.

Profile 1 had high values in all dimensions of physical-motor development, which placed it above the rest of the clusters. It comprised an equal number of boys and girls, with the lowest number of cases of prematurity (only 4 cases, 3.39%), of whom more than 30% played extracurricular physical activities. The mean age in this group was 57.7 months (SD = 7.5), and concerning weight and height, in this case, they were approximately in line with WHO standards (see Table 5), although slightly below in terms of height.

In profile 2, high values were found in all dimensions of physical-motor development except for left laterality. It stood out, above all, in right laterality compared to the rest of the groups. The mean age of this group was 50.6 months (SD = 10.16), and five children (5.26%) were preterm. The weight and height of the group were higher than WHO recommendations (Table 5), and participation in extracurricular physical activities was low (79.84% did not participate in extracurricular physical activities).

The group comprising Profile 3 obtained intermediate-low values in most of the analyzed variables. Balance stood out most negatively, followed by left laterality, while motor execution and right laterality did so more positively, together with tonic. A higher number (almost 20%) of boys versus girls composed this profile, mostly born at term (92.4%), with an average age of 48.18 months (SD = 10.15). The average childhood weight grouped in this profile was slightly higher than the standards proposed by World Health Organization (2006; Table 5), while height was slightly lower. Physical activity in this group was higher than one-fifth of the total number of children.

Profile 4 had the lowest scores on all variables analyzed. This group consisted of more boys than girls (a difference of more than 20%) with a mean age of 44.33 months (SD = 10.07). Weight in this group was slightly above WHO standards, while height was lower. This was the group with the highest number of cases of students born

TABLE 2 Descriptive statistics and Pearson correlations.

Measure	1	2	3	4	5	M	SD	Skewness	Kurtosis
1. Right laterality						4.26	0.91	−1.23	0.93
2. Left laterality	0.24**	-				3.03	1.12	0.35	−0.79
3. Coordination	0.56**	0.40**	-			3.83	0.93	−0.86	0.47
4. Balance	0.53**	0.44**	0.80**	-		3.86	0.97	−0.80	0.22
5. Motor Execution	0.56**	0.31**	0.65**	0.66**	-	4.38	0.76	−1.49	2.09
6. Tonicity	0.61**	0.42**	0.68**	0.63**	0.72**	3.85	0.95	−0.98	0.21

n = 412. **p < 0.001.

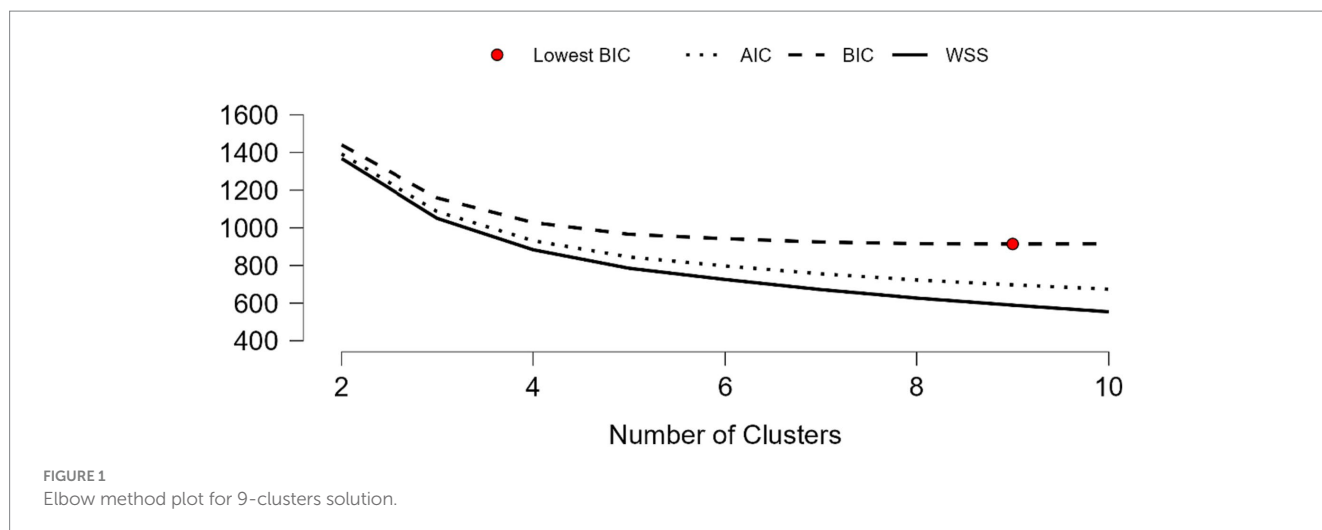


TABLE 3 Number of clusters based on agglomerative hierarchic analysis and fit index.

Clusters	N	R ²	Silhouette	CHI	Maximum diameter	Minimum separation	Dunn index
9	412	0.76	0.25	160.12	4.75	0.27	0.06
5	412	0.68	0.28	217.88	5.15	0.40	0.08
4	412	0.65	0.28	243.56	5.67	0.40	0.07
3	412	0.57	0.30	275.50	6.66	0.40	0.06

pre-term (15%) and in which they practiced the least extracurricular physical activities (92.73% did not do so; Table 7).

3.4 Relations of the four profiles to the biosocial variables

Finally, the relationships between each of the profiles described and some of the biosocial variables were analyzed (Table 8). No statistically significant differences were found concerning sex ($p=0.145$) nor concerning prematurity ($p=0.066$), but there were significant differences in terms of participation in extracurricular physical activities ($p<0.001$).

Although no overall differences were found for sex or prematurity, additional checks revealed statistically significant differences for prematurity between clusters 1 and 4 (Chi-Square = 6.79, $p<0.05$) and clusters 2 and 4 (Chi-Square = 3.57, $p<0.05$)—but the reliability of this result is scarce since the data used to calculate it is too small, as can be seen in (Table 8). Also, Table 9 shows statistically significant differences in participation in extracurricular physical activities in all clusters, except for 2 and 3.

In Table 10, the variance analysis showed differences between the groups by age between clusters 1 and the other three, with the oldest children being in cluster 1; there were also differences between clusters 2 and 4, with the average age being lower in cluster 4. As for the weight variable, there were differences between clusters 1 and 4, with the mean being higher in cluster 1 and, between clusters 2 and 4, higher in cluster 2. As for the height, differences between clusters 1 and 3, and 4, and between 2 and 3 and 4 stood out, so that cluster 1 contains the tallest children, followed in order by clusters 2, 3, and 4.

4 Discussion

The study aimed to identify physical-motor profiles throughout preschool age (3–6 years) and explore associations between profiles and selected biosocial factors (age, sex, prematurity, weight, height, BMI, and participation in extracurricular physical activities).

Cluster analysis has made it possible to define four different groups of physical-motor development in childhood. These profiles have differed fundamentally in the level of development of laterality skills (right and left), coordination, balance, motor execution, and tonic.

The type of groups that were formed showed that, although during physical-motor development, the child progressively acquires a set of differentiated skills, these must all be understood as part of a set, with development being compromised in some areas when the others are low. Proof of this is the fact that four profiles were obtained that differentiate between high overall skills (cluster 1) and high skills, except left laterality (cluster 2), medium-low skills (cluster 3), and low skills (cluster 4).

Regarding the difference between clusters 1 and 2, another of the findings of the present study is that, compared to other studies in which laterality is evaluated together, by separating the skills related to right laterality and the left, it has been found that the mastery of left laterality occurs in only one of the groups; specifically, in which it presents the greatest global development at a physical-motor level. The determination of the dominance of one or another member occurs gradually and depends on brain maturation, which can be influenced by aspects related to the environment and stimulation (Hill and Khanem, 2009; Betancourt et al., 2020). Around 90% of the world population are right-handers (Peters et al., 2006), and they tend to be more consistent in using their preferred side; that is, they tend to be more lateralized than a left-hander, who in turn is more

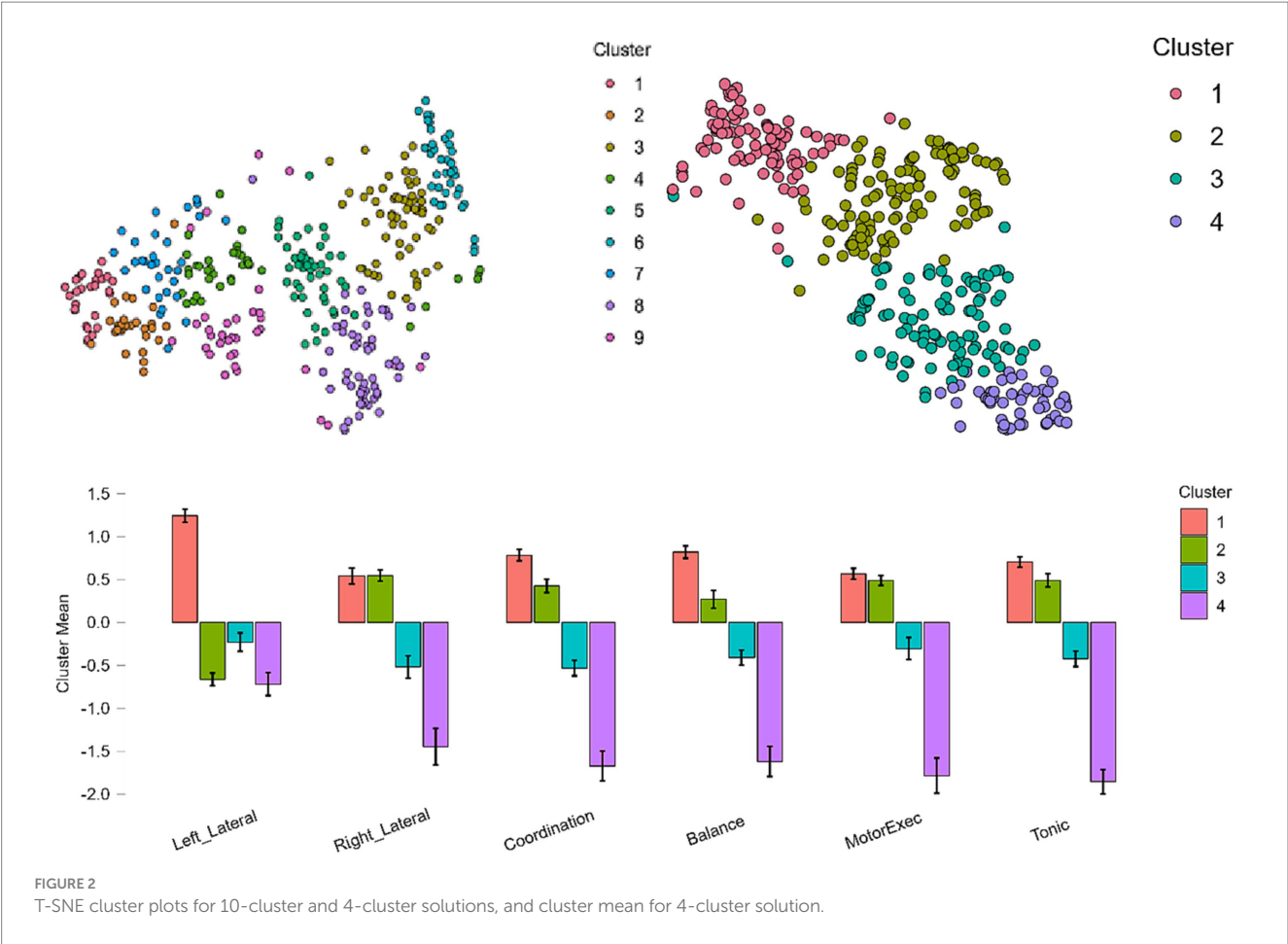


TABLE 4 Between-group differences in the dimensions of physical-motor development for each profile.

	Cluster 1 (n = 121)		Cluster 2 (n = 129)		Cluster 3 (n = 107)		Cluster 4 (n = 55)		F(3, 408)	η^2
Measure	M	SD	M	SD	M	SD	M	SD		
Right laterality	4.75	0.57	4.76	0.41	3.79	0.75	2.95	0.87	152.48**	0.53
Left laterality	4.41	0.58	2.29	0.57	2.77	0.76	2.22	0.67	282.78**	0.68
Coordination	4.55	0.41	4.22	0.51	3.33	0.53	2.27	0.73	295.13**	0.69
Balance	4.66	0.10	4.12	0.70	2.46	0.53	2.29	0.77	217.27**	0.62
Motor execution	4.81	0.32	4.75	0.30	4.15	0.62	3.02	0.70	221.41**	0.62
Tonicity	4.52	0.38	4.31	0.50	3.45	0.54	2.09	0.61	362.60**	0.73

* $p < 0.001$. High PMD (Cluster 1); High PMD, except left laterality (Cluster 2); medium-low PMD (Cluster 3); and low PMD (Cluster 4).

symmetrical since he uses the non-preferred limb more frequently than right-handers (Gurd et al., 2006; Scharoun and Bryden, 2014; Bondi et al., 2020). Thus, left-handers seem to have more advantage in combined coordination tasks using both limbs (Freitas et al., 2014). Some studies have shown that ambidexterity is more common in athletes than in the general population (Stöckel and Weigelt, 2012; Stöckel and Vater, 2014) and that, in some cases, strong lateralization can be disadvantageous in the practice of certain physical activities those who require more open skills and whose use of the preferred and non-preferred hand is essential for their success (Nuri et al., 2013; Stöckel and Vater, 2014). For this reason, although we do not have data on the dominance of laterality in the sample, we can speculate that laterality is defined by age, brain maturation, the environmental

context, and stimulus, so although the high levels of PMD the children in profile 2 are slightly younger than those in profile 1, these children also participate less in extracurricular physical activities and, as such, their left laterality may not be as improved as those in profile 1.

This study suggests that children with lower weights, especially those falling below the WHO weight standards (cluster 4), tend to have lower scores in physical-motor development when compared to children with higher development (cluster 1 and 2). An equivalent situation occurs, where boys and girls who are shorter than the WHO standards (clusters 3 and 4) also present less physical-motor development. These findings reinforce the idea that both weight and height can be important biological factors to consider when assessing and promoting the child's physical motor development. Furthermore,

TABLE 5 *Post-hoc* tests for pairwise differences between profiles according to physical-motor development variables.

	Cluster 1– Cluster 2		Cluster 1– Cluster 3		Cluster 1– Cluster 4		Cluster 2– Cluster 3		Cluster 2– Cluster 4		Cluster 3– Cluster 4	
	M diff (SE)	t	M diff (SE)	T	M diff (SE)	t	M diff (SE)	t	M diff (SE)	t	M diff (SE)	t
Right laterality	−0.01 (0.08)	−0.1	0.96 (0.09)	11.6**	1.81 (0.10)	17.7**	0.97 (0.08)	11.8**	1.81 (0.10)	17.9**	0.84 (0.11)	8.1**
Left laterality	2.13 (0.08)	26.3**	1.64 (0.08)	19.4**	2.19 (0.10)	21.1**	−0.48 (0.08)	−5.8**	0.06 (0.10)	0.6	0.55 (0.10)	5.2**
Coordination	0.33 (0.07)	4.9**	1.22 (0.07)	17.6**	2.28 (0.09)	26.8**	0.89 (0.07)	12.9**	1.95 (0.08)	23.1**	1.06 (0.09)	12.2**
Balance	0.53 (0.08)	6.9**	1.2 (0.08)	14.9**	2.37 (0.10)	24.1**	0.66 (0.08)	8.3**	1.84 (1)	18.8**	1.18 (0.10)	11.7**
Motor execution	0.06 (0.06)	1.0**	0.66 (0.06)	10.6**	1.78 (0.08)	23.4**	0.60 (0.06)	9.8**	1.72 (0.08)	22.8**	1.12 (0.08)	14.4**
Tonicity	0.20 (0.06)	3.2*	1.07 (0.07)	16.2**	2.42 (0.08)	30.0**	0.86 (0.07)	13.3**	2.22 (0.08)	27.8**	1.36 (0.08)	16.5**

* $p < 0.05$, ** $p < 0.001$. High PMD (Cluster 1); High PMD, except left laterality (Cluster 2); medium-low PMD (Cluster 3); and low PMD (Cluster 4).

TABLE 6 Description of the four clusters (profiles) regarding biosocial variables.

		Frequency (%)	M	SD	Skewness	Kurtosis	95% CI Mean	
							Upper	Lower
Cluster 1: High PMD ($n = 121$)	Sex: Male	63 (52.07)						
	Age	121 (100)	57.68	7.50	−0.66	0.34	59.01	56.34
	Weight	117 (96.69)	17.44	2.53	1.36	3.09	17.89	16.98
	Height		105.78	5.46	0.53	<0.01	102.65	97.07
	BMI		15.55	1.49	0.34	1.58	15.82	15.27
	Prematurity: Yes	4 (3.39)						
	Extracurricular physical activities: Yes	41 (33.88)						
Cluster 2 High PMD, except left laterality ($n = 129$)	Sex: Male	60 (46.51)						
	Age	129 (100)	50.60	10.16	−0.11	−1.10	46.99	41.66
	Weight	105 (81.40)	17.44	2.37	−0.57	1.02	17.89	16.98
	Height		104.86	6.80	−1.20	2.33	106.16	103.56
	BMI		15.82	1.30	0.71	1.90	16.07	15.57
	Prematurity: Yes	5 (5.26)						
	Extracurricular physical activities: Yes	26 (20.16)						
Cluster 3 Medium-Low PMD ($n = 107$)	Sex: Male	63 (58.88)						
	Age	107 (100)	48.18	10.15	0.19	−1.17	50.10	52.35
	Weight	75 (70.09)	16.63	2.82	0.79	2.03	17.27	15.99
	Height		100.81	8.83	−0.36	−0.75	102.81	98.82
	BMI		16.33	1.52	0.99	1.15	16.67	15.99
	Prematurity: Yes	7 (7.61)						
	Extracurricular physical activities: Yes	19 (17.76)						
Cluster 4 Low PMD ($n = 55$)	Sex: Male	34 (61.82)						
	Age	55 (100)	44.33	10.07	0.87	−0.46	46.99	41.67
	Weight	37 (67.27)	16.03	2.89	1.82	7.31	16.96	15.10
	Height		99.86	8.67	0.53	0.04	102.65	97.07
	BMI		16.01	1.21	−0.46	0.82	16.40	15.62
	Prematurity: Yes	6 (15.00)						
	Extracurricular physical activities: Yes	4 (7.27)						

TABLE 7 Comparison between real scores for weight and height in the study and those expected based on WHO Standard scores (2006).

		Real scores			Expected weight		Expected height		Expected BMI	
		Weight	Height	BMI	Girls	Boys	Girls	Boys	Girls	Boys
44 months	Cluster 4	16.03	99.86	16.01	15.34	15.68	100.31	101.04	15.29	15.40
48 months	Cluster 3	16.63	100.81	16.33	16.07	16.35	102.73	103.33	15.26	15.33
50 months	Cluster 2	17.69	105.26	15.82	16.43	16.68	103.90	104.45	15.25	15.30
58 months	Cluster 1	17.44	105.78	15.55	17.87	18.01	108.36	108.32	15.27	15.21

TABLE 8 Chi-square test statistics regarding sex, prematurity, and extracurricular physical activities^a.

		Cluster 1: high PMD (n = 121)	Cluster 2 high PMD except for left laterality (n = 129)	Cluster 3: medium-low PMD (n = 107)	Cluster 4 low PMD (n = 55)	Chi-square	V Cramer
Sex	Female	58 (47.93)	69 (53.49)	44 (41.12)	21 (38.18)	5.40	-
	Male	63 (52.07)	60 (46.51)	63 (58.88)	34 (61.82)		
Prematurity	No	114 (96.61)	90 (94.74)	85 (92.39)	34 (85.00)	7.18	-
	Yes	4 (3.39)	5 (5.26)	7 (7.61)	6 (15.00)		
Extracurricular physical activities	No	80 (66.12)	103 (79.85)	88 (82.24)	51 (92.73)	18.38**	0.21
	Yes	41 (33.88)	26 (20.15)	19 (17.76)	4 (7.27)		

^a*p* < 0.05, ***p* < 0.001. ^aThe valid percentage is used, excluding missing values.

TABLE 9 Chi-square test statistics for pairwise differences between profiles regarding extracurricular physical activities.

	Cluster	Chi-square test statistic	V Cramer
Extracurricular physical activities	Medium-low PMD (cluster 3) vs. High PMD (cluster 1).	7.62*	0.18
	High PMD, except left laterality (cluster 2) vs. Low PMD (cluster 4)	4.69*	0.16
	High PMD, except left laterality (cluster 2) vs. High PMD (cluster 1).	5.99*	0.16
	Low PMD (cluster 4) vs. High PMD (cluster 1).	14.07**	0.28

^a*p* < 0.05, ***p* < 0.001.

TABLE 10 Analysis of variance for the variables age, weight, height, and BMI in the different profiles: ANOVA and *post-hoc* tests.

		Age			Weight			Height			BMI		
		Sum squares	F(3,408)	η ²	Sum squares	F(3,408)	η ²	Sum squares	F(3,408)	η ²	Sum squares	F(3,408)	η ²
		8699.4	32.54*	0.19	84.09	4.19*	0.04	1816.05	11.94**	0.10	29.23	4.91*	0.04
<i>Post-hoc</i>		M diff	SE	t	M diff	SE	t	M diff	SE	t	M diff	SE	t
C1	C2	7.08	1.20	5.93*	−3.66×10 ^{−4}	0.35	−0.00	0.92	0.96	0.96	−0.28	0.19	−1.46
	C3	9.50	1.25	7.58**	0.81	0.38	2.11	4.96	1.05	4.71**	−0.79	0.21	−3.77**
	C4	13.35	1.54	8.70**	1.41	0.49	2.88*	5.92	1.34	4.41**	−0.47	0.27	−1.75
C2	C3	2.42	1.23	1.96	0.81	0.39	2.06	4.04	1.08	3.76**	−0.51	0.21	−2.39
	C4	6.27	1.52	4.12**	1.41	0.50	2.84*	5.00	1.36	3.67*	−0.19	0.27	−0.70
C3	C4	3.85	1.57	2.46	0.60	0.52	1.15	0.96	1.43	0.67	0.32	0.28	1.13

^a*p* < 0.05, ***p* < 0.001. C1: High PMD, C2: High PMD (except left laterality), C3: Medium-low PMD, and C4: Low PMD.

it reinforces the assumption that motor development in children at very young ages is more influenced by biological maturation, and after this period, it becomes more influenced by environmental factors (stimuli/practice; Barnett et al., 2012).

Regarding extracurricular physical activities practice, the results are in line with many studies (Riethmuller et al., 2009; Holfelder and Schott, 2014; Valentini et al., 2016; Wick et al., 2017) since benefits could be deduced from this development by finding

that those who practice extracurricular physical activities the most are in the high PMD group (cluster 1) and high PMD, except laterality (cluster 2).

Another element that we must consider is premature birth, although its relevance was only observed in the contrast between the high PMD group (cluster 1) and the low PMD group (cluster 4). Despite the lower number of premature children in this sample, these results are in line with several studies that indicate that premature children tend to have less motor competence than children who are born at full term (FitzGerald et al., 2021; Karageorgi et al., 2021).

While all these values may be relevant, there is no doubt that it is necessary to consider one of the critical variables in development: age. The data showed that children who have higher values of physical-motor development are also the oldest, being variable with significance in contrast between practically all groups [except the differences between high PMD (cluster 2) and medium-low PMD (cluster 3), as well as between this and low PMD (cluster 4)]. Motor skill development is a progressive process in which children typically acquire new skills and refine existing ones as they get older (Barnett et al., 2012; Valentini et al., 2016; Bolger et al., 2018; Li et al., 2023). It is also important to highlight that motor development depends on several biosocial factors, and, as such, not all children reach developmental milestones at the same time.

5 Conclusion, limitations, and future studies

These results lead us to question whether it is possible to establish profiles at such early ages. The answer we can approximate, based on the data, is considering the differences that exist in early development and the possible influence of multiple variables—not only biological but also environmental. This study allows us to identify specific trends that may be conducive to child development, such as participation in extracurricular physical activities. Other variables, such as sex, BMI, or prematurity, do not seem decisive when determining a higher or lower physical-motor development profile.

Given this data, analyzing these profiles in greater depth within broader age ranges and observing how they evolve could be interesting. This could provide a clearer idea of child development's potential facilitators and risks. In this study, children's postural status was not controlled. In future studies, it is recommended that this variable be incorporated, as well as other environmental variables that may be involved in these development profiles.

Data availability statement

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

References

- Allotey, J., Zamora, J., Cheong-See, F., Kalidindi, M., Arroyo-Manzano, D., Asztalos, E., et al. (2018). Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64061 children. *BJOG: Int J Obstet Gyn.* 125, 16–25. doi: 10.1111/1471-0528.14832
- Baidari, I., and Patil, C. (2020). A criterion for deciding the number of clusters in a dataset based on data depth. *Vietnam J Comput Sci.* 7, 417–431. doi: 10.1142/S2196888820500232
- Barnett, L., Hinkley, T., Okely, A. D., and Salmon, J. (2012). Child, family and environmental correlates of children's motor skill proficiency. *J. Sci. Med. Sport* 16, 332–336. doi: 10.1016/j.jsams.2012.08.011
- Betancourt, G., Sánchez, C., Beatriz, A. M., and Barroso, E. (2020). Patrón de lateralidad en jugadores masculinos de baloncesto, reservas escolares y juveniles de La Habana. *Podium. Rev Cienc Tecnol Cultura Fis.* 15, 449–459. doi: 10.47230/unescumciencias.v4.n1.2021.308

Ethics statement

The studies involving humans were approved by Junta de Comunidades de Castilla-La Mancha. Ethical Committee of the CRA “Calar del Mundo.” The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. 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

PG-M: Data curation, Investigation, Writing – original draft. LL-P: Formal analysis, Methodology, Writing – original draft. PM: Formal analysis, Methodology, Writing – original draft. CS: Conceptualization, Investigation, Supervision, Writing – review & editing. IS: Conceptualization, Supervision, Validation, Writing – review & editing. LS: Conceptualization, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Part of this research was funded by the Castilla-La Mancha University. Investigation Project: “Physical Activity and Health, aid for carrying out applied research projects, within the framework of the UCLM research plan, 85% co-financed by the European Regional Development Fund.” Organic number (2022-GRIN-34290).

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.

- Bolger, L. E., Bolger, L. A., O'Neill, C., Coughlan, E., O'Brien, W., Lacey, S., et al. (2018). Age and sex differences in fundamental movement skills among a cohort of Irish school children. *J Motor Learn Dev.* 6, 81–100. doi: 10.1123/jmld.2017-0003
- Bondi, D., Prete, G., Malatesta, G., and Robazza, C. (2020). Laterality in children: evidence for task-dependent lateralization of motor functions. *Int. J. Environ. Res. Public Health* 17, 1–13. doi: 10.3390/ijerph17186705
- Boonzaaijer, M., Oudgenoeg-Paz, O., Suij, I., Westers, P., Nuysink, J., Volman, M., et al. (2021). Modeling a gross motor curve of typically developing Dutch infants from 3.5 to 15.5 months based on the Alberta infant motor scale. *Early Hum. Dev.* 157:105366. doi: 10.1016/j.earlhumdev.2021.105366
- Bronfenbrenner, U. (2005). *Making human beings human: Bioecological perspectives on human development*. California, USA: Sage Publications Ltd.
- Calinski, T., and Harabasz, J. (1974) A Dendrite Method for Cluster Analysis Communications in Statistics. *Theory and Methods*, 3, 1–27.
- Cao, F., Liang, J., and Juang, G. (2009). An initialization method for the K-means algorithm using neighborhood model. *Comput Math Appl.* 58, 474–483. doi: 10.1016/j.camwa.2009.04.017
- Clarworthy, J., Buick, D., Hankins, M., Weinman, J., and Horne, R. (2005). The use and reporting of cluster analysis in health psychology: a review. *British J Health Psycho.* 10, 329–358. doi: 10.1348/135910705X25697
- Cliff, D. P., Okely, A. D., Morgan, P. J., Jones, R. A., Steele, J. R., and Baur, L. A. (2012). Proficiency deficiency: mastery of fundamental movement skills and skill components in overweight and obese children. *Obesity* 20, 1024–1033. doi: 10.1038/oby.2011.241
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Hillsdale, NJ: LEA
- Coppens, E., Rommers, N., Bardid, E., Deconinck, F. J. A., De Martelaer, K., D'Hondt, E., et al. (2021). Long-term effectiveness of a fundamental motor skill intervention in Belgian children: a 6-year follow-up. *Scand. J. Med. Sci. Sports* 31, 23–34. doi: 10.1111/sms.13898
- David, M., and Sutton, C.D. (2004). *Social research. The basics*. California, USA: SAGE Publications.
- Drenowatz, C., and Greier, K. (2019). Association of biological maturation with the development of motor competence in Austrian middle school students—a 3-year observational study. *Translational pediatrics* 8, 402–411. doi: 10.21037/tp.2019.09.03
- Escolano-Pérez, E., Sánchez-López, C. R., and Herrero-Nivela, M. L. (2022). Teacher-rated executive functions, gender and relative age: independent and interactive effects on observed fundamental motor skills in kindergarteners. *Front. Psychol.* 13:848525. doi: 10.3389/fpsyg.2022.848525
- FitzGerald, T. L., Cameron, K. L., Albeshier, R. A., Mentiplay, B. F., Lee, K. J., Clark, R. A., et al. (2021). Strength, motor skills, and physical activity in preschool-aged children born either at less than 30 weeks of gestation or at term. *Phys. Ther.* 101:pzab037. doi: 10.1093/ptj/pzab037
- Freitas, C., Vasconcelos, O., and Botelho, M. (2014). Handedness and developmental coordination disorder in Portuguese children: study with the M-ABC test. *Laterality* 19, 655–676. doi: 10.1080/1357650X.2014.897349
- Gabbard, C. (2011). *Lifelong motor development (6th ed)*. San Francisco, CA: Benjamin Cummings.
- Gabbard, G. (2014). *Psychodynamic psychiatry in clinical practice (5th ed.)*. USA: American Psychiatric Publishing, Inc.
- Gurd, J. M., Schulz, J., Cherkas, L., and Ebers, G. C. (2006). Hand preference and performance in 20 pairs of monozygotic twins with discordant handedness. *Cortex* 42, 934–945. doi: 10.1016/S0010-9452(08)70438-6
- Hadders-Algra, M. (2018). Early human motor development: from variation to the ability to vary and adapt. *Neurosci. Biobehav. Rev.* 90, 411–427. doi: 10.1016/j.neubiorev.2018.05.009
- Hardy, L. L., O'Hara, B. J., Rogers, K., St George, A., and Bauman, A. (2014). Contribution of organized and nonorganized activity to children's motor skills and fitness. *J. Sch. Health* 84, 690–696. doi: 10.1111/josh.12202
- Hill, E. L., and Khanem, F. (2009). The development of hand preference in children: the effect of task demands and links with manual dexterity. *Brain Cogn.* 71, 99–107. doi: 10.1016/j.bandc.2009.04.006
- Holfelder, B., and Schott, N. (2014). Relationship of fundamental movement skills and physical activity in children and adolescents: a systematic review. *Psycho Sport Exercise.* 15, 382–391. doi: 10.1016/j.psychsport.2014.03.005
- Ikedo, T., and Aoyagi, O. (2008). Meta-analytic study of gender differences in motor performance and their annual changes among Japanese preschool-aged children. *Japanese J Sch Health.* 4, 24–39. doi: 10.1016/j.biortech.2007.01.046
- Jaakkola, T., Yli-Piipari, S., Huhtiniemi, M., Salin, K., Hakonen, H., and Gråstén, A. (2021). Motor competence and health-related fitness of school-age children: a two-year latent transition analysis. *Med. Sci. Sports Exerc.* 53, 2645–2652. doi: 10.1249/MSS.0000000000002746
- Jones, D., Innerd, A., Giles, E. L., and Azevedo, L. B. (2020). Association between fundamental motor skills and physical activity in the early years: a systematic review and meta-analysis. *J. Sport Health Sci.* 9, 542–552. doi: 10.1016/j.jshs.2020.03.001
- Karageorgi, N., Charitou, S., Asonitou, K., and Koutsouki, D. (2021). Motor profile of late preterm infants: a systematic review of the last decade. *J Biosci Meds.* 9, 195–206. doi: 10.4236/jbm.2021.97018
- Kaufman, L., and Rousseeuw, P. (1990). *Finding groups in data: An introduction to cluster analysis*. USA: John Wiley & Sons, Inc.
- Krebs, C. (2009). *Ecology: The experimental analysis of distribution and abundance*. San Francisco, USA: Pearson Benjamin Cummings.
- Lai, J. C., and Huang, T.-J. (2011). An agglomerative clustering algorithm using a dynamic k-nearest-neighbor list. *Inf. Sci.* 181, 1722–1734. doi: 10.1016/j.ins.2011.01.011
- Li, K., Chen, S., Ma, J., Martins, C., Duncan, M., Sheng, X., et al. (2023). The relative age effect on fundamental movement skills in Chinese children aged 3–5 years. *BMC Pediatr.* 23:150. doi: 10.1186/s12887-023-03967-6
- Livonen, S., and Sääkslahti, A. (2014). Preschool children's fundamental motor skills: a review of significant determinants. *Early Child Dev. Care* 184, 1107–1126. doi: 10.1080/03004430.2013.837897
- Lohman, T.G., Roche, A.F., and Martorell, R. (1988). *Anthropometric standardization reference manual*. Chicago, USA: Human Kinetics Books.
- Lopes, V. P., Malina, R. M., Maia, R., and Rodrigues, L. P. (2018). Body mass index and motor coordination: non-linear relationships in children 6–10 years. *Childright* 44, 443–451. doi: 10.1111/cch.12557
- Lubans, D., Morgan, P., Cliff, D., Barnett, L., and Okley, A. (2010). Fundamental movement skills in children and adolescents. *Sports Med.* 40, 1019–1035. doi: 10.2165/11536850-000000000-00000
- Mojena, R. (1977). Hierarchical grouping methods and stopping rules: an evaluation. *Aust. Comput. J.* 20, 359–363. doi: 10.1093/comjnl/20.4.359
- Moreira, R. S., Magalhães, L. C., and Alves, C. R. (2014). Effect of preterm birth on motor development, behavior, and school performance of school-age children: a systematic review. *J. Pediatr.* 90, 119–134. doi: 10.1016/j.jpeds.2013.05.010
- Navarro-Patón, R., Lago-Ballesteros, J., Arufe-Giraldez, V., Sanmiguel-Rodríguez, A., Lago-Fuentes, C., and Meciás-Calvo, M. (2021). Gender differences on motor competence in 5-year-old preschool children regarding relative age. *Int. J. Environ. Res. Public Health* 18:3143. doi: 10.3390/ijerph18063143
- Nuri, L., Shadmehr, A., Ghotbi, N., and Attarbashi Moghadam, B. (2013). Reaction time and anticipatory skill of athletes in open and closed skill-dominated sport. *Eur. J. Sport Sci.* 13, 431–436. doi: 10.1080/17461391.2012.738712
- Okely, A. D., Booth, M. L., and Chey, T. (2004). Relationships between body composition and fundamental movement skills among children and adolescents. *Res. Q. Exerc. Sport* 75, 238–247. doi: 10.1080/02701367.2004.10609157
- Oudgenoeg-Paz, O., Mulder, H., Jongmans, M. J., van der Ham, I. J. M., and Van der Stigchel, S. (2017). The link between motor and cognitive development in children born preterm and/or with low birth weight: a review of current evidence. *Neuro Sci Biobehav Rev.* 80, 382–393. doi: 10.1016/j.neubiorev.2017.06.009
- Pereira, S., Bustamante, A., Santos, C., Hedeker, D., Tani, G., Garganta, R., et al. (2021). Biological and environmental influences on motor coordination in Peruvian children and adolescents. *Sci. Rep.* 11:15444. doi: 10.1038/s41598-021-95075-7
- Peters, M., Reimers, S., and Manning, J. T. (2006). Hand preference for writing and associations with selected demographic and behavioral variables in 255, 100 subjects: the BBC internet study. *Brain Cogn.* 62, 177–189. doi: 10.1016/j.bandc.2006.04.005
- Riethmuller, A. M., Jones, R., and Okely, A. D. (2009). Efficacy of interventions to improve motor development in young children: a systematic review. *Pediatrics* 124, e782–e792. doi: 10.1542/peds.2009-0333
- Romero Martínez, S. J., Ordóñez Camacho, X. G., and Gil Madrone, P. (2018). Development of the checklist of psychomotor activities for 5- to 6-year-old children. *Percept. Mot. Skills* 125, 1070–1092. doi: 10.1177/0031512518804359
- Saraiva, L., Rodrigues, L. P., Cordovil, R., and Barreiros, J. (2013). Influence of age, sex and somatic variables on the motor performance of pre-school children. *Ann. Hum. Biol.* 40, 444–450. doi: 10.3109/03014460.2013.802012
- Scharoun, S. M., and Bryden, P. J. (2014). Hand preference, performance abilities, and hand selection in children. *Front. Psychol.* 5, 5–82. doi: 10.3389/fpsyg.2014.00082
- Schubert, E. (2022). Stop using the elbow criterion for k-means and how to choose the number of clusters instead. *Statistics. Machine Learn.* 1, 1–7. doi: 10.48550/arXiv.2212.12189
- Stöckel, T., and Vater, C. (2014). Hand preference patterns in expert basketball players: interrelations between basketball-specific and everyday life behavior. *Hum. Mov. Sci.* 38, 143–151. doi: 10.1016/j.humov.2014.09.002
- Stöckel, T., and Weigelt, M. (2012). Plasticity of human handedness: decreased one-hand bias and inter-manual performance asymmetry in expert basketball players. *J. Sports Sci.* 30, 1037–1045. doi: 10.1080/02640414.2012.685087
- Valentini, N., Logan, S., De Spessato, B., Souza, M., Pereira, K., and Rudisill, M. (2016). Fundamental motor skills across childhood: age, sex, and competence outcomes of Brazilian children. *J Motor Learn Dev.* 4, 16–36. doi: 10.1123/jmld.2015-0021
- Verbecque, E., Coetsee, D., and Smits-Engelsman, B. (2022). Underweight children are agile but lack power. *BMC Pediatr.* 22:490. doi: 10.1186/s12887-022-03544-3

- Webb, A., O'Reilly, A., Clayden, D., Seunarine, K., Chong, K., Dale, N., et al. (2012). Effect of growth hormone deficiency on brain structure, motor function and cognition. *Brain J. Neurol.* 135, 216–227. doi: 10.1093/brain/awr305
- Werneck, O., Fernandes, A., and Silva, R. (2020). Understanding biological maturation and motor competence for physical activity promotion during the first years of life. *Translational pediatrics* 9, 1–3. doi: 10.21037/tp.2020.01.02
- Wick, K., Leeger-Aschmann, C. S., Monn, N. D., Radtke, T., Ott, L. V., Rebholz, C. E., et al. (2017). Interventions to promote fundamental movement skills in childcare and kindergarten: a systematic review and Meta-analysis. *Sports Med.* 47, 2045–2068. doi: 10.1007/s40279-017-0723-1
- World Health Organization. (2006). *WHO child growth standards length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. Methods and development.* WHO Library Cataloguing-in-Publication Data.
- Xin, F., Chen, S. T., Clark, C., Hong, J. T., Liu, Y., and Cai, Y. J. (2020). Relationship between fundamental movement skills and physical activity in preschool-aged children: a systematic review. *Int. J. Environ. Res. Public Health* 17:3566. doi: 10.3390/ijerph17103566
- Xu, J., and Núñez, J. C. (2023). Homework purposes in eighth grade students: identifying student profiles and their relationship with homework effort, completion, and achievement. *Psicothema* 35, 111–118. doi: 10.7334/psicothema2022.334



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do
Castelo, Portugal

REVIEWED BY

Zhuoying Wang,
The University of Texas at Austin, United States
Elizabeth Jane Newton,
London South Bank University,
United Kingdom

*CORRESPONDENCE

David Tzuriel
✉ david.tzuriel@biu.ac.il

RECEIVED 16 November 2023

ACCEPTED 31 January 2024

PUBLISHED 28 February 2024

CITATION

Tzuriel D (2024) Analogical thinking
modifiability and math processing strategy.
Front. Psychol. 15:1339591.
doi: 10.3389/fpsyg.2024.1339591

COPYRIGHT

© 2024 Tzuriel. 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.

Analogical thinking modifiability and math processing strategy

David Tzuriel^{1,2*}

¹Talpiot Academic College, Graduate Program of Learning Disabilities, Holon, Israel, ²Faculty of Education, Bar-Ilan University, Ramat Gan, Israel

A sample of 48 children in Grade 2 was randomly assigned to an experimental ($n = 24$) and a control group ($n = 24$). Both groups were administered the Analogical Modifiability Puzzle Test (AMPT) and Math Accuracy and Processing Strategy (MAPS) test before and after a teaching phase of the AMPT. The MAPS test includes scores for Accuracy, Processing Strategy, and a Math-Total. The findings reveal significant treatment $\times$ time interactions for AMPT (near-transfer) and MAPS (far-transfer) scores. Hierarchical regression analysis showed that AMPT post-teaching score added significantly to Math Total. The findings indicate that Math Accuracy and Math Processing Strategies are affected by mediation for analogical thinking and that modifiability of analogical thinking significantly predicts Math-Total.

KEYWORDS

analogical thinking, dynamic assessment, mathematical performance, mediated learning, executive functions, mathematical processing strategy

Introduction

Analogical thinking is a fundamental cognitive operation that plays a role in executing many academic and daily life skills (Gentner et al., 2003), especially skills related to STEM domains (Alexander, 2017). Analogical thinking is a cognitive operation that helps individuals to represent information and objects as the structures of relationships. These structures can be compared and united in innovative ways depending on the contextual objectives (e.g., Gentner et al., 2003; Holyoak, 2004). Learning by analogy involves flexible conceptual learning and problem-solving (Goswami, 1992), especially in mathematics (Gelman, 2000). It was reported, for example, that, when US high-school students engage in solving math problems, they are found to be retrieving procedures rather than perceiving the relationship and similarities between concepts and strategies. In other words, their processing strategies for solving math problems lack analogical thinking (Stigler et al., 2010; Givvin et al., 2011).

Analogical thinking was reported to be central to mathematical learning and thinking (e.g., Richland and McDonough, 2010; Vamvakoussi and Vosniadou, 2012; Vamvakoussi, 2017, 2019; Zhao et al., 2021). Most studies on the use of analogy in math education refer to within-domain (i.e., mathematics) transfer of concepts such as the complex relation between geometry and arithmetic (Dantzig, 2005). The objectives of this study were (a) to examine the transfer of analogical thinking in a figural modality of a cognitive non-mathematical domain to math performance and math processing strategies and (b) to find the prediction of improvement in math performance by analogical thinking modifiability. The transfer is examined after a short-term teaching of figural analogies within a dynamic assessment (DA) procedure (Tzuriel and George, 2009; Tzuriel and Shamir, 2010; Tzuriel, 2021a). It should be noted that the teaching of figural analogies involves, in addition to transformative rules of analogical thinking, mediation of executive functions such as control of impulsivity, systematic exploratory behavior, and working memory (Feuerstein et al., 2002; Baddeley, 2012).

Analogical thinking and mathematical skills

Analogical thinking and math achievements have been found to be intimately related both at theoretical and practical levels (e.g., Bassok and Holyoak, 1989; Goswami, 1992; Gelman, 2000; Resing, 2000; English, 2004; Richland and Simms, 2015). Learning by analogy enables the use of common concepts and strategies between mathematical representations and enhances the understanding of new mathematical constructs and synchronization between mathematical components (Gelman, 2000; Gonzales et al., 2001; National Mathematics Advisory Panel, 2008). Using analogies in a learning process involves finding a systematic correspondence or “mapping” two domains: an acquainted *source* analog and a novel *target*, which is concrete or abstract (Sternberg and Nigro, 1980). For example, in mathematics, the acquainted source could refer to solving inequalities, and the projected novel target is solving equations (Gick and Holyoak, 1987). Many researchers agree that mathematical reasoning involves projecting (transfer) the abstract mathematical relations (e.g., proportion, equality, integral) into different contexts (e.g., Gelman, 2000; Gentner et al., 2003). Several researchers suggest the concepts of *conceptual metaphors* (e.g., Núñez and Lakoff, 2005), *schema construction* (Gick and Holyoak, 1983), and *bridging analogies* (e.g., Clement, 1993, 2013) to refer to cross-domain mappings (i.e., transfer concept from one domain to another).

In one of the intriguing cross-cultural studies, Richland et al. (2007) compared the effects of the use of analogies in math teaching among teachers in the United States, Hong Kong (HK) and Japan. The reasons for the comparison were based on two main earlier findings: (a) students in HK and Japan consistently outperformed students from the United States on the Trends in International Mathematics and Science Study (TIMSS) international achievement tests (e.g., Mullis et al., 2004) and (b) different classroom instructional practices were observed between the teachers of the USA and two Asian countries. The goal of Richland et al.'s (2007) study was to ascertain what cognitive and memory factors may explain why US students achieve lower scores in international math tests. The findings indicated that the US teachers provide the same number of analogies as the Asian teachers; however, they provide much less *cognitive processing support* that facilitates relational learning. For example, Asian teachers used far more gestures and more than twice as many prompts of mental and visual imagery (in HK) and spatial supports for comparison (in HK and Japan) than US teachers. The authors interpreted these findings by attributing diverse cultural orientations to analogical reasoning. Teachers in Japan and HK are found to be more attentive to processing demands using analogical reasoning than US teachers. The processing demands include memory search to understand the analogy source, accessibility to a visual image that supports working memory (WM) during the construction of the analogy, and reliance on cues that guide the analogy process.

Richland and McDonough (2010) showed that reducing the processing load during instructional analogies facilitated learning and relational reasoning. These authors showed, in a sample of undergraduate students, that supporting comparative reasoning with cues facilitated the students' ability to discriminate

between relevant analogs at a test given later. Similarly, Rittle-Johnson and Star (2007) compared two teaching strategies: simultaneous visual-spatial representations of analogs (i.e., problems solved in two ways were presented simultaneously and an explicit requirement to compare them) and solutions presented sequentially one after another. The authors reported that simultaneous representation enhanced the students' procedural and conceptual understanding and that their retention of the procedures, were better transferred to novel analogy problems.

The dynamic assessment of cognitive modifiability

The term dynamic assessment (DA) is defined as “an assessment of thinking, perception, learning, and problem-solving by an active teaching process aimed at modifying cognitive function” (Tzuriet, 2021a, p. 69). The main goal of DA is to assess the *cognitive modifiability* and learning potential using mediated learning strategies. Cognitive modifiability affords an indication of learning potential. However, the learning potential may be actualized depending on the provision of appropriate intervention aimed at the actualization of the individual's abilities (Tzuriet, 2001, 2002, 2021a,b; Feuerstein et al., 2002). The idea upon which DA is based is that measures of cognitive modifiability are intimately related to mediational processes by which the child learns how to process information far more than they are to the standardized measures of intelligence. In other words, the strategies of mediation used within the DA procedure are similar much more to the learning processes in an academic domain and other life situations compared to standardized testing methods and, therefore, provide better indications about future changes in cognitive structures.

In a previous study with the *Analogical Thinking Modifiability Test* (AMPT, Tzuriet, 2022), the goals were (a) to investigate whether mediation with Construction analogies is more powerful than mediation with Closed analogies to improve analogical thinking. In Construction analogies, the child must construct the answer using tactile colored puzzle-like parts and a plate with four (2×2) “windows” in which the puzzle-like design is created. The solution is done in a sequential process that does not require the activation of WM. In the Closed analogies version, the child is presented with three parts of the analogy (A:B::C:___) and is asked to choose the correct answer from six designs presented at the bottom of the page. (b) The second goal was to assess whether the mediation of solving analogies improves spatial WM, as measured by the *Children's Spatial Working Memory* test (CSWM, Tzuriet, 2021c). More specifically, the question was which type of mediation—mediation of Construction or mediation of Closed analogies—transfers better to mathematics? (c) The last goal was to explore the correlation between analogical thinking and WM. The findings indicated clearly that mediation with Construction analogies was more effective in enhancing analogical thinking than mediation with Closed analogies and that both groups improved their spatial WM. The reason for the superiority of Construction analogies is that the format of Construction analogies contains imminent characteristics of self-regulation and systematic sequential and progressive activity, which reduce cognitive load and

WM requirements. In Closed analogies, on the other hand, the examinee is required to maintain the solution in his/her *memory sketchpad* (Baddeley, 2012) during the search for the correct answer and sorting out distractors. The explanation was further supported by the differential correlation between AMPT performance time and CSWM. The correlations found for Construction analogies were 0.14 and 0.22 ($p > 0.05$), for pre-teaching and post-teaching, respectively. The correlations for Closed analogies were 0.47 and 0.47 ($p < 0.001$), for pre- and post-teaching, respectively. It seems that, in Closed analogies, a longer performance time helps individuals to cope with the task's requirement for WM. In contrast, Construction analogies do not require the activation of WM and hence its low correlation with spatial WM. The findings of the previous study raised the question of whether the improvement in analogical thinking was caused by the mediation strategies *per se* or the mere exposure and practice of the analogy problems. This question is explored further in the current study by contrasting a group that received the mediation of analogical thinking with a group that practiced the same teaching items without mediation. I chose to focus on Closed analogies, which are found to be more related to WM than to Construction analogies. Furthermore, in the current study, the focus was extended to include the far-transfer effects of mediation of analogical thinking on math performance and math processing strategies conceived to be related to the executive functions of WM (e.g., Richland et al., 2007).

In the current study, the measurement/research version of DA was applied (Tzuriel, 1997, 2001, 2020, 2021a). This version includes the preliminary, pre-teaching, teaching, and post-teaching phases; the last three phases consist of parallel items. In the preliminary phase, the child is acquainted with the dimensions of the tasks and cognitive strategies to solve the problems (e.g., referring to task dimensions). The focus of the teaching phase is on processing the analogical rules of transformation, efficient strategies (e.g., use of one dimension at a time), applying verbal anticipation of a solution, and making efficient use of executive functions such as self-regulation and WM. Cognitive modifiability is measured quantitatively and is indicated by the pre- to post-teaching improvement. Previous DA findings revealed that children in kindergarten and first years of primary school could solve complex analogies after a teaching phase than what might have been expected developmentally of children their age (Tzuriel and Klein, 1985; Tzuriel, 2000, 2001, 2020, 2021a). Children in Grade 2 were randomly assigned to the experimental (EX) and control (CN) groups. Both groups received the AMPT and the *Math Accuracy and Processing Strategy (MAPS)* Test (see Measures). Children in the EX group received a teaching phase of analogies using mediation strategies by trained experimenters. Children in the CN group solved the same analogies without teaching. Both groups were evaluated again after the intervention with the AMPT and MAPS test.

Hypotheses

1. Children in the EX group who received the mediation of analogical thinking will demonstrate higher improvement on

AMPT analogies from pre- to post-teaching than children in the CN group.

2. Children in the EX group who received the mediation of analogical thinking will demonstrate higher improvement in the MAPS test than children in the CN group.
3. Cognitive modifiability of analogical thinking will significantly predict the modifiability in MAPS scores.

Method

Participants

The participants were 48 children (25 boys and 23 girls) in Grade 2 who were randomly selected from six classes from six different schools in the central region of Israel. Students with learning problems (e.g., learning disability, attention-deficit hyperactivity disorder, autism spectrum disorder) were excluded from the sample. The children were randomly assigned to the EX ($n=24$) and CN ($n=24$) groups. The mean age of children in the EX group was 87.96 months ($SD = 2.90$) and in the CN group was 89.33 months ($SD = 3.17$). Both groups were similar on variables of age, $t_{(46)} = -1.56$, $p > 0.05$, gender distribution, $\chi^2 = 0.08$, $p > 0.05$, occupation level of fathers, $\chi^2 = 0.22$, $p > 0.05$, and of mothers $\chi^2 = 0.93$, $p > 0.05$ (on a 1–5 scale). The trainers were graduate students in a Learning Disabilities Program of a college for teachers with 4–15 years of teaching experience. All trainers received 8 h of training prior to the administration of tests and mediation strategies. The training included proficiency in test administration, specific mediation strategies of the analogies in the teaching phase, practice of administration with two children prior to the start of the study, and accurate recording and scoring.

Measures

The analogical modifiability puzzle test

The AMPT (Tzuriel, 2021a,b, 2022) is a DA measure of cognitive modifiability in the domain of analogical thinking. The AMPT contains three parallel sets of problems, each containing 12 items. The AMPT can be administered either in the *Construction* analogies or *Closed* analogies version (Tzuriel, 2022). In the current study, the Closed analogies version was applied. Before the start of the test, children are presented with two introductory problems aimed at familiarizing the children with the dimensions of AMPT (i.e., color, shape of puzzle parts) and the rules of problem-solving. The items are arranged progressively from easy to difficult problems. In the Closed analogies version, which is used in the current study, the child is presented at the top of the page with three sections of the analogy (see Figure 1), and the child must choose the fourth section, from six answers, of which five are distractors. To solve the problem and complete the puzzle, the child must consider the changes from the left to the right section and from the top to the bottom section. The correct solution requires considering changes in the color and position of the panels and shapes from left to right and from top to bottom.

In Figure 1, the colors of the background panels change from left to right (white and blue to blue and white). The hearts on the left sections that are empty are filled up with red hearts on

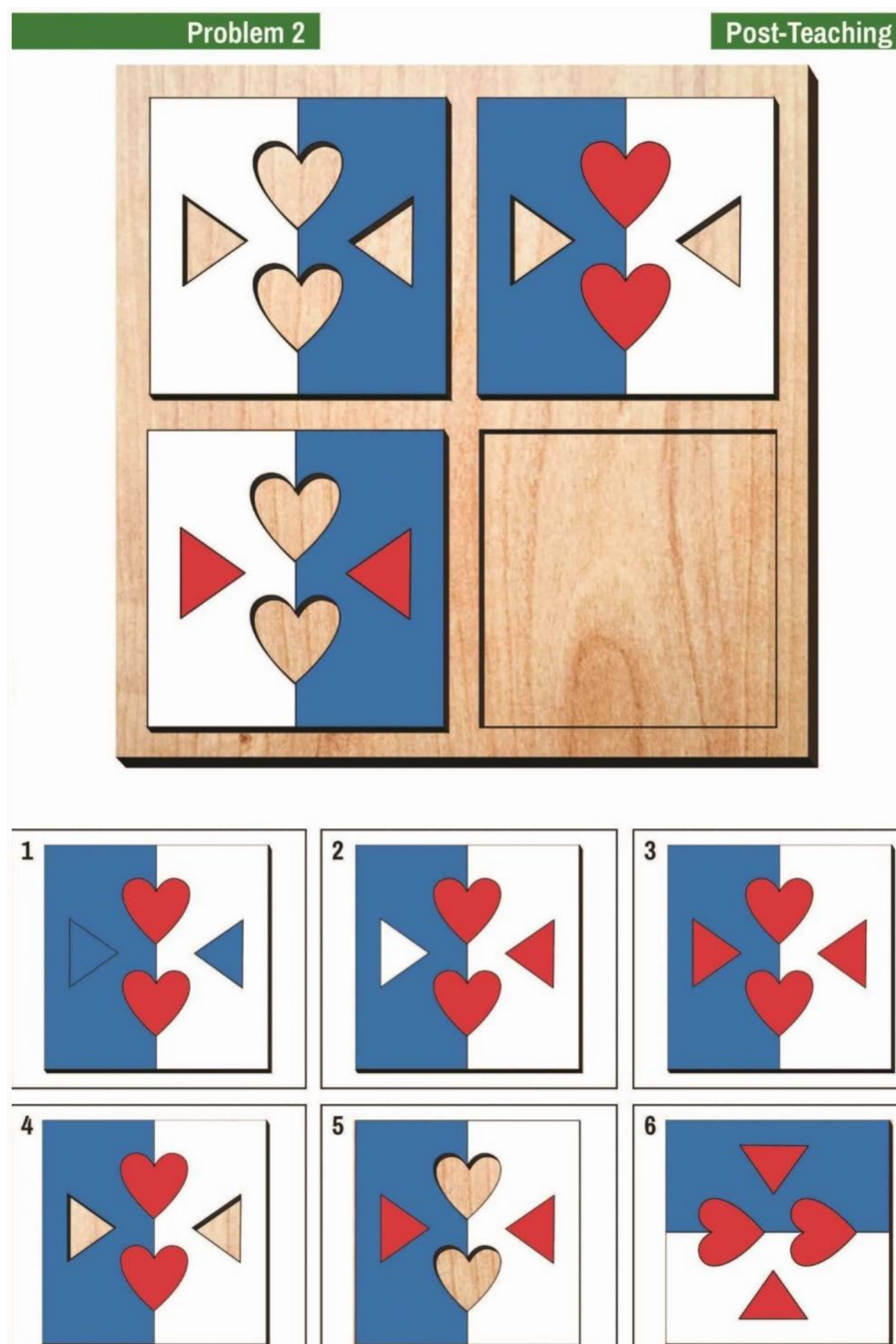


FIGURE 1
Example of an easy item from the Analogical Modifiability Puzzle Test (AMPT) (by permission of the author).

the right, and the empty triangles on the top are filled up with red triangles on the bottom. The correct solution is number 3 (blue and white panels and red hearts and triangles). Cronbach's alpha reliability coefficients for the pre- and post-teaching phases of the AMPT based on a sample of kindergarten children ($n = 96$) were 0.94 and 0.95, respectively (Tzuriel, 2022). Cronbach's alpha coefficients for the Pre- and Post-Teaching phases in the current sample were 0.94 and 0.73, respectively. Previous findings showed that a short-term intervention within a DA procedure using the AMPT improved spatial working memory (SWM) and that the

improvement in SWM was even more powerful ($\eta_p^2 = .61$) than the effect on analogical thinking ($\eta_p^2 = .40$) *per se*, which was the targeted mediation activity (Tzuriel, 2022).

Math accuracy and processing strategy test (MAPS)

The MAPS test (Tzuriel, 2024), designed specifically for the current study, is aimed at the examination of both performance accuracy and processing strategy. In each item, the children are

TABLE 1 Examples of problems from the Math Accuracy and Processing Strategy (MAPS) Test.

Item #	Problem	Solution	Processing strategy: explain how you solved the problem?
3	Here is a math problem with a solution: $19 + 34 = 53$. Please use this solution to solve the following problem without using a computation $29 + 24 =$		
5	You are asked to count from 2 to 32 skipping 2 each time. Will the number 87 appear later in the series?		
7	In how much the sum of $64 + 8$ is smaller than $64 + 28$?		
10	Ron is 35 years old, and Dan is 16 years old. How old is Tammy if we know that the sum of all three ages is 60?		

presented a math problem and asked to solve it and to write how they solved the problem (processing strategy). The test is composed of 10 items and is aimed at children in the second grade. In each item, the child is asked first to solve a mathematical problem and then to explain the strategy used to reach a solution. The explanation was aimed at verifying that the answer is based on a strategy and not on a technical performance. Separate scores are given to accurate answers and to explanations revealing an efficient strategy. A separate score of 1 was given in each item to an accurate answer and to an efficient processing strategy. Examples of problems are shown in Table 1. The Cronbach's alpha coefficients based on the current sample, for pre- and post-intervention, respectively, were Math Accuracy, 0.82 and 0.77; Math Processing Strategy, 0.84 and 0.83; Math Total, 0.77 and 0.77.

Procedure

All children were administered the AMPT individually using the measurement/research format and the MAPS test before and after the intervention phase. Children in the EX group received the mediation with 12 analogy teaching items for 1 h. Mediation was focused on the systematic exploration of task dimensions, the deletion of impossible alternatives, the delay of rapid answer and inhibition of behavior, the comparison of the different options, and the verbal anticipation of the answer. Children in the CN group practiced the teaching items with no mediation for the same amount of time. The test was administered in a quiet room in the school during regular school time. The pre-intervention phase was administered to all participants for 1 h followed by the intervention and post-intervention phases that were given between 1 and 5 days later; both tests were administered on the same day with a small break of 10 min in-between.

Results

The effects of teaching analogies on analogical thinking and mathematics performance

Table 2 presents the means and standard deviations of the study's variables. To examine the effect of teaching on analogical

TABLE 2 Means and standard deviations of pre- and post-teaching scores of AMPT, math accuracy, math processing strategy, and math total in the EX and CN groups.

Test	Phase	Group	M	SD
AMPT	Pre-	EX	4.63	2.65
		CN	5.25	2.45
	Post-	EX	7.04	2.35
		CN	5.42	3.17
Math accuracy	Pre-	EX	6.50	2.95
		CN	6.38	2.90
	Post-	EX	7.79	2.11
		CN	6.54	2.62
Math processing strategy	Pre-	EX	5.96	3.04
		CN	5.63	3.05
	Post-	EX	7.50	2.15
		CN	6.29	3.20
Math total	Pre-	EX	12.46	3.87
		CN	12.00	4.54
	Post-	EX	15.29	3.01
		CN	12.83	4.55

thinking modifiability (near transfer), a repeated measures ANOVA of Treatment $\times$ Time (2×2) for the AMPT scores was conducted. Similarly, repeated measures ANOVAs for the variables (far-transfer) of math were conducted. The ANOVA findings for the AMPT are shown in Table 3 and the ANOVA findings for the math variables are shown in Table 4. Finally, we examined the prediction of Math-Total scores by analogical modifiability using the hierarchical regression procedure. The findings (Tables 3, 4) reveal significant Treatment $\times$ Time interactions for all the dependent variables. The interactions are portrayed in Figures 2–5. A summary of the simple effects analyses of the interactions is shown in Table 5.

Treatment $\times$ time interaction on AMPT scores

The significant Treatment $\times$ Time interaction (Figure 2) revealed that children in the EX group improved their scores from

pre- to post-teaching more than children in the CN group. Simple effects analyses (Table 5) revealed no significant group differences in the pre-teaching phase. However, in the post-teaching phase, children in the EX group outperformed children in the CN group. Within-group analyses revealed a significant improvement from pre- to post-teaching only in the EX group.

Treatment × time interaction on math accuracy scores

The significant interaction of Treatment × Time (Figure 3) indicates that children in the EX group showed higher pre- to post-teaching improvement than children in the CN group. Simple effects analyses (Table 5) revealed no significant group differences in the pre- and post-teaching phases. However, within-group analyses showed that only the EX group showed significant pre- to post-teaching improvement.

Treatment × time interaction on math processing strategy scores

The significant interaction of Treatment × Time (Figure 4) indicates that children in the EX group showed higher pre- to post-teaching improvement than children in the CN group. The simple effects of between-group analyses (Table 5) revealed no significant group differences in either pre- or post-teaching phases. Within-group analyses showed significant pre- to post-teaching improvement in both the EX and CN groups, though the slope of change was steeper in the EX group.

TABLE 3 Repeated measures ANOVA's of AMPT scores by treatment and time.

Source of variation	df	MS	F	η_p^2
Treatment (A)	1	6.00	0.53	0.01
Error	46	11.29		
Time (B)	1	40.04	13.20***	0.22
A × B	1	30.38	10.01**	0.18
Error	46	3.03		

p < 0.01; *p < 0.001.

Treatment × time interaction on math total scores

As can be seen in Figure 5, children in the EX group demonstrated higher pre- to post-teaching improvement than children in the CN group. Simple effects between-group analyses (Table 5) revealed no significant group differences in the pre-teaching phase, but in the post-teaching phase, children in the EX group outperformed children in the CN group. Simple effects of within-group analyses showed that both groups revealed pre- to post-teaching improvement, though the EX group demonstrated a much higher improvement.

Hierarchical regression analysis: prediction of post-teaching MAPS-total by AMPT scores

An intriguing objective of the current study was to assess whether analogical thinking modifiability predicts the modifiability of math performance. According to Hypothesis 3, the modifiability of Total-Math will be predicted by analogical thinking modifiability. In other words, the hypothesis is that a short-term

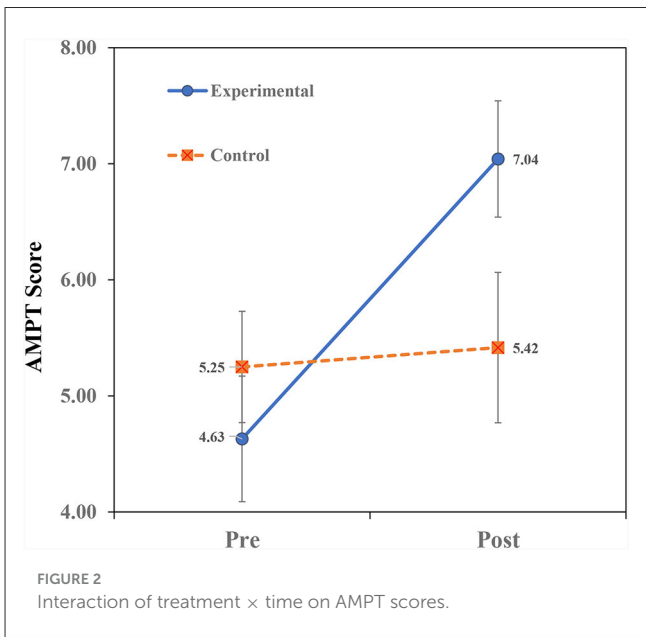


FIGURE 2 Interaction of treatment × time on AMPT scores.

TABLE 4 Repeated measures ANOVAs of math variables: math accuracy, math processing strategy, and math total by treatment and time.

	Math accuracy			Math processing strategy			Math total			
Source of variation	df	MS	F	η_p^2	MS	F	η_p^2	MS	F	η_p^2
Treatment (A)	1	36.26	4.88*	0.10	14.26	0.91	0.02	51.04	1.66	0.04
Error	46	7.44			15.61			30.78		
Time (B)	1	12.76	24.31***	0.35	29.26	27.11***	0.37	80.67	42.98***	0.48
A × B	1	7.59	14.47***	0.24	4.59	4.26*	0.09	24.00	12.79***	0.22
Error	46	0.53			1.08			1.88		

*p < 0.05; ***p < 0.001.

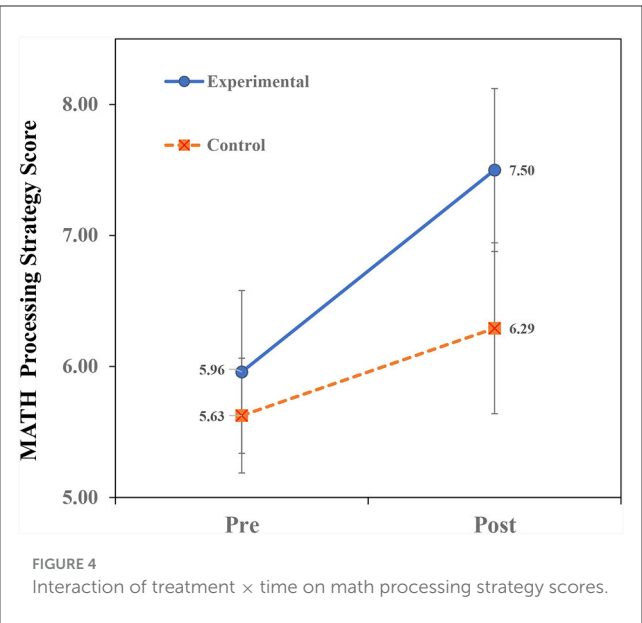
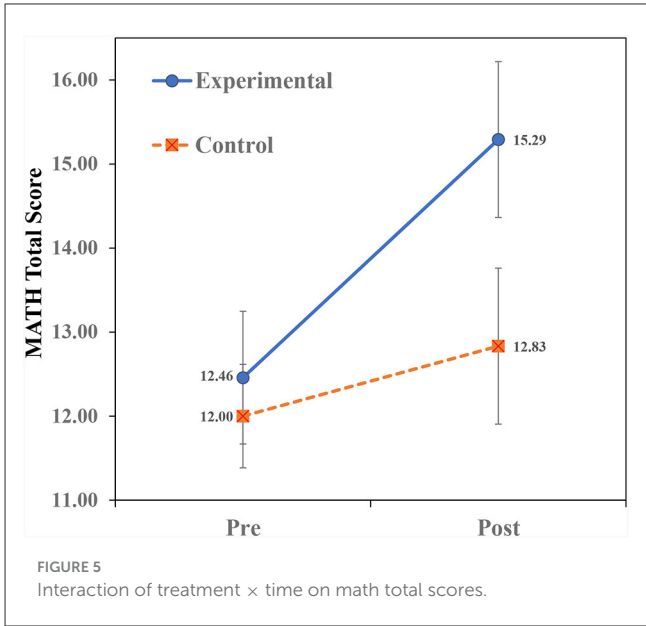
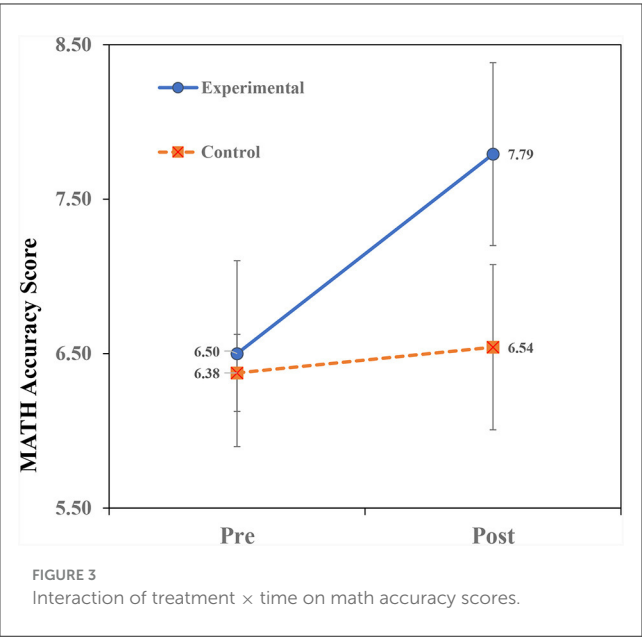


TABLE 5 Simple effects analyses of treatment × time interactions of AMPT, math accuracy, math processing strategy, and math total variables.

Test	Between group		Within groups	
	Phase of testing	<i>t</i> ₍₄₆₎	Group	<i>t</i> ₍₂₃₎
AMPT	Pre-	0.85	EX	5.86***
	Post-	2.02*	CN	0.29
Math accuracy	Pre-	0.15	EX	5.13***
	Post-	1.82	CN	1.07
Math processing strategy	Pre-	0.38	EX	4.05***
	Post-	1.54	CN	3.56**
Math total	Pre-	0.38	EX	5.85***
	Post-	2.21*	CN	2.97**

p* < 0.05, *p* < 0.01, ****p* < 0.001.

intervention of analogical thinking within a DA procedure will affect not only analogical thinking but will be transferred to math thinking. This hypothesis was evaluated by a three-step hierarchical regression analysis where the predicted variable (criterion) was post-intervention MAPS-Total. In Step 1, the variable of age (in months) was introduced. In Step 2, the variables of pre-teaching variables of MAPS-Total and AMPT were introduced, and in Step 3, the post-teaching variables of MAPS-Total and AMPT were introduced. The modifiability of MAPS-Total score was indicated by the post-teaching score after statistically “washing out” the variance contributed by the MAPS-Total pre-teaching score. Similarly, the modifiability of AMPT score was demonstrated by the post-teaching AMPT score after statistically “washing out” for the variance contributed by the AMPT pre-teaching score. The

results, as shown in Table 6, indicate that Age was not a significant predictor of the post-intervention math performance. The variance of Age was negligible, $F_{(1,46)} = 0.01$, $p < 0.92$, and $R^2 = 0.00$. In Step 2, however, as expected, the pre-intervention math score was a significant predictor of the post-intervention math score ($\beta = 0.86$). The pre-intervention AMPT score was not found as a significant predictor ($\beta = 0.04$). The significant prediction of Step 2 added 75% to the prediction of Math-Total score, $F_{(3,44)} = 45.63$, $p < 0.0001$, $\Delta R^2 = 0.75$, $p < 0.0001$.

In Step 3, after controlling for the variance of the pre-intervention variables, the post-intervention AMPT variable has emerged as a significant predictor ($\beta = 0.24$) of post-teaching MAPS-Total (criterion variable), thus confirming Hypothesis 3. The overall prediction of post-intervention MAPS score was significant, $F_{(4,43)} = 41.46$, $p < 0.0001$, $R^2 = 0.79$. The AMPT post-intervention score, as an indicator of analogical modifiability, added significantly to the prediction of post-intervention Math-Total, $\Delta R^2 = 0.04$, $p < 0.008$.

TABLE 6 Hierarchical regression analyses of post-teaching math-total score by pre- and post-intervention AMPT scores.

Step	Predicting variables	β	t	p
1	Age	0.01	0.11	0.92
		$R^2 = 0.00, \Delta R^2 = 0.00$		
		$F_{(1,46)} = 0.01, p = 0.92$		
2	Age	-0.12	-0.157	0.12
	Pre-AMTP	0.04	0.53	0.60
	Pre-math-total	0.86	10.42	0.0001
		$R^2 = 0.75, \Delta R^2 = 0.75, p < 0.0001$		
		$F_{(3,44)} = 45.63, p < 0.0001$		
3	Age	-0.08	-1.06	0.30
	Pre-AMTP	-0.06	-0.68	0.50
	Pre-math-total	0.81	10.18	0.0001
	Post-AMPT	0.24	2.79	0.008
		$R^2 = 0.79, \Delta R^2 = 0.04, p < 0.008$		
		$F_{(4,43)} = 41.46, p < 0.0001$		

Discussion

The results of this study support the hypothesis that the mediation of analogical thinking within a DA procedure will significantly improve both analogical thinking (Hypothesis 1, near transfer) and math performance (Hypothesis 2, far transfer). These findings agree with the previous findings that showed an intimate relationship between analogical thinking and math thinking, especially with math processing strategies (Núñez and Lakoff, 2005; Richland et al., 2007; Vamvakoussi, 2017, 2019). However, the results of this study go beyond previous studies by indicating that a short-term intervention in analogical thinking can be transferred to math performance, especially to math processing strategies. Math problem-solving has been conceptualized as a transfer challenge requiring students to apply strategies, skills, and knowledge to novel problems (Fuchs and Fuchs, 2005). This idea is congruent with the schema construction theory according to which students' development of schemas occurs when students group problems into types that require the same solution (e.g., Gick and Holyoak, 1983; Chen, 1999). It should be emphasized that the intervention for solving the AMPT analogies was not limited only to specific strategies of solving analogies *per se* but included mediation for the control of impulsivity, considering two or more sources of information simultaneously, activation of WM, systematic exploratory behavior, and verbal anticipation of a solution. There was no way of separating these cognitive functions from the content aspects of the visual analogies of the AMPT.

It is important to note that math processing has much in common with analogical thinking. For example, the question of

“In how much the sum of $64 + 8$ is smaller than $64 + 28$?” requires a simple analogy—if 8 is smaller than 28 in 20 and 64 is a constant, then $64 + 8$ is smaller in 20 than $64 + 28$. The analogical processing is a much more efficient way to solve the problem than the meticulous procedure of adding up the sums, calculating each segment, and subtracting one from the other.

The teaching of AMPT items within the DA procedure was found in an earlier study on Grade 1 children to be powerful in improving not only their analogical thinking but was also transferred to significant improvement of spatial WM (Tzuriel, 2022). The improvement in spatial working memory was impressive, as indicated by a more powerful effect ($\eta_p^2 = 0.61$) than that found on analogical thinking ($\eta_p^2 = 0.40$), which was the targeted domain of mediation strategies. Moreover, the correlations of AMPT analogies (both pre- and post-teaching) with spatial WM was.47 ($p < 0.001$). It should be emphasized that the contents of the MAPS tasks as well as the spatial WM task were different from the contents of the intervention contents, as represented by the AMPT analogies; thus, the improvement would be attributed to the mediation characteristics of the AMPT. Our findings might pave the way for long-term intervention studies aimed at establishing the effects of analogical thinking intervention on the development of math skills. The question is, of course, to what extent is this transfer durable, especially since the intervention was short within a DA procedure and the well-known phenomenon of fading-out of intervention effects (e.g., Bailey et al., 2020).

The results of the hierarchical regression analyses confirm Hypothesis 3. The results show that the variable of post-teaching AMPT significantly contributed to the prediction of math ability after controlling for the variables of pre-teaching scores of analogical thinking and math. These intriguing findings imply that children who demonstrate analogical thinking modifiability in a DA procedure are also those who show pre- to post-intervention modifiability in mathematical thinking. The meaning of this finding is that a short-term mediation within a DA procedure was transferred to math domain enhancing both math accuracy and math processing strategies. This conclusion, however, should be studied further with different age groups, varying the duration of the intervention and employing different analogy tests. It should be emphasized that the finding of a far transfer after a short-term intervention does not mean that the change is durable. One should bear in mind that the intervention in the current study was within a DA procedure, which provides an important indication for the transfer of intervention of analogical thinking to mathematical processing. For a durable long-lasting effect of the intervention, a more comprehensive intervention and a follow-up assessment should be performed. The finding of a far-transfer of analogical thinking to math processing strategies is of special importance for teaching math to students who experience difficulties in processing math problems. Teaching analogical problem-solving might facilitate math processing even more than the direct teaching of math.

Limitations

The study has several limitations. First, the study was conducted on a limited sample of 48 Grade 2 students. Despite the

impressive findings of far-transfer effects from analogical thinking training to math performance, more research involving bigger samples is required. Second, the MAPS test is composed of only 10 items, which limits the validity of findings. Although there is a separate score for accuracy and for processing strategy (a total of 20) for each item, a much larger scale with more items would help to establish the findings. In fact, in a further unpublished study (Tzuriel, 2024), steps were made to enlarge the number of items in MAPS. The third limitation is related to the lack of recording of the specific mediation strategies used in the intervention phase of the analogies. The assumption is that all examiners use the same mediation strategies. A higher resolution of the specific strategies used and their effect on the transfer to math domain would facilitate not only theoretical construction but also practical implications for teachers.

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 involving human participants was reviewed and approved by the Ethical Committee of Talpiot Academic College. The study was conducted in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the parents of children prior to start of data gathering and based on Israel Ministry of Education requirements.

References

- Alexander, A. (2017). Relational reasoning in STEM domains: a foundation for academic development. *Educ. Psychol. Rev.* 29, 1–10. doi: 10.1007/s10648-016-9383-1
- Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422
- Bailey, H., Duncan, C. J., Cunha, F., Foorman, B. R., and Yeager, S. (2020). Persistence and fade-out of educational intervention effects: mechanisms and potential solutions. *Psychol. Sci. Public Int.* 21, 55–97. doi: 10.1177/1529100620915848
- Bassok, M., and Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *J. Exp. Psychol. Learn. Mem. Cognit.* 15, 153–166. doi: 10.1037//0278-7393.15.1.153
- Chen, Z. (1999). Schema induction in children's analogical problem-solving. *J. Educ. Psychol.* 91, 703–715. doi: 10.1037/0022-0663.91.4.703
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *J. Res. Sci. Teach.* 30, 1241–1257. doi: 10.1002/tea.3660301007
- Clement, J. (2013). "Roles for explanatory models and analogies in conceptual change," in *International Handbook of Research on Conceptual Change, 2nd Edn*, ed S. Vosniadou (Abingdon-on-Thames, Oxfordshire: Routledge), 412–446.
- Dantzig, T. (2005). *Number: The Language of Science, 4th Edn*. London: Pi Press.
- English, L. D. (2004). "Mathematical and analogical reasoning in early childhood," in *Mathematical and Analogical Reasoning of Young Learners*, ed L. D. English (Abingdon-on-Thames, Oxfordshire: Routledge), 13–34.
- Feuerstein, R., Feuerstein, R. S., Falik, L. H., and Rand, Y. (2002). *The Dynamic Assessment of Cognitive Modifiability: The Learning Propensity Assessment Device*. Jerusalem: ICELP Press.
- Fuchs, L., and Fuchs, D. (2005). Enhancing mathematical problem solving for students with disabilities. *J. Spec. Educ.* 39, 45–57. doi: 10.1177/00224669050390010501
- Gelman, R. (2000). The epigenesis of mathematical thinking. *J. Appl. Dev. Psychol.* 21, 27–37. doi: 10.1016/S0193-3973(99)00048-9
- Gentner, D., Loewenstein, J., and Thompson, L. (2003). Learning and transfer: a general role for analogical encoding. *J. Educ. Psychol.* 95, 393–408. doi: 10.1037/0022-0663.95.2.393
- Gick, M. L., and Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cogn. Psychol.* 15, 1–38. doi: 10.1016/0010-0285(83)90002-6
- Gick, M. L., and Holyoak, K. J. (1987). "The cognitive basis of knowledge transfer," in *Transfer of Learning: Contemporary Research and Application*, eds S. M. Cormier, and J. D. Hagman (Cambridge, MA: Academic Press), 9–45.
- Givvin, K. B., Stigler, J. W., and Thompson, B. J. (2011). What community college developmental mathematics students understand about mathematics. Part II: the interviews. *MathAMATYC Educ.* 2, 4–18.
- Gonzales, P., Calsyn, C., Jocelyn, L. J., Mak, K., Kastberg, A.rafah, S., Williams, T., et al. (2001). *Pursuing Excellence: Comparisons of International Eighth-Grade Mathematics and Science Achievement from a U.S. Perspective, 1995–1999. Initial Findings from the Third International Mathematics and Science Study–Repeat. Statistical Analysis Report*. Washington, DC: National Center for Education Statistics, US Department of Education.
- Goswami, U. (1992). *Analogical Reasoning in Children*. Mahwah, NJ: Erlbaum.
- Holyoak, K. J. (2004). "Analogy," in *The Cambridge Handbook of Thinking and Reasoning*, eds K. J. Holyoak, and R. J. Forrison (Cambridge: Cambridge University Press), 117–142.

Author contributions

DT: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing.

Funding

The author declares that financial support for this publication was received from Bar-Ilan University and Talpiot Academic College.

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.

- Mullis, I. V. S., Martin, M. O., Gonzales, E. J., and Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report: Findings From IEAs Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Washington, DC: TIMSS and PIRLS International Study Center.
- National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.
- Núñez, R., and Lakoff, G. (2005). "The cognitive foundations of mathematics: the role of conceptual metaphor," in *Handbook of Mathematical Cognition*, ed J. I. D. Campbell (Hove, East Sussex: Psychology Press), 109–124.
- Resing, W. C. M. (2000). "Assessing the learning potential for inductive reasoning (LIR) in young children," in *Dynamic Assessment: Prevailing Models and Applications*, eds C. S. Lidz, and J. G. Elliott (JAI/Elsevier Science), 229–262.
- Richland, E. A., and Simms, N. (2015). Analogy, higher order thinking, and education. *WIREs Cogn. Sci.* 6, 177–192. doi: 10.1002/wcs.1336
- Richland, L. E., and McDonough, I. M. (2010). Learning by analogy: discriminating between potential analogs. *Contemp. Educ. Psychol.* 35, 28–43. doi: 10.1016/j.cedpsych.2009.09.001
- Richland, L. E., Zur, O., and Holyoak, K. J. (2007). Cognitive supports analogies in the mathematics classroom. *Science* 316, 1128–1129. doi: 10.1126/science.1142103
- Rittle-Johnson, B., and Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *J. Educ. Psychol.* 99, 561–574. doi: 10.1037/0022-0663.99.3.561
- Sternberg, R. J., and Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Dev.* 51, 27–38. doi: 10.2307/1129586
- Stigler, J. W., Givvin, K. B., and Thompson, B. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educ.* 10, 4–16.
- Tzuriel, D. (1997). A novel dynamic assessment approach for young children: major dimensions and current research. *Educ. Child Psychol.* 14, 83–108.
- Tzuriel, D. (2000). Dynamic assessment of young children: educational and intervention perspectives. *Educ. Psychol. Rev.* 12, 385–435.
- Tzuriel, D. (2001). *Dynamic Assessment of Young Children*. New York, NY: Kluwer Academic/Plenum Press.
- Tzuriel, D. (2002). "Dynamic assessment of learning potential," in *Encyclopedia of Education*, 2nd edn. (McMillan Press), 127–131.
- Tzuriel, D. (2020). "Dynamic cognitive assessment for preschool age children," in *Oxford Research Encyclopedia of Education*, ed L. F. Zhang (Oxford University Press). doi: 10.1093/acrefore/9780190264093.013.942
- Tzuriel, D. (2021a). *Mediated Learning Experience and Cognitive Modifiability*. Springer Nature.
- Tzuriel, D. (2021b). *Analogical Modifiability Puzzle Test-Instruction Manual*. Bar-Ilan University.
- Tzuriel, D. (2021c). *The Childrens' Spatial Working Memory (CSWM) test. Instruction Manual*. Faculty of Education, Bar-Ilan University, Israel.
- Tzuriel, D. (2022). Construction versus closed analogies: Effects of mediation and relation to working memory. *Formazione Insegnamento.* 20, 34–49. doi: 10.7346/-feis-XX-02-22_04
- Tzuriel, D. (2024). *Math Accuracy and Processing Strategy (Maps) Test* (Unpublished Test). Talpiot Academic College, Israel.
- Tzuriel, D., and George, T. (2009). Improvement of analogical reasoning and academic achievement by the Analogical Reasoning Programme (ARP). *Educ. Child Psychol.* 26, 71–94. doi: 10.53841/bpsecp.2009.26.3.71
- Tzuriel, D., and Klein, P. S. (1985). Analogical thinking modifiability in disadvantaged, regular, special education, and mentally retarded children. *J. Abnorm. Child Psychol.* 13, 539–552. doi: 10.1007/BF00923140
- Tzuriel, D., and Shamir, A. (2010). Mediation strategies and cognitive modifiability in young children as a function of peer mediation with young children program and training in analogies versus math tasks. *J. Cogn. Educ. Psychol.* 9, 48–72. doi: 10.1891/1945-8959.9.1.48
- Vamvakoussi, X. (2017). Using analogies to facilitate conceptual change in mathematics learning. *ZDM* 49, 497–507. doi: 10.1007/s11858-017-0857-5
- Vamvakoussi, X. (2019). The use of analogies in mathematics instruction: affordances and challenges. *Math. Cognit. Learn.* 5, 247–268. doi: 10.1016/B978-0-12-815952-1.00010-4
- Vamvakoussi, X., and Vosniadou, S. (2012). Bridging the gap between the dense and the discrete: the number line and the "rubber line" bridging analogy. *Math. Think. Learn.* 14, 265–284. doi: 10.1080/10986065.2012.717378
- Zhao, H., Alexander, P. A., and Sun, Y. (2021). Relational reasoning's contributions to mathematical thinking and performance in Chinese elementary and middle-school students. *J. Educ. Psychol.* 113, 279–303. doi: 10.1037/edu0000595



OPEN ACCESS

EDITED BY

Luisa Losada-Puente,
University of A Coruña, Spain

REVIEWED BY

Kevin Mark Roessger,
University of Arkansas, United States
Pedro Simão Mendes,
University Institute of Lisbon, Portugal

*CORRESPONDENCE

Juana M. Ortega-Tudela
✉ jmortega@ujaen.es

RECEIVED 02 September 2023

ACCEPTED 27 February 2024

PUBLISHED 14 March 2024

CITATION

Lechuga MT, Ortega-Tudela JM and
Gómez-Ariza CJ (2024) Retrieval-based
concept mapping makes a difference as a
retrieval practice activity: a study with high
school students.
Front. Educ. 9:1287744.
doi: 10.3389/feduc.2024.1287744

COPYRIGHT

© 2024 Lechuga, Ortega-Tudela and
Gómez-Ariza. 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.

Retrieval-based concept mapping makes a difference as a retrieval practice activity: a study with high school students

María Teresa Lechuga¹, Juana M. Ortega-Tudela^{2*} and
Carlos J. Gómez-Ariza¹

¹Department of Psychology, University of Jaén, Jaén, Spain, ²Department of Pedagogy, University of Jaén, Jaén, Spain

Introduction: While specific sequences of retrieval-based activities have been shown to have a powerful effect on learning, no previous study has examined this issue in children or adolescents. Here, we aimed to determine whether the benefit of concept mapping as an initial retrieval activity observed previously with college students in a lab setting may also be found in younger and less experienced students in a more naturalistic school setting.

Methods: After a short training in concept mapping, participants (N = 60) read an educational text and then engaged in a sequence of two learning activities that required them to retrieve relevant ideas from the text. One of the activities involved free recall by writing down as many ideas as possible and the other involved creating a concept map, both of them in the absence of educational texts. Critically, we manipulated the order in which the activities were performed.

Results: A mediation analysis with success during retrieval practice as the mediator revealed both a direct effect of retrieval sequence and an indirect effect. Creating a concept map first and then freely recalling by writing paragraphs significantly improved performance on a 2-week delayed learning test, as compared to performing the same activities in the inverse order, even when doing concept mapping first led to lower success rates during practice.

Discussion: These results support the idea that concept learning from instructional texts can be modulated by the sequence of retrieval activities performed and point to the educational value of retrieval-based concept mapping as a first learning activity within a series across different learners and educational settings.

KEYWORDS

concept mapping, retrieval-based learning, teaching strategies, adolescents, testing effect

1 Introduction

Findings from a good number of studies show that retrieval-based study activities can be highly effective in promoting long-term learning across learner characteristics, materials, and contexts (Karpicke, 2017; Moreira et al., 2019; Yang et al., 2021). In retrieval-based learning, students engage in activities that require them to repeatedly recall previously studied information (e.g., answering a teacher's questions or self-created flashcards) once study materials are set aside. Importantly, it has been shown that retrieval practice may lead to better

student performance (the so-called “testing effect”) on subsequent learning tests than do other traditional learning activities such as rereading or concept mapping (Adesope et al., 2017, for a meta-analysis; Nunes and Karpicke, 2015, for a review), and that this benefit can be observed either in laboratory or natural school settings, including when genuine curriculum materials are used (Moreira et al., 2019; Lamotte et al., 2021, for reviews; see also Ortega-Tudela et al., 2021).

A common retrieval strategy that has been systematically considered in studies is free recall [i.e., participants are asked to write down all of the information they can remember once learning material has been removed; see reviews by Nunes and Karpicke (2015) and Karpicke (2017)], which has typically been compared with extra study (i.e., rereading) of materials. Roediger and Karpicke (2006), for example, compared two retrieval practice conditions (one involving 3 study periods and 1 retrieval period, and another comprising 1 period of study and 3 of retrieval) with rereading (4 periods of study), matching total time of learning activities across the different conditions. Participants’ learning was tested either 5 min or 1 week after they undertook the learning activities. Although there was a greater benefit from the rereading condition in the immediate test, it was the group that had engaged in repeated retrieval (free recall) that exhibited better retention in the delayed final assessment.

In addition to rereading, free recall has also been compared to more elaborative study strategies such as concept mapping (Karpicke and Blunt, 2011; Lechuga et al., 2015; O’Day and Karpicke, 2021; for other elaborative strategies see Fritz et al., 2007; Karpicke and Smith, 2012). Elaborative processes are thought to boost the encoding of relational information as well as the distinctiveness of specific ideas and facts (Craik and Tulving, 1975; Hunt, 2012), and concept map creation is considered to rely to a large extent on the contribution of such processes to meaningful learning (Novak and Gowin, 1984; Novak and Cañas, 2006). Indeed, over the course of recent years, concept maps, which are node-and-link diagrams that hierarchically represent relevant concepts and relations among them within a knowledge domain, have gained extraordinary popularity among educators and students (Chevron, 2014; Kinchin, 2014; Karpicke, 2018; Schroeder et al., 2018). Strikingly, studies that have compared retrieval practice (i.e., free recall by writing paragraphs) to concept mapping have mostly failed to show the latter to be superior as a learning tool (Karpicke and Blunt, 2011; Lechuga et al., 2015; O’Day and Karpicke, 2021). By using educational texts as materials, Karpicke and Blunt (2011; see also O’Day and Karpicke, 2021, Exp. 2) showed that repeated free recall produced better learning than creating a concept map (while viewing the text) on a final test 1 week later. Lechuga et al. (2015) essentially replicated this effect, even though they also showed that it was less pronounced (only affected verbatim questions, not inferential questions) for participants with experience in using concept maps than it was for those who undertook a short training in concept mapping.

A matter that has recently begun to draw researchers’ attention is the identification of the most effective ways to enhance student learning by combining different learning (including retrieval-based) activities (e.g., see McDaniel, 2023; Roelle et al., 2023, for recent reviews). Thus, for example, Karpicke and Smith (2012) examined whether combining an imagery-based elaborative strategy (the keyword method; McDaniel and Pressley, 1984) with either restudy or retrieval practice improved the learning of uncommon words. The

results revealed that practicing retrieval was the most effective strategy to learn new words (as measured on a 1-week delayed final test) regardless of whether it was coupled with the elaborative strategy. In other words, the combination of elaborative encoding with repeated retrieval failed to produce a benefit beyond that of simply practicing retrieval (see Fritz et al., 2007, for a related finding with 12–13 year-old children). However, and of special relevance here, research has highlighted that distinct learning activities can offer unique benefits to the learning process (see McDaniel, 2023; Roelle et al., 2023). While generative learning activities could foster the development of coherent conceptual representations that are well integrated with prior knowledge, retrieval practice activities could promote knowledge consolidation in memory. The idea that generative activities, such as concept mapping, enhance meaningful learning is well supported by a number of studies (see Fiorella and Mayer, 2016). Thus, because concept mapping seems to rely on elaborative processing and enhance the organizational representation of concepts in memory (Novak and Gowin, 1984; Nesbit and Adesope, 2006; Chevron, 2014; see Karpicke, 2018 for a critical analysis), it is thus not surprising that research is being conducted to determine whether the coupling of concept mapping and retrieval practice may be particularly effective to promote long-term learning.

To date, although research on this issue is still limited, retrieval and concept mapping activities have been combined in two different ways. On the one hand, one study examined the effectiveness of engaging students in concept mapping and free recall, as separate activities, within a sequence of learning events. Thus, O’Day and Karpicke (2021; Exp. 2) compared retrieval practice (5-min initial study followed by free recall for 20 min, restudy for 5 min, and free recall for 10 min; 30 min of total time) against a combination of concept mapping and retrieval practice (initial study for 5 min followed by concept mapping for 20 min, free recall for 10 min, restudy for 5 min, and free recall for 10 min; 50 min of total time). One reason to expect that the combination of creating a concept map first and then practicing retrieval might be particularly effective is that it is assumed that concept mapping enhances students’ relational representation of concepts within materials. Hence, this organizational representation might subsequently provide more elaborated retrieval structures that would guide search processes in memory and facilitate access to relevant information (either in subsequent retrieval practice periods or in learning tests) (Hunt, 2012; Kahana, 2017; Roelle et al., 2022, 2023; McDaniel, 2023). One week after the learning session, participants took a short answer test comprising verbatim and inference questions. Results revealed that coupling concept mapping and free recall did not improve learning as compared to repeated retrieval alone, despite the fact that the combined condition significantly increased (in 20 min) the total time spent on the learning session. Again, adding an elaborative study strategy (concept mapping in this case) in a sequence of learning activities did not produce a benefit beyond that of simply carrying out retrieval practice.

In other cases, retrieval practice has been combined with concept mapping by embedding the act of recall in the process of creating a concept map. While students typically construct concept maps in the presence of materials (the texts aid them in identifying main concepts and relations; as in O’Day and Karpicke’s study, for example), an alternative way of elaborating a concept map is to draw it up in the absence of materials, which necessarily requires students to engage in retrieving information from memory. Interestingly, Blunt and

Karpicke (2014) tested the potential effectiveness of retrieval-based concept mapping across two experiments. In their first one, and after a 5-min reading period, participants practiced retrieval of information from a short science text. Importantly, half of participants did so by writing down as many ideas as they could, whereas the other half were required to recall by creating a concept map. In both cases retrieval was done in the absence of texts and with two periods of practice. One week later, the participants returned to the lab to complete a short-answer learning test that assessed verbatim and inferential knowledge. The results showed performance from both retrieval formats on the final test to be virtually equivalent. In the second experiment, the researchers factorially crossed retrieval format and availability of texts. Again, no difference between the two retrieval formats was observed, even though both concept mapping and writing down ideas in the absence of materials produced better learning than completing the same activities with the materials at hand.

From their findings, Blunt and Karpicke (2014) argued that retrieval-based concept mapping is as effective as freely recalling by writing paragraphs, thus suggesting that the benefit attributed to both the relational and item-specific processing required in concept mapping is redundant with the processing associated with repeated retrieval. Altogether, the results of the above-mentioned studies that combined concept mapping (and also other elaborative strategies such as Fritz et al., 2007; Karpicke and Smith, 2012 did) with retrieval practice would in fact seem to indicate that no additional benefit should be expected beyond the one provided by the very fact of practicing retrieval. Hence, an interim conclusion that might be drawn from the described evidence is that it is the very act of retrieving information that promotes learning the most; accordingly, freely recalling (by writing down, for example) might be the simplest and best choice when it comes to learning from educational texts, and concept mapping (either retrieval-based or not) might not be a worthwhile learning activity to engage in since it is effortful and time-consuming (Quinn et al., 2003; Mintzes et al., 2011). The results of a recent study by Ortega-Tudela et al. (2019), however, suggest a different picture.

In their initial experiment, Ortega-Tudela et al. (2019) replicated the main finding from Blunt and Karpicke (2014), such that paragraph writing (PW) and concept mapping (CM), both as retrieval-based activities, led to similar levels of performance in a 2-week (instead of 1-week) delayed learning test. After successfully replicating Blunt and Karpicke's main finding, Ortega-Tudela et al. (2019) designed a second experiment to further test the idea that the relational and item-specific processing involved in concept mapping is redundant with retrieval practice. Their prediction was straightforward: If both concept mapping and paragraph writing down, as retrieval-based activities, engage redundant cognitive processing with equivalent learning outcomes (as Blunt and Karpicke, 2014 suggested), then one would expect two groups of participants doing retrieval practice in both formats but performing the activities in a different order (CM + PW vs. PW + CM) to exhibit similar performance on a learning test. One might predict, however, that if the fact of elaborating a concept map provides a unique opportunity for later retrieval [i.e., because of the organizational/relational processing that it is assumed to demand; see O'Day and Karpicke (2021) for a similar hypothesis when it comes to creating a concept map with materials made available], then performance on the final learning test should be dependent on the position that retrieval-based concept mapping takes in the sequence

of learning activities. Specifically, it might be expected that the sequence CM + PW would give rise to better learning than PW + CM. This was, precisely, the main result observed by Ortega-Tudela et al. (2019). On average, those participants who first elaborated concept maps outperformed the participants who first retrieved ideas by writing them down as paragraphs, with this difference not being attributable to the quality of the concept maps created by participants from both groups. Interestingly, these results are in line with what Roelle et al. (2023) call "construction-before-consolidation," suggesting the potential advantages of practicing generative learning prior to retrieval practice activities (see also Roelle and Nückles, 2019). Nevertheless, results from other recent studies seem to favor retrieval-practice-first sequences (e.g., Roelle et al., 2022), highlighting the need for further exploration. As suggested by Roelle et al. (2023), there may not be a singular optimal learning sequence, and the effects of generative-first or retrieval-practice-first sequences could be influenced by moderating variables. In any case, the sequencing effect obtained by Ortega-Tudela et al. (2019) might be of particular interest in gaining an understanding of effective combinations of learning strategies and merits further investigation. Although it stands in stark contrast to results from other (still scarce) studies showing that concept mapping does not contribute to retrieval practice as a learning tool (Karpicke and Blunt, 2011; Lechuga et al., 2015; O'Day and Karpicke, 2021), it lends empirical support to the widespread idea that concept mapping relies on cognitive operations that may be especially effective for learning from educational texts (Novak and Gowin, 1984; Chevron, 2014; Karpicke, 2018). In addition, the finding points to the need to further consider the nature of distinct retrieval activities when it comes to boosting learning, and is particularly suggestive of the potential educational value of creating concept maps in the absence of educational texts as a first stage in a sequence of learning activities.

The educational implications of this sequencing effect could be of special relevance in the case of primary and secondary education students, in whom performance on learning tasks has a greater potential for improvement. However, to date, no other studies have addressed the issue of combining and sequencing in respect of retrieval-based concept mapping, aside from the original work by Ortega-Tudela et al. (2019). Thus, in the present study, we aimed to determine whether the sequencing effect observed in Ortega-Tudela et al. (2019) with college students may also be found in learners attending high school. The replication of the main finding obtained by Ortega-Tudela et al. (2019)—with a sample of younger and less experienced students—in a more naturalistic school setting would contribute significantly to identifying how retrieval-based activities (especially concept mapping) may be used more effectively and in distinct educational contexts.

While we expected to find the above-mentioned sequencing effect in 16–17-year-old students, we also considered that several factors could modulate such an effect in the present study. Firstly, episodic memory functioning and its neural underpinnings undergo changes throughout adolescence and young adulthood, particularly regarding strategic components of encoding and retrieval (Shing et al., 2008, 2010; Mechie et al., 2021). Thus, we could anticipate limited benefits of elaborating concept maps first in our sample due to the lower success rates during retrieval practice. Furthermore, as self-regulation abilities improve with age across the middle and high school grades, so does the spontaneous deployment of retrieval practice as a learning

activity (e.g., Tullis and Maddox, 2020). Consequently, the benefit of concept mapping first observed among college students may be less prominent in younger learners. On the other hand, conducting the study in a naturalistic educational context (the entire research process was conducted in classrooms as part of academic activities) poses important technical challenges that could reduce the sensitivity to observe experimental effects. However, we took advantage of an outstanding opportunity to conduct research beyond the laboratory because the Spanish Education System includes a cross-curricular competence in the 11th and 12th grades curricula that focuses on 'learning to learn', which allowed us to more easily justify our presence in the classrooms. Thus, following closely the procedure used in the second study experiment from Ortega-Tudela et al. (2019), but embedding it into the curricular content of one academic activity oriented to enhance the students' skills in preparing for exams, in the present study we directly compare the effect of the sequence of retrieval activities on a learning test taken 2 weeks later. If creating a concept map first in the absence of texts facilitates the retrieval of relevant isolated concepts, with a focus on relations among the reached concepts, thus increasing their accessibility and organization over future retrieval attempts, then better learning should be observed in students who are able to create concept maps of sufficient quality as the first retrieval activity. Although research among young students might not support the preference for concept mapping as a learning tool, primary and secondary education students as well as adults have been shown to benefit from concept mapping training and guidance (Habók, 2012; Gallenstein, 2013; Roessger et al., 2018). Thus, in the present study a brief training in concept mapping was provided to ensure understanding and practice on why and how to elaborate concept maps. Hence, our prediction was that adolescents attending high school would also benefit more from the sequence CM + PW than from the sequence PW + CM.

In addition to performance (either on the retrieval practice periods or the learning test), we assessed participants' own outcome expectations as well as academic self-concept and overall use of learning strategies, in order to control for their potential effect on final learning. Because the format of the retrieval activity and prior learning strategy experience may bias participants' metacognitive judgements and expectations about their final performance, we asked them to provide judgments of learning immediately following their completion of the first retrieval trial, in any format. Although retrieval-based activities generally do not generate high expectations for learning outcomes (Rivers, 2021), we intended to explore whether such judgments were reliant on the type of retrieval activity that learners performed. Additionally, we included a short questionnaire to assess academic self-concept due to its reciprocal relationships with the use of learning strategies and achievement (McInerney et al., 2012). Finally, we also assessed the frequency with which participants utilized various study techniques to control for potential differences between the groups in this regard (Miyatsu et al., 2018).

2 Method

2.1 Participants

The study was conducted at a public high school located in a middle-class neighborhood in the city of Jaén (Spain), as part of a

school time activity organized by the Educational Orientation Department. Partial information about the study was first presented to 101 students from the first and second years of the bachelor's degree in the Spanish educational system (Grades 11 and 12), who were participating at that moment in a series of academic activities on learning strategies (as part of the transversal curricular content promoted by the school)¹. Ultimately, 64 students agreed to participate in the study. However, we detected 4 outliers (two participants from each experimental group) who were no longer considered for analyses. These were students who produced less than two ideas (out of 25) from the studied text in any of the two retrieval conditions during the first practice trial. Thus, the final sample consisted of 60 participants. None of them had specific learning difficulties. The school's team approved the activities carried out in the study as part of the academic program, and the University's Research Ethics Committee approved the study. Informed consent and assent form were obtained prior to participation.

2.2 Materials and procedure

The experiment was carried out over three independent sessions (one training session and two assessment sessions) with a delay of 14 days between them. In the first session, participants were informed in general about the study and required to provide their informed consent to participate in the activities.

Before the training session started, the academic self-concept scale and the frequency of using study strategies were completed. An academic self-concept scale validated for use with Spanish-speaking adolescents was selected (Academic Self-Concept for Adolescents, ASCA; Ordaz-Villegas et al., 2013). The ASCA scale consists of 28 Likert items ranging from 1 (never) to 5 (always), loading on four different factors: self-regulation, general intellectual abilities, motivation, and creativity. With the original validation sample ($n=347$; ranging 14–18 years old), Ordaz-Villegas et al. (2013) reported a global Cronbach's alpha of 0.828. In our sample, the global Cronbach's alpha was 0.750 (specific alphas: self-regulation = 0.555; general intellectual abilities = 0.517; motivation = 0.478, and creativity = 0.569).

The students also filled out a short questionnaire about their use of diverse learning strategies (Hábitos y Técnicas de Estudio, CHTE; Álvarez and Fernández, 1999). The CHTE comprises 12 Likert-type items (scoring frequency of use from 1, never, to 4, always) assessing various student practices. We specifically selected those that directly relate to common study strategies for preparation of subsequent learning test. These include underlining, schema creation,

¹ Although we were aware of the difficulty of collecting data from a large sample in a naturalistic setting, the ideal sample size was determined in advance by assuming a medium effect size (like the one observed in Ortega-Tudela et al., 2019). A power analysis performed with G*Power (Faul et al., 2007) determined that a minimal total sample size of 98 participants would be necessary to obtain a statistically significant effect from the sequencing of the retrieval-based learning activities (power = 0.80 and $\alpha = 0.05$; ANOVA: Repeated measures, between factor).

summarizing, concept mapping, and self-generation of questions. Only information on frequency of use of these strategies was collected.

After completing the questionnaires, students were provided with a short training on concept mapping in their natural classroom groups (see also [Lechuga et al., 2015](#); [Ortega-Tudela et al., 2019](#)). The program consisted of three steps. First, a general explanation of concept maps was given as well as a brief description of their usefulness for students. Then, participants were taught how to elaborate a concept map. They were given a text (about different types of forests) and one of the researchers created a concept map from the text by drawing on the blackboard, emphasizing the various characteristics and possibilities. Finally, the participants were given a second text (on factors to be considered when carrying out an academic task) and, in groups of two, they were instructed to work on creating their own concept map from this particular text. Two of the researchers supervised the students to ensure that the maps were constructed according to the instructions given and the researchers also answered any questions on the process. The students' usual teacher was also involved in supervising and in answering questions. Both texts and procedures used during the training were adapted from [Repetto et al.'s \(2004\)](#) program, as used in [Lechuga et al. \(2015\)](#). Session 1 lasted approximately 55 min.

The second session took place 2 weeks later and lasted about 25 min. Only students who expressed interest in participating attended sessions 2 and 3. Students belonging to the same class performed the study session as a natural group. All of them were encouraged to read a text for 5 min. The text (taken from [Repetto et al.'s \(2004\)](#) program and also used in previous studies: [Lechuga et al., 2015](#); [Ortega-Tudela et al., 2019](#)), consisted of 25 ideas or propositions on the use of fibers and the manufacturing of fabrics. This is an expository text with clear organizational structures, educationally relevant content in the natural sciences, and corresponding to the first year of the Spanish Secondary Education (ESO) system (7th grade in the American system).

Subsequently, the students undertook two 10-min retrieval trials in the absence of the previously studied text. Critically, participants were randomly assigned to one of the two experimental conditions involving the same retrieval activities but in reverse order: Paragraph writing first and then concept map (PW + CM) or concept map first and then paragraph writing (CM + PW). This assignment took place upon arrival at the rooms designated for the study, regardless of participants' class affiliation. Both groups were similar in gender distribution and age (Group CM + PW: $n = 32$, 18 women, 3 unidentified; $M_{\text{age}} = 16.65$, $SD_{\text{age}} = 0.55$; Group PW + CM: $n = 28$, 12 women, 4 unidentified; $M_{\text{age}} = 17.00$, $SD_{\text{age}} = 0.78$).

Those participating in the PW + CM group were given a blank piece of paper and asked to write down everything they could recollect about the text (paragraph format). After 10 min the pieces of paper were collected and the students had a 1-min break. Then, they were provided with a new blank piece of paper and asked to elaborate a concept map in the way that they learned in the first session. In the CM + PW group, however, the participants first engaged in concept mapping and then they were moved onto the paragraph-based activity. Immediately after the 1-min break, and before taking the second retrieval trial, participants from both groups were given a sheet of paper and asked to make judgments of learning, a subjective estimate on a scale ranging from 0 to 100 of how much they would be able to recall from the text in a final test to be administered 2 weeks later.

The third experimental session took place 2 weeks later. In it, the students took a learning test that comprised 15 short-answer questions to be answered in writing in the space available. Ten of the questions required students to recall content directly presented in the text (verbatim). An example of a verbatim question is, "What is the origin of the natural fibers that are used to create fabrics?," with "animal and vegetables" being the correct answer that explicitly appeared in the text. The other five questions elicited inferences about the studied content (inference). An example of an inference question is, "Why are China and Egypt good countries for cotton cultivation?," with "Due to the hot and humid climate" as the correct answer. Different sentences from the text must be integrated in order to answer the question ("Thousands of years ago, the Chinese and the Egyptians knew the method of making fiber from the hairs of cotton seeds." and "Most cotton is grown in countries that have a hot, humid climate"). The students were given 15 min to answer the questions.

The responses to the learning test were scored by using a correction template containing the ideas of the text in addition to the acceptable responses for each question. This template has been widely used across studies and made the scoring process straightforward, highly reliable and free of ambiguity (see [Lechuga et al., 2015](#); [Ortega-Tudela et al., 2019](#)). A graduate Psychology student was briefly trained in the application of the template on the basis of examples of hits and errors from previous studies. This student, who was blind to all aspects of the study, scored the responses to verbatim and inferential questions from the learning test (up to 1 point per question). One of the authors (J.M.O.-T.) supervised the correct application of the template.

As in related studies ([Lechuga et al., 2015](#); [Ortega-Tudela et al., 2019](#)), two judges (graduate students who were also totally blind to the aims of the study) were asked to score the number of ideas from the text (1 point per unit) that were recalled during the retrieval trials in either format (paragraphs or concept maps). They also used the aforementioned template with the 25 ideas from the text to score the responses. Since this scoring procedure did not lead to significant disagreements between the two raters (see also [Lechuga et al., 2015](#); [Ortega-Tudela et al., 2019](#)), it was decided to proceed collaboratively after scoring a few participants. Thus, both judges scored each participants' response face-to-face, and a consensus on scoring was reached when disagreements arose (less than 1% of responses). Finally, these two judges rated the quality of the participants' concept maps on a 10-point scale regarding 5 dimensions: number of represented primary ideas, number of represented secondary ideas, hierarchical structure, use of connectors, and concept organization. The global intraclass correlation coefficient (ICC) between raters was 0.844. The ICCs for the dimensions were: primary ideas = 0.758, secondary ideas = 0.796, hierarchical structure = 0.759, use of connectors = 0.694, and concept organization = 0.753.

3 Results

Mediation analyses, independent samples *t*-tests, analyses of variance, and correlation/regression analyses were conducted on the data from the self-reported (use of learning strategies, academic self-concept, and judgments of learning; [Table 1](#)), and performance measures (ideas produced at retrieval practice and correct responses on the learning test; [Table 2](#); and quality of concept maps; [Table 3](#)). All

TABLE 1 Mean proportion (and standard deviations) of ideas produced in each trial of retrieval and correct responses in the final test as a function of retrieval sequence.

Retrieval sequence	Retrieval practice		Learning (final) test			
	Trial 1	Trial 2	Overall	Verbatim	Inferential	Overall
PW + CM	0.34 (0.17)	0.22 (0.10)	0.28 (0.13)	0.17 (0.12)	0.20 (0.21)	0.18 (0.13)
CM + PW	0.24 (0.09)	0.30 (0.18)	0.27 (0.13)	0.24 (0.17)	0.26 (0.20)	0.25 (0.16)
Overall	0.28 (0.14)	0.26 (0.15)		0.21 (0.15)	0.23 (0.21)	

CM, concept mapping; PW, paragraph writing. For the PW + CM group, Trial 1 refers to PW and Trial 2 refers to CM. For the CM + PW group, Trial 1 refers to CM and Trial 2 refers to PW.

TABLE 2 Mean scores (and standard deviations) of self-reported measures (learning strategies, as measured with CTHE; academic self-concept as measured with ASCA, and judgments of learning) as a function of retrieval sequence.

		Retrieval sequence		<i>p</i>	Cohen's <i>d</i>
		PW + CM	CM + PW		
Learning strategies (0–4)	Underlining	3.09 (1.04)	3.32 (1.01)	0.41	0.23
	Schemas creation	2.08 (1.04)	2.29 (1.13)	0.50	0.19
	Summarizing	2.79 (1.25)	2.90 (1.16)	0.73	0.09
	Concept mapping	1.46 (0.65)	1.42 (0.67)	0.83	−0.06
	Self-generating questions	3.17 (0.91)	3.23 (0.88)	0.81	0.07
Academic self-concept (0–20)	Self-regulation	13.04 (2.91)	12.29 (3.19)	0.37	−0.24
	Cognitive ability	14.21 (2.78)	13.68 (2.96)	0.50	−0.18
	Motivation	14.17 (3.73)	13.58 (2.96)	0.52	−0.18
	Creativity	11.65 (2.81)	11.45 (2.34)	0.80	−0.07
Learning judgment (0–100)		38.57 (20.85)	33.75 (19.96)	0.36	−0.24

CM, concept mapping; PW, paragraph writing.

TABLE 3 Mean scores (and standard deviations in a 0–10 scale) in the considered dimensions on the quality of the concept maps as a function of retrieval sequence.

Assessed dimensions	Retrieval sequence		<i>p</i>	Cohen's <i>d</i>
	PW + CM	CM + PW		
Primary ideas	6.15 (1.90)	6.22 (1.81)	0.88	0.04
Secondary ideas	5.63 (1.78)	5.48 (2.19)	0.80	−0.07
Hierarchical structure	5.85 (1.77)	5.83 (1.79)	0.96	−0.01
Use of connectors	4.71 (2.65)	5.78 (2.40)	0.59	0.15
Conceptual organization	6.04 (1.74)	6.09 (1.51)	0.91	0.03
Overall quality (mean score)	5.67 (1.79)	5.74 (1.71)	0.89	0.04

CM, concept mapping; PW, paragraph writing.

analyses were performed by using JASP software (JASP team, 2023, version 0.17.3). The full data set is available at <https://osf.io/82acz/>.

We report analyses of performance on retrieval activities (paragraph writing and concept mapping) at the group level and their potential association with performance on the learning test. Subsequently, we report analyses aimed at determining whether the sequence of retrieval-based activities had a direct effect on learning outcomes in the 2-week delayed test and whether this effect was

mediated by retrieval success during the first retrieval activity. Finally, we describe the results of analyses aimed at discovering: (a) whether students' judgments of learning were predictive of actual performance on the learning test; and (b) whether the use of learning strategies and the academic self-concept modulated performance either in the retrieval activities or in the learning test.

3.1 Success during retrieval practice

A mixed analysis of variance (ANOVA) on the proportion of ideas produced (out of the total of 25) during retrieval practice was performed with retrieval sequence (PW + CM vs. CM + PW) as the between-participants factor and retrieval period as the within-participant factor (trial 1 vs. trial 2). Neither sequence [$F(1, 58) < 1$, $p\eta^2 < 0.01$] nor retrieval trial reached statistical significance, [$F(1, 58) = 3.37$, $p = 0.07$, $p\eta^2 = 0.05$]. However, the interaction retrieval sequence $\times$ trial did, [$F(1, 58) = 38.08$, $p < 0.01$, $p\eta^2 = 0.40$]. Follow-up analyses revealed that during trial 1 the PW + CM group showed better performance ($M = 0.34$; $SD = 0.16$) than the CM + PW group [$M = 0.24$; $SD = 0.10$; $F(1, 58) = 8.38$, $p < 0.01$, $p\eta^2 = 0.13$]. During trial 2, it was the CM + PW group that produced more ideas ($M = 0.30$; $SD = 0.18$) than the PW + CM group [$M = 0.22$; $SD = 0.10$; $F(1, 58) = 4.75$, $p = 0.03$, $p\eta^2 = 0.08$]. In other words, writing paragraphs led participants to recall more ideas than concept mapping regardless of the order in which activities were done (see Table 1). No reliable difference between the groups emerged either when writing

paragraphs or when elaborating concept maps (both with $F < 1$ and $p\eta^2 < 0.01$).

We also examined whether performance during the retrieval practice periods predicted final learning considering the whole set of participants. First, we wanted to learn whether overall success during retrieval practice (collapsing performance across retrieval periods) correlated with learning outcomes (averaging performance in both types of test questions) 2 weeks later. This analysis revealed a reliable positive Pearson correlation ($r = 0.64$, $p < 0.001$). Then, we examined correlations for each type of retrieval activity. Both retrieval success in paragraph writing ($r = 0.62$, $p < 0.001$) and concept mapping ($r = 0.58$, $p < 0.001$) predicted learning outcomes. In addition, we conducted separate stepwise multiple regression analyses over both types of test questions to identify the best predictor in each case. Interestingly, the analyses revealed that success in paragraph writing periods uniquely predicted performance [31% of variance; $t(59) = 5.22$, $p < 0.01$] on verbatim questions, whereas success in concept mapping periods uniquely predicted performance [29% of variance, $t(59) = 4.95$, $p < 0.01$] on inferential questions.

3.2 Effect of sequence of retrieval activities on learning test performance

Since the main goal of the present study was to determine whether retrieval sequence had an effect on students' learning, and following the main analysis performed by Ortega-Tudela et al. (2019), we firstly conducted the originally-planned mixed ANOVA on correct responses in the learning test with retrieval sequence (between-groups) and type of question (within-participant) as factors. The analysis failed to show statistically significant effects [retrieval sequence: $F(1,58) = 2.55$, $p = 0.12$, $p\eta^2 = 0.04$; question type: $F(1,58) < 1$, $p\eta^2 = 0.01$; interaction: $F(1,58) < 1$, $p\eta^2 < 0.01$]. Importantly, however, the fact that in the present study with adolescents (a) the two retrieval activities had different retrieval success rates, with paragraph writing leading to the production of more ideas, especially in trial 1, which is an important deviation from the results previously found with college students, and that (b) success during retrieval practice was predictive of final learning, clearly pointed to the need of considering performance in the first trial of retrieval practice as a relevant factor when accounting for learning in the final test. Differences in success rates (stemming from distinct factors) during retrieval practice may modulate learning outcomes (e.g., Minear et al., 2018; Leggett et al., 2019; Racsmany et al., 2020; Ortega-Tudela et al., 2021; see also Karpicke, 2017). Specifically, in the present study elaborating retrieval-based concept maps first seemed to have very different effects on retrieval success (fewer ideas produced during retrieval practice) and performance on the delayed learning test (more correct responses), but both retrieval sequence (more benefit from concept mapping first) and retrieval success (more benefit from paragraph writing first) might have opposing effects on learning. If so, failing to account for such facing effects could make it difficult to observe differences between the two sequences in the final test. In other words, the expected direct effect of sequencing the retrieval activities may have been masked by the indirect larger effect of retrieval success during paragraph writing. To deal with this possibility, we conducted a mediation analysis to examine whether the lack of effect of retrieval sequence was mediated by the effect of success during retrieval

practice on performance on the learning test. If this were the case, a reliable direct effect of retrieval sequence should be found after accounting for the mediating (indirect) effect of retrieval practice success on learning. The mediation analysis was conducted by introducing retrieval sequence (CW + PW vs. PW + CW) as the predictor, success during retrieval practice as the mediator, and correct responses to verbatim and inferential questions as the outcomes. Statistical significance was tested by using 5,000 (bias-corrected percentile) bootstrap samples. The analysis confirmed direct and indirect effects. As can be seen in Figure 1, retrieval sequence had a direct effect on the proportion of correct responses to both verbatim (c'_1 : $B = -0.12$ [95% CI, -0.19 , -0.05], $p = 0.001$, $\beta = -0.77$) and inference questions (c'_2 : $B = -0.13$, 1.12 [95% CI, -0.23 , -0.03], $p = 0.01$, $\beta = -0.63$) that was independent of success during retrieval practice. Thus, CM + PW led to better learning scores than PW + CM. There was also a (smaller) indirect effect that was mediated by success during the first retrieval trial on verbatim (ab_1 : $B = 0.05$ [95% CI, 0.02, 0.10], $p = 0.02$, $\beta = 0.34$) and inference questions (ab_2 : $B = 0.07$ [95% CI, 0.02, 0.14], $p = 0.02$, $\beta = 0.34$). This mediation effect indicates that the benefit of paragraph writing for learning relies on its capability to facilitate retrieval, although this effect was smaller than the direct effect produced by concept mapping first. Confirming the differential influence of concept mapping and paragraph writing as first retrieval activities on learning, the direct and indirect effects had estimates of a different sign, which accounts for the lack of a statistically significant total effect (please, note that total effect = direct effect + indirect effect) of retrieval sequence on correct responses to verbatim (c : $B = -0.07$ [95% CI, -0.14 , 0.01], $p = 0.08$, $\beta = -0.44$) and inferential questions (c_2 : $B = -0.06$ [95% CI, -0.16 , 0.05], $p = 0.25$, $\beta = -0.29$).

3.3 Learning strategies and academic self-concept

Only the self-reported frequency of use of two learning strategies showed to be predictive of successful performance. On the one hand, greater frequency of use of underlining was shown to be inversely correlated with the proportion of ideas recalled when writing paragraphs ($r = -0.34$, $p = 0.01$). On the other hand, summarizing negatively correlated with the number of ideas recalled either when writing paragraphs ($r = -0.37$, $p < 0.01$) or creating concept maps ($r = -0.40$, $p < 0.01$), and with the overall score on the final learning test ($r = -0.36$, $p < 0.01$)².

2 Since underlining and summarizing are considered scarcely valuable learning techniques (although students may have the ability to implement them effectively, they do not always do so; Dunlosky et al., 2013), these associations might be simply reflecting individual differences in motivation or general abilities to engage in academic tasks (such as the ones being required in our study). Hence, those participants who reported more use of underlining and summarizing as study techniques could have devoted less attention while reading the text or when told to think back on its main ideas during the retrieval periods. Because there were no differences between the groups in the use of these techniques, this finding is not further discussed here.

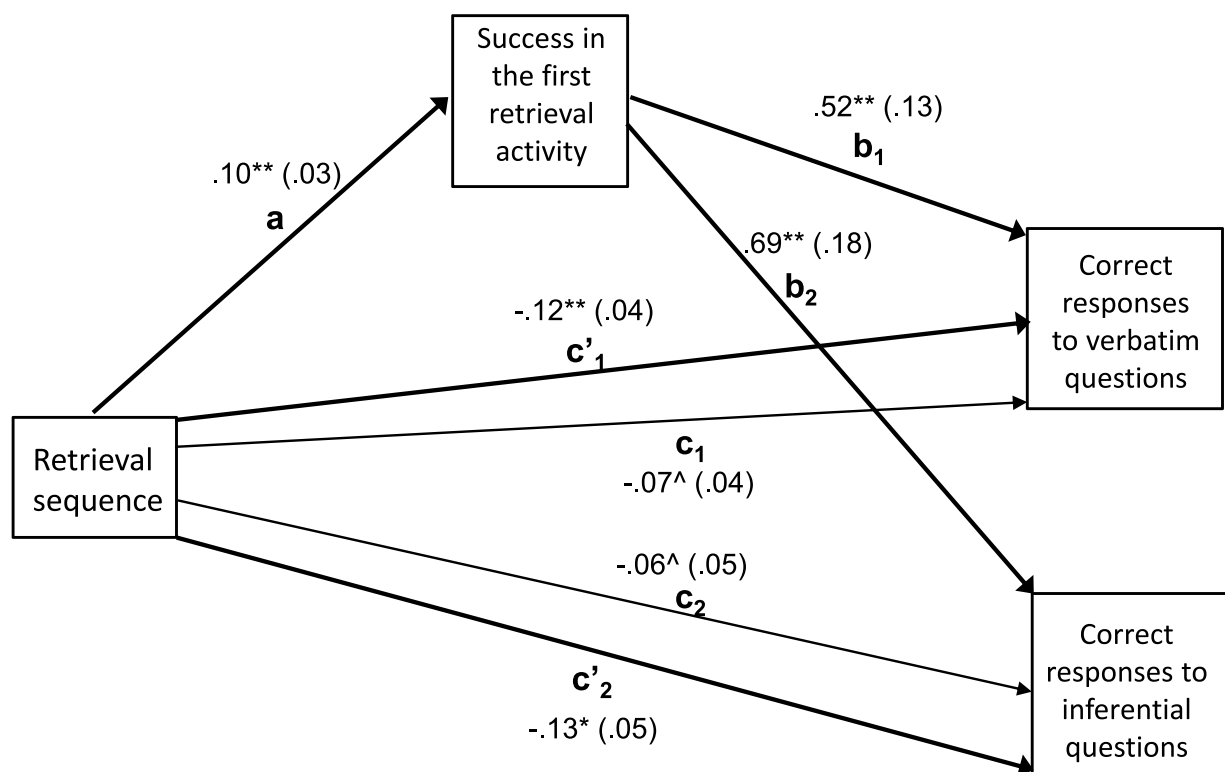


FIGURE 1

Path model diagram with the results of the mediation analysis. Total effect (paths c_1 and c_2) represents the effect of retrieval sequence on final test performance with no mediator in the model. Direct effect (paths c'_1 and c'_2) represents the effect of retrieval sequence on performance that is independent of retrieval success as a mediator in the model. Indirect effect (paths ab_1 and ab_2) represents the effect of retrieval on performance that is mediated by retrieval success. The figure shows unstandardized coefficients (and standard errors). See main text for further information. $^*p < 0.01$, $^{**}p < 0.01$, $^{\wedge}$ non-significant.

No associations emerged between any of the dimensions of the ASCA scale and student performance either in the retrieval activities or in the learning test (all Pearson coefficients with $p > 0.19$). For the sake of completeness, correlations between the self-reported measures are reported in [Supplementary material](#), although they will be not further discussed.

[Table 2](#) reports means and standard deviations of learning strategies and the academic self-concept scales for both groups of participants. Two-tailed t -tests for independent samples failed to show differences between the two groups in frequency of use of learning strategies. While concept mapping was reported as the least used study strategy, underlining and self-generation of questions were the strategies that students reported using the most. No group differences were also observed when it came to responding to items regarding the different dimensions included in the ASCA scale (see [Table 2](#)).

3.4 Judgment of learning

A bivariate Pearson correlation between the judgments of learning made by the students after the first retrieval activity and their overall performance on the (2-weeks delayed) learning test showed no reliable association ($r = 0.22$, $p = 0.08$; partial correlation after controlling for overall success at retrieval practice: $r = 0.07$, $p = 0.60$). A two-tailed

t -test failed to show a difference in the judgments of learning made by the groups (see [Table 2](#)).

3.5 Quality of the concept maps

Two-tailed t -tests for independent samples showed that there were no differences between both groups in any of the dimensions considered when assessing the quality of the concept maps (see [Table 3](#)). Finally, we also examined whether the quality of concept maps (overall score) was predictive of performance in the final test. Pearson correlation analyses showed positive and reliable associations between overall quality and proportion of correct responses to verbatim ($r = 0.37$, $p < 0.01$) and inferential questions ($r = 0.32$, $p = 0.02$).

4 Discussion

With the aim of further investigating the potential benefit of elaborating a concept map (in the absence of texts) as an initial learning activity within a series, in the present study, we assessed the effect of the sequence of such retrieval-based activities in a sample of high school students. While there is an increasing interest in determining the effectiveness of distinct combinations of learning

activities (with an emphasis on those relying on retrieval practice to boost outcomes), to the best of our knowledge no previous study has examined the effect of such combinations on learners other than college students.

Extending previous findings among college students, our results with high-school students support the idea that conceptual learning from educational texts can be modulated by the sequence of retrieval-based activities performed. Specifically, these results reveal that creating a concept map in the absence of educational texts followed by free recall by writing paragraphs led to better responses to verbatim and inferential questions (on a learning test that took place 2 weeks later) than performing the same activities in the reverse order. Additionally, our results indicate that this effect is not related to prior experience with the use of study techniques or academic self-concept. Hence, that specific benefit, originally observed in college students in a lab setting (Ortega-Tudela et al., 2019), has been shown to be generalizable in a more naturalistic educational setting to learners who are less experienced as students than college adults and who were, in fact, less successful at retrieving relevant ideas from the text and less skilled at elaborating concept maps³. A few aspects of these results merit particular consideration and will be discussed below.

Aside from the age of the participants, a remarkable difference between the present study and the one with college learners (Ortega-Tudela et al., 2019) is that the set of research-related tasks (concept mapping training, retrieval practice, and learning assessment) were all implemented in the classroom as part of a series of academic activities oriented to enhance the students' skills in preparing for exams. Hence, all these activities were carried out in a genuine educational setting⁴. Despite these differences with the original study, our results largely replicate the finding of better learning outcomes resulting from the sequence that involves concept mapping first, which suggests that such a sequencing effect might work across a variety of learners and contexts. Although the latter is an empirical question that would be worth addressing in future studies, the present results contribute to this idea.

It has turned out to be of special practical significance that we observed the benefit of concept mapping first with learners who reported that they hardly ever created concept maps to deal with exams (concept mapping was reported as the least used learning strategy by participants) and who, as a matter of fact, elaborated concept maps of rather low-moderate quality during their retrieval

periods in our study. Thus, from our own findings with adolescents it would seem that the value of retrieval-based concept mapping does not depend on expertise in concept mapping. However, we acknowledge that as concept mapping is a developmental skill that improves with repeated practice and feedback (Roessger et al., 2018), the potential benefits of retrieval-based concept mapping as a learning activity may increase with training and practice. This is also an important question to address in future studies. Studies that have compared concept mapping with other learning strategies (e.g., retrieval practice) have typically introduced concept map creation as a one-off activity (Karpicke and Blunt, 2011; O'Day and Karpicke, 2021) which means that learners did not necessarily master the required skill, which could have minimized the potential benefits of concept mapping (see Lechuga et al., 2015).

Instead, and also of practical relevance, the benefit of retrieval-based concept mapping would seem to depend on the nature of the retrieval activity that follows in the sequence of learning tasks. As Blunt and Karpicke (2014, Exp. 2; see also Exp. 1 from Ortega-Tudela et al., 2019) showed, two consecutive periods of concept mapping in the absence of materials (i.e., CM + CM) failed to improve learning as compared to either open-book concept mapping or repeated free recall by writing paragraphs (i.e., PW + PW). Therefore, it seems to be the specific configuration of concept mapping plus free recall by writing (in this order) that it has been shown to be particularly effective in improving learning outcomes. It remains to be explored whether other learning activities (retrieval-based or not) beyond free recall might also make a difference when it comes to learning from educational materials. In addition, it would also be necessary to show that CM + PW improves learning more than PW + PW, given the difficulty of constructing concept maps with closed books (especially for adolescents and children) and the apparent ease of writing paragraphs. While an indirect (cross-experiment) comparison between both sequences is possible from the data of Ortega-Tudela et al. (2019) with younger adults, which reveals better learning after CM + PW than after PW + PW, future studies should examine this issue more closely. In any case, the present results, together with previous findings, argue against the general idea that all retrieval-based activities are similarly effective simply because they involve the fact of recalling (Blunt and Karpicke, 2014; O'Day and Karpicke, 2021). Rather, different retrieval demands may interact with each other and with other cognitive processes underpinning the specific learning task at hand to modulate knowledge acquisition and long-term retention. Creating a concept map without materials might promote a more conceptually-guided retrieval orientation and the generation of more elaborated mental models that, in turn, would facilitate the adoption of more effective retrieval strategies at subsequent attempts to access relevant knowledge (Ortega-Tudela et al., 2019). This interpretation of the benefit of CM + PW over PW + CM as retrieval activities would align with the elaborative retrieval account (Carpenter, 2009, 2011; see also Hinze et al., 2013). According to this view, when retrieval attempts are demanding and the learners have to get engaged in reconstructive process reliant on additional information (e.g., during concept mapping in the absence of texts), these additional pieces of information are then integrated with existing memory traces (Carpenter, 2011). In other words, elaborated memories may be created during demanding retrieval episodes that might help to re-organize or supplement initially encoded information, thus facilitating future access to these memories. Nevertheless, we would like to emphasize that these findings do not

³ Cross-experiment two-tailed *t*-tests confirmed that the (adult) participants in Ortega-Tudela et al.'s study (Exp. 2): (a) recalled more ideas in the first retrieval activity (CM+PW: $M=0.45$; $SD=0.09$; PW+CM: $M=0.42$; $SD=0.09$) than the (adolescent) participants in the present study (CM+PW: $M=0.24$; $SD=0.09$; PW+CM: $M=0.34$; $SD=0.17$), both tests with $p<0.01$, and (b) elaborated concept maps of overall better quality ($M=6.60$; $SD=1.45$) than the younger participants ($M=5.71$; $SD=1.73$), $p<0.001$.

⁴ Closely following the definition of classroom research considered by Agarwal et al. (2021) in their recent meta-analysis on retrieval practice in educational settings, in the present study the materials and activities to be performed on them were directly related to the assigned course materials. In addition, retrieval practice was individual, not collaborative, so that all participants engaged in either concept mapping or free recall individually under the supervision of instructors. Finally, in all cases retrieval-based activities were performed without the use of external learning aids.

necessarily indicate a general superiority of performing elaborative/generative learning tasks first within a sequence. As a matter of fact, a notable difference between the present study and previous research with young adults is that high school students recalled fewer ideas when using concept mapping as a retrieval task compared to paragraph writing. This suggests that concept mapping with closed books may be a challenging retrieval task for younger students. It is worth noting that adults showed similar retrieval success across retrieval formats in [Ortega-Tudela et al. \(2019\)](#). Furthermore, in the present study, success in the first retrieval activity played a role as a mediator in the relationship between retrieval sequence and learning scores. This implies that, in addition to the inherent benefit of creating concept maps first, which was evident from the reliable direct effect of the retrieval sequence favoring CM + PW over PW + CM, the rate of success during retrieval practice may modulate its effectiveness. While this fact does not come as a surprise (e.g., [Leggett et al., 2019](#)), from a more practical perspective it points to the need to consider the utilization of procedures (e.g., providing hints during concept mapping) that ensure acceptable success rates for students facing challenging retrieval tasks. Additionally, there is evidence that having learners engaged in retrieval practice (without elaborative components) before engaging in a generative learning activity can also lead to better learning outcomes ([Roelle et al., 2022](#)). Hence, the main question to be answered is not which specific learning sequence is most effective, but under which conditions a specific sequence can deliver superior learning outcomes. Further research is clearly necessary on this issue.

Finally, an unexpected, but not surprising, finding in the present study was a retrieval activity-based dissociation when it came to responding to factual and inference questions 2 weeks after doing retrieval practice. While it was performance during the periods of paragraph writing that better predicted the proportion of correct responses to verbatim questions on the final test, retrieval success during concept mapping was a better predictor of correct responses to inference questions (thought to demand more elaborative processes) on the learning test. Indeed, this finding is suggestive of the elaborative/relational nature of the cognitive processing that underlies concept mapping, with particular consequences for learning, and is consistent with results from recent studies that compared the effects of generative learning and retrieval practice (e.g., [Roelle and Nückles, 2019](#)). The benefit of generative/elaborative tasks (in the present case, concept mapping) is thought to be reliant on the quantity or quality of mental processes that make sense of the information provided, and one would expect such a benefit to be particularly evident in inference or comprehension questions ([Fiorella, 2023](#)). Nevertheless, our participants did not appear to be aware of the potential impact of retrieval-based concept mapping as an initial learning activity, as both groups made similar judgments of learning but had different learning outcomes.

5 Limitations and recommendations

A limitation of our results, even when converging with previous findings concerning students of a different nature, is the reliance on the restricted educational materials worked on by the students. It would be of practical relevance to determine in future studies the

effectiveness of initial retrieval-based concept mapping with more varied and curriculum-related materials. A second limitation is that the present study included a relatively small sample of students, so that the present results should be interpreted with caution even when they are in line with results from previous findings. Naturalistic studies usually come at the expense of reduced number of participants even when ideal sample size was estimated in advance. In addition, we recognize that our participants' prior knowledge of the subject matter of the learning text was not assessed. While the random assignment of students to experimental conditions is expected to control for prior knowledge, differences between groups on this variable may still be possible. A fourth limitation concerns the quality of the concept maps. Although we assessed a number of relevant dimensions, we did not consider the number of connections that a given concept had with other concepts, which has been shown to be a relevant parameter mediating immediate and delayed recall ([Roessger et al., 2021](#)). Finally, our study only focused on the students. However, given the role that teachers may play when designing and implementing retrieval-based activities in the classroom ([Ortega-Tudela et al., 2021](#)), considering their expectations about the effectiveness of these activities as well as their feelings on the benefit–cost ratio after implementing them would contribute to our understanding of the practical value of such activities.

6 Conclusion

Our study with high school students provides evidence, consistent with the results from some previous studies, for the importance of learning tasks that promote the construction of coherent mental representations of to-be-learned materials ([Roelle et al., 2023](#)). Specifically, the present findings lend support to the benefit of concept mapping as a retrieval-based learning activity and should help instructors and students to think of tests as knowledge modifiers and learning rather than just as assessment tools (for recent approaches see [Bae et al., 2019](#); [Roelle et al., 2022](#)). Specifically, the present results contribute to the identification of an effective combination of retrieval practice activities that may be implemented among high school students with little experience of concept maps. Future studies should identify potential factors that may moderate the efficacy of such a combination.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://osf.io/82acz/>.

Ethics statement

The study was approved by Comité de Bioética de la Universidad de Jaén and conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

ML: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. JO-T: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Writing – review & editing. CG-A: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing.

Funding

The authors declare that financial support was received for the research and publication of this article from Consejería de Economía, Conocimiento, Empresas y Universidades from Junta de Andalucía, Spain [Programa Operativo FEDER Andalucía 2014–2020. Proyectos_FEDER_1262768].

Acknowledgments

We thank the staff and children from the IES Fuente de la Peña (Jaén, Spain) for their implication and help.

References

- Adesope, O. O., Trevisan, D. A., and Sundararajan, N. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Rev. Educ. Res.* 87, 659–701. doi: 10.3102/0034654316689306
- Agarwal, P. K., Nunes, L. D., and Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: a systematic review of applied research in schools and classrooms. *Educ. Psychol. Rev.* 33, 1409–1453. doi: 10.1007/s10648-021-09595-9
- Álvarez, M., and Fernández, R. (1999). *Cuestionario de Hábitos y Técnicas de Estudio (CHTE)*. Madrid: TEA.
- Bae, C. L., Theriault, D. J., and Redifer, J. L. (2019). Investigating the testing effect: retrieval as a characteristic of effective study strategies. *Learn. Instr.* 60, 206–214. doi: 10.1016/j.learninstruc.2017.12.008
- Blunt, J. R., and Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *J. Educ. Psychol.* 106, 849–858. doi: 10.1037/a0035934
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J. Exp. Psychol. Learn.* 35, 1563–1569. doi: 10.1037/a0017021
- Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: support for the mediator effectiveness hypothesis of the testing effect. *J. Exp. Psychol. Learn.* 37, 1547–1552. doi: 10.1037/a0024140
- Chevron, M. P. (2014). A metacognitive tool: theoretical and operational analysis of skills exercised in structured concept maps. *Persp. Sci.* 2, 46–54. doi: 10.1016/j.pisc.2014.07.001
- Craik, F. I., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* 104, 268–294. doi: 10.1037/0096-3445.104.3.268
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Publ. Int.* 14, 4–58. doi: 10.1177/1529100612453266
- 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. doi: 10.3758/BF03193146
- Fiorella, L. (2023). Making sense of generative learning. *Educ. Psychol. Rev.* 35:50. doi: 10.1007/s10648-023-09769-7
- Fiorella, L., and Mayer, R. E. (2016). Eight ways to promote generative learning. *Educ. Psychol. Rev.* 28, 717–741. doi: 10.1007/s10648-015-9348-9
- Fritz, C. O., Morris, P. E., Acton, M., Voelkel, A. R., and Etkind, R. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Appl. Cognit. Psych.* 21, 499–526. doi: 10.1002/acp.1287
- Gallenstein, N. (2013). Concept mapping for learners of all ages. *J. Educ. Teach. Train.* 4, 59–72.
- Habók, A. (2012). Evaluating a concept mapping training programme by 10- and 13-year-old students. *Int. Electr. J. Element. Ed.* 4, 459–472.
- Hinze, S. R., Wiley, J., and Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *J. Mem. Lang.* 69, 151–164. doi: 10.1016/j.jml.2013.03.002
- Hunt, R. R. (2012). Distinctive processing: the co-action of similarity and difference in memory. *Psychol. Learn. Mot.* 56, 1–46. doi: 10.1016/B978-0-12-394393-4.00001-7
- Kahana, M. J. (2017). *Memory search, in learning and memory: a comprehensive reference*, Oxford, England: Academic Press, 2, 181–200
- Karpicke, J. D. (2017). "Retrieval-based learning: A decade of progress, in cognitive psychology of memory" in *Learning and memory: a comprehensive reference*. ed. J. Wixted (Oxford: Elsevier)
- Karpicke, J. D. (2018). "Concept mapping" in *The SAGE encyclopaedia of educational research, measurement and evaluation*. ed. B. Frey (Thousand Oaks, CA: SAGE Publishing, Inc.), 351–354.
- Karpicke, J. D., and Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 331, 772–775. doi: 10.1126/science.1199327
- Karpicke, J. D., and Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *J. Mem. Lang.* 67, 17–29. doi: 10.1016/j.jml.2012.02.004
- Kinchin, I. M. (2014). Concept mapping as a learning tool in higher education: a critical analysis of recent reviews. *J. Cont. High. Ed.* 62, 39–49. doi: 10.1080/07377363.2014.872011
- Lamotte, M., Izaute, M., and Darnon, C. (2021). Can tests improve learning in real university classrooms? *J. Cogn. Psychol.* 33, 974–992. doi: 10.1080/20445911.2021.1956939
- Lechuga, M. T., Ortega-Tudela, J. M., and Gómez-Ariza, C. J. (2015). Further evidence that concept mapping is not better than repeated retrieval as a tool for learning from texts. *Learn. Instr.* 40, 61–68. doi: 10.1016/j.learninstruc.2015.08.002
- Leggett, J. M. I., Burt, J. S., and Carroll, A. (2019). Retrieval practice can improve classroom review despite low practice test performance. *Appl. Cognit. Psychol.* 33, 759–770. doi: 10.1002/acp.3517
- McDaniel, M. A. (2023). Combining retrieval practice with elaborative encoding: complementary or redundant? *Educ. Psychol. Rev.* 35:75. doi: 10.1007/s10648-023-09784-8
- McDaniel, M. A., and Pressley, M. (1984). Putting the keyword method in context. *J. Educ. Psychol.* 76, 598–609. doi: 10.1037/0022-0663.76.4.598

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2024.1287744/full#supplementary-material>

- McInerney, D. M., Cheng, R. W., Mok, M. M. C., and Lam, A. K. H. (2012). Academic self-concept and learning strategies: direction of effect on student academic achievement. *J. Adv. Acad.* 23, 249–269. doi: 10.1177/1932202X12451020
- Mechie, I. R., Plaisted-Grant, K., and Cheke, L. G. (2021). How does episodic memory develop in adolescence? *Learn. Mem.* 28, 204–217. doi: 10.1101/lm.053264.120
- Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., and Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *J. Exp. Psychol. Learn.* 44, 1474–1486. doi: 10.1037/xlm0000486
- Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., et al. (2011). Comment on retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 334:453. doi: 10.1126/science.1203698
- Miyatsu, T., Nguyen, K., and McDaniel, M. A. (2018). Five popular study strategies: their optimal implementation and pitfalls. *Perspect. Psychol. Sci.* 13, 390–407. doi: 10.1177/1745691617710510
- Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., and Jaeger, A. (2019). Retrieval practice in classroom settings: a review of applied research. *Front. Educ.* 4:5. doi: 10.3389/feduc.2019.00005
- Nesbit, J. C., and Adesope, O. O. (2006). Learning with concept and knowledge maps: a meta-analysis. *Rev. Educ. Res.* 76, 413–448. doi: 10.3102/00346543076003413
- Novak, J. D., and Cañas, A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Inform. Visual.* 5, 175–184. doi: 10.1057/palgrave.ivs.9500126
- Novak, J. D., and Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press
- Nunes, L. D., and Karpicke, J. D. (2015). “Retrieval-based learning: research at the interface between cognitive science and education” in *Emerging trends in the social and behavioral sciences*. eds. R. A. Scott and S. M. Kosslyn (Chichester: Wiley), 1–16.
- O'Day, G. M., and Karpicke, J. D. (2021). Comparing and combining retrieval practice and concept mapping. *J. Educ. Psychol.* 113, 986–997. doi: 10.1037/edu0000486
- Ordaz-Villegas, G., Acle-Tomasini, G., and Reyes-Lagunes, L. I. (2013). Development of an academic self-concept for adolescents (ASCA) scale. *J. Beh. Health Soc. Iss.* 5, 117–130. doi: 10.5460/jbhsi.v5.2.42304
- Ortega-Tudela, J. M., Lechuga, M. T., Bermúdez-Sierra, M., and Gómez-Ariza, C. J. (2021). Testing the effectiveness of retrieval-based learning in naturalistic school settings. *SAGE Open* 11:215824402110615. doi: 10.1177/21582440211061569
- Ortega-Tudela, J. M., Lechuga, M. T., and Gómez-Ariza, C. J. (2019). A specific benefit of retrieval-based concept mapping to enhance learning from texts. *Instr. Sci.* 47, 239–255. doi: 10.1007/s11251-018-9476-y
- Quinn, H., Mintzes, J., and Laws, R. A. (2003). Successive concept mapping: assessing understanding in college science classes. *J. Coll. Sci. Teach.* 33, 12–17.
- Racsmány, M., Szöllősi, Á., and Marián, M. (2020). Reversing the testing effect by feedback is a matter of performance criterion at practice. *Mem. Cogn.* 48, 1161–1170. doi: 10.3758/s13421-020-01041-5
- Repetto, E., Sutil, I., and Manzano, N. (2004). *Comprender y aprender en el aula. Programa para la integración en el curriculum de las estrategias de comprensión lectora*. Madrid: UNED.
- Rivers, M. L. (2021). Metacognition about practice testing: a review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educ. Psychol. Rev.* 33, 823–862. doi: 10.1007/s10648-020-09578-2
- Roediger, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x
- Roelle, J., Endres, T., Abel, R., Obergassel, N., Nückles, M., and Renkl, A. (2023). Happy together? On the relationship between research on retrieval practice and generative learning using the case of follow-up learning tasks. *Educ. Psychol. Rev.* 35:102. doi: 10.1007/s10648-023-09810-9
- Roelle, J., Froese, L., Krebs, R., Obergassel, N., and Waldeyer, J. (2022). Sequence matters! Retrieval practice before generative learning is more effective than the reverse order. *Learn. Instr.* 80:101634. doi: 10.1016/j.learninstruc.2022.101634
- Roelle, J., and Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: the cohesion and elaboration of the text matters. *J. Educ. Psychol.* 111, 1341–1361. doi: 10.1037/edu0000345
- Roessger, K. M., Daley, B. J., and Hafez, D. A. (2018). Effects of teaching concept mapping using practice, feedback, and relational framing. *Learn. Instr.* 54, 11–21. doi: 10.1016/j.learninstruc.2018.01.011
- Roessger, K. M., Moy, G. R., Phan, T. V., Campbell, H., and Parker, D. (2021). Multi-level factors affecting immediate and delayed concept recall in concept mapping with adult learners. *Educ. Psychol.* 41, 224–244. doi: 10.1080/01443410.2020.1806211
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., and Adesope, O. O. (2018). Studying and constructing concept maps: a meta-analysis. *Educ. Psychol. Rev.* 30, 431–455. doi: 10.1007/s10648-017-9403-9
- Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S., and Lindenberger, U. (2010). Episodic memory across the lifespan: the contributions of associative and strategic components. *Neurosci. Biobehav. R.* 34, 1080–1091. doi: 10.1016/j.neubiorev.2009.11.002
- Shing, Y. L., Werkle-Bergner, M., Li, S.-C., and Lindenberger, U. (2008). Associative and strategic components of episodic memory: a life-span dissociation. *J. Exp. Psychol. Gen.* 137, 495–513. doi: 10.1037/0096-3445.137.3.495
- Tullis, J. G., and Maddox, G. B. (2020). Self-reported use of retrieval practice varies across age and domain. *Metacogn. Learn.* 15, 129–154. doi: 10.1007/s11409-020-09223-x
- Yang, C., Luo, L., Vadillo, M. A., Yu, R., and Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: a systematic and meta-analytic review. *Psychol. Bull.* 147, 399–435. doi: 10.1037/bul0000309



OPEN ACCESS

EDITED BY

Luisa Losada-Puente,
University of A Coruña, Spain

REVIEWED BY

Brett Bligh,
Lancaster University, United Kingdom
Agnieszka Szplit,
Jan Kochanowski University, Poland

*CORRESPONDENCE

Davide Parmigiani
✉ davide.parmigiani@unige.it

RECEIVED 05 January 2024

ACCEPTED 14 March 2024

PUBLISHED 22 March 2024

CITATION

Parmigiani D, Nicchia E, Murgia E and
Ingersoll M (2024) Formative assessment in
higher education: an exploratory study within
programs for professionals in education.
Front. Educ. 9:1366215.
doi: 10.3389/feduc.2024.1366215

COPYRIGHT

© 2024 Parmigiani, Nicchia, Murgia and
Ingersoll. 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.

Formative assessment in higher education: an exploratory study within programs for professionals in education

Davide Parmigiani^{1*}, Elisabetta Nicchia¹, Emiliana Murgia¹ and
Marcea Ingersoll²

¹University of Genoa, Genoa, Italy, ²St. Thomas University, Fredericton, NB, Canada

This study explores how prospective professionals in higher education can learn about and apply formative assessment methods relevant to their future educational workplaces. In the academic year 2022–23, 156 pre-service teachers, social workers, and heads of social services took part in a three-stage mixed-method study on university learning experiences involving formative assessment practices. They were exposed to self-, peer-, and group-assessment strategies. Data collected after each stage revealed participants' perspectives on each method. Findings show that students who engaged in formative assessment comprehended assessment complexity and were motivated to use diverse assessment forms. Formative assessment proves effective for both evaluation and development, supporting higher education students in honing assessment competencies for future professional roles in educational and social sectors.

KEYWORDS

formative assessment, professionals in education, higher education, teacher education, social workers

1 Introduction and rationale

The growing emphasis on formative assessment represents a challenge for universities, schools, and educational institutions within traditionally summative assessment cultures. In 2005, the Organization for Economic Cooperation and Development (OECD) declared that schools could support academic development through formative assessment, mainly for underachieving students (OECD, 2005). In addition, the OECD (2005) underscored that formative assessment improves the retention of learning and the quality of students' work. Additionally, the United Nations Educational, Scientific and Cultural Organization (UNESCO) emphasized how formative assessment can help students' learning during emergencies and crises (Bawane and Sharma, 2020). A recent study commissioned by the European Commission underscored that "formative assessment needs to place the learners themselves at the center of the learning and assessment processes, taking a more active and central role both as individual, self-regulated learners, and as critical peers" (Cefai et al., 2021, p. 7). These studies indicate the potential of formative assessment strategies to create and build meaningful learning environments in a variety of educational contexts.

In Europe, there has been discussion of the correspondence between the 1999 implementation of the Bologna Process and the increase in learner-centered assessment

approaches in educational systems that typically employ top-down, test-based methods (Pereira et al., 2015). Founded on a skills-based learning model, the European Higher Education Area (EHEA) currently guides the operations of European universities, prompting the adoption of new methodologies and assessment systems. Systematic reviews of formative assessment reveal that despite evidence that learner-centred formative assessment practices are effective, such practices have been researched predominantly in school-based contexts, unevenly implemented in European higher education, and dominated by studies from the United Kingdom (Pereira et al., 2015; Morris et al., 2021). Moreover, a recent study of Italian higher education assessment practices underscores a culture of assessment that places emphasis on the summative or concluding phase of teaching rather than developmental supports for ongoing learning through formative assessment (Doria et al., 2023). In this study, we answer calls for broader empirical work on formative assessment in higher education; specifically, we focus our attention on the implementation of embedded formative assessment within programs that prepare university students to work as pre-service teachers, social workers and heads of social services in Italy.

This study addresses a gap in the research on how formative assessment can be used to develop and evaluate students who are preparing for professional contexts where formative assessment knowledge is both a beneficial and expected skill. This goal is achieved through the following specific research questions: can formative assessment help students to achieve deep learning of several assessment methods and strategies, support students to understand the complexity of assessment procedures, promote students' growth from both a personal and professional point of view, and motivate them to use multiple forms of assessment when they become professionals in their respective fields?

The next section provides an overview of how conceptual understandings from the literature on formative assessment provide a rationale for the research and inform the study design. Drawing on existing literature on formative assessment in higher education contexts, we explore how such concepts and practices might be used to explore how to foster pre-service teachers, social workers, and heads of social services understanding of formative assessment. Thereafter, the study aims are explored through qualitative and quantitative data that support and sustain reasonable and realistic ways to introduce formative assessment strategies and activities in higher education contexts (Crossouard and Pryor, 2012). Finally, the limitations and difficulties of combining and effectively balancing summative and formative assessment methods (Lui and Andrade, 2022) are discussed, and recommendations for future practice are made.

2 Theoretical framework and literature review

This research derives from a theoretical understanding of learning as an ongoing and developmental process that is effective when the learner is engaged in and metacognitive about the assessment process (Bruner, 1970). Like Crossouard and Pryor (2012), we resist the separation of theory from practice and consider them entangled. This framework underpins both the concepts explored in the next sections

of this paper, and how these concepts have influenced the design, analysis and reporting of the study.

When exploring the potential of formative assessment in a higher education context that is largely summative, we considered three main areas of literature to support and situate the work: formative assessment as a developmental process in education, the potential of formative assessment within programs for professionals in education, and the principles of formative assessment that could inform the study design. Throughout this section we demonstrate how, in alignment with the learner-centred nature of formative assessment, this study aims to create and investigate the conditions for building higher education learning environments where students can experience several assessment types and can reflect on and improve their own learning processes and competence development (Dann, 2014; Ibarra-Sáiz et al., 2020). We also establish how the student profile that inspires us to conduct this study is represented by learners who are ultimately able to monitor, self-reflect, modify their learning strategies according to different situations, and seek multiple creative solutions suitable for their future professional tasks (Tillema, 2010; Evans et al., 2013; Ozan and Kincal, 2018).

2.1 Formative assessment as a developmental process for higher education

Educational theorist Jerome Bruner (1970) stated that “learning depends on knowledge of results, at a time when, and at a place where, the knowledge can be used for correction” (p. 120), indicating that learning processes are not only aimed at remembering content and information but also involve modification and understanding the best ways to learn. Higgins et al. (2010) defined formative assessment as a task to be carried out while students are learning so that they can have several forms of feedback aimed at improving their own learning, whether marked or not. From this basis, and like Petty (2004), we assert that the main goal of formative assessment is developmental. Essentially, formative assessment is intended to assist students in diagnosing and monitoring their own progress by identifying their strengths and weaknesses and spending their efforts on trying to improve their learning processes (Petty, 2004, p. 463). To summarize, the key concept of formative assessment is represented by the fact that the reflexivity of students is not static but it can be enhanced or further developed at any time (Hadrill, 1995).

To be effective in higher education contexts, formative assessment requires certain characteristics. First of all, formative assessment should be continuous (Brown, 1999). An episodic and occasional activity with formative assessment cannot allow students a meaningful opportunity to self-reflect on their own learning. Instead, “regular formative assessment can be motivational, as continuous feedback is integral to the learning experience, stimulating and challenging students” (Leach et al., 1998, p. 204). The effectiveness of formative assessment in higher education contexts is well described by Yorke (2003) when the author specifies that, through formative assessment practices, students have the opportunity to understand the meaning of formative comments made either by the teachers or their peers. Additionally, students can further modify their learning approaches based on their own understandings.

2.2 Formative assessment within programs for professionals working in educational contexts

A key question regarding the use of formative assessment within programs for professionals in educational contexts relates to how understanding informs future practice. If students have the opportunity to experience formative assessment strategies while studying at university, will they be more likely to practice formative assessment when they are teachers or social workers? This inquiry was made by Hamodi et al. (2017), who identified a need for greater understanding in this under-researched area. They found that formative assessment within programs for professionals in educational contexts represents a three-fold opportunity. Firstly, formative assessment can be used as a strategy to improve all students' learning processes during their university programs. Secondly, students in initial teacher education programs should learn and practice several formative assessment strategies because they will have to create many educational opportunities in their future practice. Lastly, students should be encouraged to use formative assessment when they become professionals in their respective working fields, inside and outside school. While the first opportunity is beneficial for all university students, the last two are particularly evident for the educational professionals who are the participants of our study: pre-service teachers, social workers, and heads of social services.

Formative assessment has cognitive and metacognitive benefits for all university students, and for prospective teachers, social workers and heads of social services, there are additional professional benefits, including the ability to implement formative assessment in educational contexts (Kealey, 2010; Montgomery et al., 2023). Specifically, formative assessment represents a vital strategy for professionals working in educational settings because it provides several feedback opportunities to develop self-regulated learning (Clark, 2012; Xiao and Yang, 2019).

Feedback and self-regulation are two important dimensions of formative assessment. Regarding the first dimension, feedback, Jensen et al. (2023) emphasized that feedback should be directed toward the development of instrumental and substantive learning goals. Instrumental feedback relates to whether the work has accomplished the task criteria, while substantive learning feedback, which is less common, directs students to "reflect critically on their own assumptions and leads to a new level of understanding or quality of performance" (p. 7). Furthermore, feedback should be based on comments that stimulate reflection (Dekker et al., 2013) with the following characteristics. According to Suhoyo et al. (2017), the comments should be: strength (what students did well); weakness (aspects which need improvement); comparison to standard (similarities and differences with the requested task); correct performance (whether students' performance approximates optimal performance); action plan (indicate future improvements). Ultimately, feedback should be timely, specific, actionable, respectful, and non-judgmental (Watson and Kenny, 2014).

Regarding self-regulation, formative assessment is aimed at setting up activities to allow students to react to feedback so as to enrich and increase the final outcomes (Ng, 2016). In other words, Pintrich and Zusho (2002) suggested that self-regulation is based on an active learning process in which learners are able to monitor and regulate their own cognition (p. 64). To do so, Nicol and Macfarlane-Dick (2006) indicated seven main principles for feedback that facilitates

positive self-regulation: clarifying what good performance is (goals, criteria, expected standards); facilitating the development of self-assessment (reflection) in learning; delivering high quality information to students about their learning; encouraging teacher and peer dialogue around learning; encouraging positive motivational beliefs and self-esteem; providing opportunities to close the gap between current and desired performance; providing information to teachers that can be used to help shape teaching (p. 205).

While summative assessment is necessary to indicate the level of student performance, it remains an extrinsic assessment that does not foster students' deep reflection on their own learning progressions (Ismail et al., 2022). Formative assessment represents an opportunity for prospective workers in education to deeply reflect on both their personal and professional development through cognitive and metacognitive benefits and feedback that facilitates positive self-regulation.

2.3 Formative assessment strategies: self-, peer-, and group-assessment

From the literature, we identified three main formative assessments strategies: self-, peer-, and group-assessment to be incorporated into our study. Panadero et al. (2015) defined self-assessment as an activity through which students can explain and underline the "qualities of their own learning processes and products" (p. 804). As specified by Andrade (2019), it is always important to clarify the purpose of formative assessment: in this study, self-assessment was not aimed at giving a grade. The activity was intended to support students' reflection on their learning processes, to improve student learning and, in particular, develop students' capacities for giving feedback to themselves or others, as described by Wanner and Palmer (2018).

Regarding peer-assessment, Biesma et al. (2019) specified that peer activities can allow students to analyze the quality of a task completed by other students highlighting the quality of their learning processes and their professional development. Similarly, van Gennip et al. (2010) defined peer assessment as a learning intervention, consequently, peer assessment can be considered as a supplemental strategy in the development of assessment as learning. Moreover, Yin et al. (2022) specified that peer-assessment "is not merely about transferring information from a knowledgeable person to a rookie learner, but an active process where learners are engaged with continuous assessment of knowledge needs and learn to re-construct relevant cognitive understanding in context" (p. 2).

The third main formative assessment strategy is a particular form of peer-assessment carried out in groups. In this case, feedback is provided by a group of students for developmental purposes (Baker, 2007). This form of peer-assessment is particularly useful for professionals in education when they have to present, argue and discuss with peers the description and the design of educational activities (Homayouni, 2022). Essentially, group-assessment is a form of "co-operative or peer-assisted learning that encourages individual students in small groups to coach each other in turn so that the outcome of the process is a more rounded understanding and a more skillful execution of the task in hand than if the student was learning in isolation" (Asghar, 2010, p. 403). For all formative assessment strategies, a rubric designed by the teacher to lead the students' reflection and assessment is fundamental (Andrade and Boulay, 2003).

Morris et al. (2021) highlighted the evidence on university students' academic performance when instructors use formative

assessment strategies on a regular basis. These authors identified the four main points of an effective formative assessment strategy: (a) content, detail and delivery, (b) timing and spacing, (c) peers, and (d) technology. They found that formative activities can support positive educational experiences and enhance student learning and development.

2.4 Critical views of formative assessment in higher education contexts

In addition to the above aspects of formative assessment, it is important to consider criticisms related to the use of formative assessment strategies in higher education and how they influence this work. For instance, [Brown \(2019\)](#) posed the question: is formative assessment really assessment? According to this author, formative assessment has some aspects of assessment but it cannot be considered assessment. A series of reasons included these points: feedback occurs in ephemeral contexts; it is not possible to recognize the interpretations of teachers and, additionally, it is not possible to understand if those interpretations are adequately accurate; ultimately, stakeholders cannot know if the conclusions are valid or not.

[Hamodi et al. \(2017\)](#) emphasized that formative assessment strategies risk conflicts between students when aimed at giving or affecting grades. However, students also “recognize that the formative assessment they experienced as university students has proved valuable in their professional practice in schools” (p. 186). The social difficulties that can arise from formative assessment is a main concern ([Biesma et al., 2019](#)). Additionally, both [Koka et al. \(2017\)](#) and [Bond et al. \(2020\)](#) stressed that formative assessment is not effective if not used regularly and [Crossouard and Pryor \(2012\)](#) emphasized that unexamined theories related to formative assessment can be implicit in practice and potentially narrow the educational possibilities of an intervention. Similarly, [Morris et al. \(2021\)](#), following [Yorke \(2003\)](#), question the most effective assessment approaches for higher education students, because there is a lack of clarity across higher education.

Ultimately, the critical views of formative assessment are focused on the following questions: is it necessary to link formative assessment with a grading scale? In what ways should the informal interactions raised during the formative assessment activities be connected with instructors’ formal grading? Is there evidence regarding the effectiveness of formative assessment strategies? How might university students be motivated to use formative assessment in their future workplace environments? Lastly, how can summative and formative strategies be effectively integrated throughout the learning activities? Situated within these theories, concepts, critiques, and questions, this study aims to provide views and clarity from the Italian context.

3 Context of the study and research questions

3.1 Context

Before presenting the aims and research questions, it is useful to describe Italian university programs for pre-service teachers, social workers, and heads of social services. Pre-service teachers must have

a 5-year degree to teach either kindergarten (pupils aged 3–5) or the first five grades of primary school (pupils aged 6–10). Social workers must have a 3-year bachelor degree in educational sciences with a focus in one of two main programs. The first program trains professionals to work in educational contexts outside schools: educational services for minors with family difficulties; educational services for migrants; counselling centers; educational centers for unaccompanied foreign minors; anti-violence centers; centers for minors who committed crimes; etc. The second program is dedicated to social workers who will be employed as early childhood educators for pupils aged 0 to 3 in kindergarten. It is important to note that these professionals are called educators for early childhood services, but do not hold teacher status. Lastly, heads of social services must have a 2-year masters degree in educational sciences. These students go on to work in two main fields: as designers of local/national/international educational projects carried out by private and public bodies/centers; or, as coordinators of private and public social and educational services focused on different sectors such as early childhood, minors, migrants, etc. As indicated earlier, the Italian context is particularly suitable for this study since it has a higher education system based mainly on summative assessment. Further, given the learner-centred nature of formative assessment, it also conceptually and methodologically fitting to solicit the voices of and feedback comments made by the student participants regarding the effectiveness of formative strategies in this predominantly summative assessment culture.

3.2 Aims and research questions

This exploratory study aimed to investigate the role of formative assessment strategies carried out in university programs for pre-service teachers, social workers, and heads of social services. The overall aim of this study was to explore the benefits and the limitations related to the use of formative assessment methods in the higher education contexts. Specifically, the study aimed to examine the characteristics and practices of self-, peer- and group-assessment. Consequently, the overall research question can be expressed as follows: to what extent can formative assessment methods help higher education students reflect on their assessment competences as future professionals in educational and social fields? Further, we identified specific research questions: did the use of formative assessment

(RQ #1) help students to achieve deep learning of several assessment methods and strategies?

(RQ #2) support students to understand the complexity of assessment procedures?

(RQ #3) promote the students’ growth from both a personal and professional point of view?

(RQ #4) motivate the students to use multiple forms of assessment when they become professionals in their respective fields?

In addition, we specified two supplementary research questions: (RQ #5) were there differences and/or specificities in the use of formative assessment in the programs for pre-service teachers, social

workers, and heads of social services? (RQ #6) were there differences, in terms of effectiveness, between the formative assessment methods: self-, peer- and group-assessment?

4 Research design

To answer the research questions, a mixed method research design was chosen. Specifically, we followed the indications by Tashakkori and Teddlie (2009), adopting a mixed method multistrand design since the exploratory study provided several stages. Additionally, as stated by Creswell and Clark (2011), the timing of data collection was concurrent since we implemented both the quantitative and qualitative strands during each phase of the study. The interpretation of the results was based on a triangulation of quantitative and qualitative data (Creswell et al., 2003) since there were no specific priorities between the kinds of data. Both data typologies were utilized to gain a deep understanding of the phenomenon. We have chosen a sequential design because it allowed us to implement the different formative assessment strategies with all students at the same time. In this way, the students could express their ideas on the effectiveness of each strategy. Conceptually this methodological approach serves to amplify the points from the theoretical framework: we sought student-centred data that was sequential, over time, with multiple access points for understanding to gain a deeper understanding of this student-centred assessment approach for higher education.

4.1 Participants, procedure, and instruments

4.1.1 Participants

As explained previously, we involved four main groups of students in the study: pre-service teachers, social workers, early childhood educators, and heads of social services. Each of these groups had specific courses included in their programs during the academic year 2022/23. Pre-service teachers (including kindergarten teachers 3-6/Primary teachers) were divided into two sub-groups. The first one had a course named “Curriculum development” scheduled at the second year of the five-year teacher education program. The second course was called “Play as educational strategy” scheduled in the third year of the same program. The social workers also took the “Curriculum development” course, but in the first year of their program. The early childhood educators had the course “Play as educational strategy” scheduled in the second year of their program. Lastly, the heads of social services had a course called “Designing and assessing learning environments” scheduled in the first year of the masters degree. All the courses contained common content and topics (learning environments, educational and assessment strategies, etc.) in addition to themes and issues relevant to the specific professional development of the students’ programs.

Table 1 indicates the demographic and educational characteristics of participants. Almost all participants are female between 19 and 22 years of age. In fact, 70% of participants were in the first two years of the bachelor degree. The groups of students, divided according to prospective job, were mainly pre-service teachers (36.54%), followed by social workers (28.21%) and early childhood educators (21.15%).

More than half participants had no prior teaching or work experiences before the study. Around 30% of participants already had some teaching or work experiences and 17.31% of participants were experienced teachers or social workers. All students were invited to take part in the study, and out of 106 pre-service teachers in the courses 57 of them (53.76%) accepted to be involved. Similarly, 44 out of 105 social workers (41.91%); 33 out of 81 early childhood educators (40.74%); and 22 out of 35 heads of social services (62.86%) participated in the study.

The principles of research ethics were strictly followed. All students enrolled in the different programs were informed about the aims, activities and procedures for the study. Participation was optional, and those who agreed to participate gave written informed consent.

4.1.2 Procedure

The procedure was split into three main phases carried out throughout the academic year 2022/23. The first phase consisted of a self-assessment activity. After a mid-term written test, composed of open-ended questions and assigned four weeks from the beginning of the course, the students had to self-assess their own exam performance following a rubric designed by the instructor. The second phase consisted of a peer-assessment activity. After another mid-term written test, similar to the first, assigned after eight weeks from the beginning of the course, the students had to assess the answers written by another student following a rubric designed by the instructor.

Table 2 summarizes the procedure, specifying that the first and the second mid-term written tests were taken by each student individually and composed of three open-ended questions focused on the topics covered by the instructor in the respective period. For instance, the first test included the concepts of learning environment and educational space. The second test contained the idea of competence, and strategies to design and assess an educational path. Each question required the students to develop two main aspects: content and argumentation. Students had to write responses that both presented topic-related content concerning the topics and that argued ideas and connections with a high-level of coherence. The first two phases involved the following levels of Bloom’s taxonomy: remembering, understanding, analyzing, and evaluating.

After the first mid-term test, the students had to reflect on their own answers, following the rubric in Table 3. They had to self-assess their three answers, so they completed the rubric three times. The levels followed the levels used in the Italian educational system: from A (advanced level) to D (beginning level). The sub-division into two sub-levels was necessary to give more opportunities to assess the nuances of their own learning. In addition, the students could add qualitative comments. Similarly, after the second mid-term test, the students had to peer-assess the answers of a classmate, using the same rubric. The peer-assessment was random and blinded, so students could not identify their peer-reviewer.

The third phase consisted of a group-assessment activity. In the last mid-term test, the students had to design an educational action plan based on a strategy such as cooperative learning, problem-based learning, case study, etc. The strategy was randomly assigned to the students by the instructor the day before the test. During the test, the students had to present their action plan to a group of peers (each group was composed of 4–5 students). After the presentation, the

TABLE 1 Demographic and educational characteristics of participants.

	Participants	%
Gender		
Male	4	2.56
Female	152	97.44
Age		
19–20	63	40.38
21–22	50	32.05
23–25	27	17.31
>26	16	10.26
Year of attendance		
1 st year bachelor	38	24.37
2 nd year bachelor	74	47.45
3 rd year bachelor	22	14.09
1 st year master	22	14.09
Education area		
Kindergarten teacher 3-6/Primary teacher	57	36.54
Social worker	44	28.21
Kindergarten educator 0–3	33	21.15
Head of social services	22	14.10
Working/teaching experience		
Never	81	51.92
Few experiences (some days/weeks)	23	14.74
Many experiences (some months/1 year)	25	16.03
Full experience (2 or more years)	27	17.31

TABLE 2 The detailed procedure.

Weeks after the beginning	Typology of mid-term test	Structure of mid-term test	Questions' contents	Typology of formative assessment	Bloom's taxonomy levels
4	Written test	Individual – Three open-ended questions	Topics faced in the first 4 weeks	Self-assessment	Remembering Understanding Analyzing Evaluating
8	Written test	Individual – Three open-ended questions	Topics faced in the second 4 weeks	Peer-assessment	Remembering Understanding Analyzing Evaluating
12	Simulation	Group – Presentation of an educational strategy	Educational strategies	Group-assessment	Applying Creating

group students had to assess the presentation following the rubric shown in [Table 4](#).

4.1.3 Instrument

The research procedure involved three observational moments. After each mid-term test, the students were asked to fill in an online questionnaire focused on the research questions. The questionnaire was composed of three sections. The first section contained the

demographic and educational characteristics of participants (see [Table 1](#)). The second included three scales dedicated to each typology of formative assessment (self-, peer-, and group-). Each scale was composed of five items aimed to underscore the participants' ideas regarding the different techniques of formative assessment. The items were linked to the following factors: this formative assessment activity (self, peer, or group) supported students' development in: (a) learning assessment strategies (Learning); (b) understanding complexity of

TABLE 3 The rubric for self- and peer-assessment.

Bloom's taxonomy levels	Rubric's dimensions	Rubric's indicators	Levels							
			A1	A2	B1	B2	C1	C2	D1	D2
Remembering Understanding	Topics' presentation	Spelling and syntax								
		Clarity of exposition								
		Exhaustiveness of contents								
Analyzing Evaluating	Processing information related to the topics'	Focusing on the requested topic								
		Connection and coherence among contents								
		Argumentation level								
If you want, you can add qualitative comments										

TABLE 4 The rubric for group-assessment.

Bloom's taxonomy levels	Rubric's dimensions	Rubric's indicators	Levels							
			A1	A2	B1	B2	C1	C2	D1	D2
Applying	Activities' description	The strategy has been well described								
		The strategy's sequences are coherent and well connected								
Creating	Educational impact of the activities	The strategy supports complex learning processes								
		The strategy involves deeply the participants								
If you want, you can add qualitative comments										

assessment procedures (Complexity); (c) growing from a personal point of view since I'm more aware of my capacities and limitations (Personal); (d) growing from a professional point of view since I'm more aware of my competences as teacher/social worker (Professional); (e) increasing my motivation to use formative assessment strategies in the future (Motivation).

These items were rated by the students with a four-point Likert scale, from 4 (Yes, this formative assessment activity has been very useful/effective) to 1 (No, this formative assessment activity has not been useful/effective at all). In addition, in the third section, they were able to add free qualitative comments and, ultimately, they were asked to suggest whether the instructor should use this strategy again the following year with new students.

4.2 The qualitative–quantitative data analysis procedure

The data analyses were performed both from a qualitative and quantitative point of view. The qualitative data were coded with NVivo

14 following the three steps suggested by grounded theory: open coding, axial coding and selective coding (Charmaz, 2014; Corbin and Strauss, 2015; De Smet et al., 2019). The quantitative data were analyzed with SPSS 29 and focused on reliability analyses, ANOVA for repeated measures, the Friedman test, Exploratory Factor Analysis (EFA), and non-parametric tests to highlight potential statistically significant differences between the demographic and educational characteristics of the participants, considering gender, age, year of attendance, education area, working/teaching experience as variables.

5 Data analysis and findings

The data analysis is structured as follows. On the basis of the study aims indicated in the introduction, both quantitative and qualitative findings will underline, firstly, which formative assessment activities are more suitable in the higher education contexts. So, the findings will highlight the favorite techniques by the students: self-, peer-, and/or group-assessment. Secondly, the findings will emphasize which factors (Learning, Complexity, Personal, Professional, Motivation),

TABLE 5 Reliability coefficients.

Scale	Cronbach α	McDonald ω	Average inter-item correlation
Self-assessment (5 items)	0.796	0.813	0.432
Peer-assessment (5 items)	0.783	0.805	0.421
Group-assessment (5 items)	0.894	0.911	0.625
Critical values	Good $0.800 < \alpha < 0.900$ Excellent $\alpha > 0.900$		0.400 to 0.500 (Spiliotopoulou, 2009) 0.300 to 0.700 (DeVon et al., 2007)

linked with the formative assessment, have been more developed. So, the findings can tell us in which way formative assessment can become a basic and crucial concept for the students. Finally, the findings will stress if the use of formative assessment supports the students' motivation to use formative assessment in their own future professional fields.

5.1 Quantitative findings

The first quantitative data analysis was focused on the instrument's reliability, so we used the following coefficients: Cronbach's Alpha (α); McDonald's Omega (ω); and average inter-item correlation. Table 5 summarizes the results.

Then, the quantitative analysis concentrated on the potential differences, in terms of effectiveness, between the formative assessment methods: self-, peer- and group-assessment. The Friedman test revealed that there is a statistically significant difference among these three methods ($\chi^2 = 9.726$ $df = 2$ $p < 0.008$). Specifically, the Conover's *post hoc* comparisons showed that there are not differences between peer- and group-assessment ($t = 0.174$ $p < 0.862$) but there are significant differences between self- and peer-assessment ($t = 2.786$ $p < 0.005$) and between self- and group-assessment ($t = 2.612$ $p < 0.009$). The results were specifically higher for peer- and group-assessment.

Analyzing each scale, we found significant differences between the items related to the research questions. Particularly, within the scales regarding the self- and the peer-assessment methods, we found that the scores for the factors named "Complexity" and "Motivation" were higher than the others.

Similarly, analyzing the potential differences related to each factor between the scales, we found that the scores for the factor "Complexity" were higher ($\chi^2 = 11.647$ $df = 2$ $p < 0.003$) in the self- ($t = 2.054$ $p < 0.041$) and peer-assessment ($t = 3.407$ $p < 0.001$) compared to those in the group-assessment. In addition, the scores for the factor "Professional" are higher ($\chi^2 = 14.630$ $df = 2$ $p < 0.001$) in the peer- ($t = 2.692$ $p < 0.007$) and group-assessment ($t = 3.718$ $p < 0.001$) compared to those in the self-assessment.

Ultimately, analyzing which are the factors with the highest scores across the scales, we found that "Complexity" and "Motivation" were more appreciated by the participants compared to other factors ($\chi^2 = 73.945$ $df = 4$ $p < 0.001$).

5.1.1 Differences among participants

The differences between the demographic and educational characteristics of the participants were performed with ANOVA for repeated measures. Tables 6A–C summarize the results. The

assessment methods are indicated with SA (self-assessment), PA (peer-assessment) and GA (group-assessment).

Table 6A presents the statistically significant differences between the variable 'Age' with its dimensions and each assessment method. Table 6B displays the differences between all variables with related dimensions and each factor comparing the assessment methods. Lastly, Table 6C reveals the differences between all variables with related dimensions and each factor within each assessment method.

Additionally, we found interesting differences regarding the importance of factors for some variables in general. Regarding the variable 'Age', the students 19–20 years old appreciated the factor 'Complexity' more than the factors 'Personal' and 'Professional' (respectively, MD = 0.501 $p < 0.006$ and MD = 0.621 $p < 0.000$). Again, the students 19–20 years old highly valued the factor 'Motivation' compared to the factor 'Professional' (MD = 0.436 $p < 0.001$). Lastly, the students older than 26 appreciated the factor 'Complexity' more than 'Professional' (MD = 0.685 $p < 0.030$).

Considering the variable 'Year of attendance', the students at the II year of bachelor degree gave higher scores to the factors 'Complexity' and 'Motivation' compared to 'Professional' (respectively, MD = 0.360 $p < 0.036$ and MD = 0.319 $p < 0.014$).

Within the variable 'Education area', pre-service teachers appreciated the factor 'Complexity' more than 'Personal' and 'Professional' (respectively, MD = 0.454 $p < 0.037$ and MD = 0.444 $p < 0.034$). Similarly, social workers appreciated 'Complexity' more than 'Professional' (MD = 0.501 $p < 0.019$).

Regarding the variable 'Working/teaching experience', students with 'never' work experiences, appreciated 'Complexity' more than 'Professional' (MD = 0.449 $p < 0.005$). Ultimately, students with 'full experience' appreciated 'Complexity' and 'Motivation' more than 'Professional' (respectively, MD = 0.542 $p < 0.012$ and MD = 0.444 $p < 0.010$).

5.1.2 Exploratory factor analysis (EFA)

We decided to perform an Exploratory factor analysis (EFA) because it was interesting to identify potential common factors that might explain the structure of the instrument and the validity of the measured variables (Watkins, 2018). The EFA was completed with varimax rotation and Kaiser normalization, using principal components extraction, because we were interested in highlighting with eigenvalues > 1 to emphasize the presence of latent factors.

The results indicate that the sample was adequate since the Kaiser-Meyer-Olkin test was 0.844; additionally, Bartlett's Test of Sphericity revealed a p -value of < 0.000 ($\chi^2 = 1041.176$; $df = 105$). Lastly, the goodness of fit test was 81.104 ($df = 51$; $p < 0.005$).

TABLE 6 Differences between the participants' demographic and educational characteristics.

A				
Variable		Between assessment methods		
		SA	PA	GA
Age	19–20	No differences (nd)	nd	Compared to >26 (MD = 0.694 $p < 0.004$)
	21–22	nd	Compared to GA higher than 23–25 (MD = 0.694 $p < 0.004$)	nd
	23–25	nd	nd	nd
	>26	nd	Compared to GA (MD = 0.498 $p < 0.022$)	nd

B				
Variable		Between assessment methods for each factor		
		SA	PA	GA
Age	19–20	nd	nd	Compared to SA for professional (MD = 0.576 $p < 0.000$) Compared to SA for motivation (MD = 0.480 $p < 0.001$)
	21–22	nd	Compared to GA for complexity (MD = 0.289 $p < 0.012$)	nd
	23–25	nd		nd
	>26	nd	Compared to GA for learning (MD = 0.501 $p < 0.024$) Compared to GA for complexity (MD = 0.701 $p < 0.000$)	nd
Year of attendance	I year bachelor	Compared to GA for complexity (MD = 0.501 $p < 0.010$)	Compared to GA for complexity (MD = 0.549 $p < 0.000$)	nd
	II year bachelor	nd	nd	Compared to SA for professional (MD = 0.378 $p < 0.006$) Compared to PA for professional (MD = 0.215 $p < 0.038$)
	III year bachelor	nd	Compared to GA for complexity (MD = 0.556 $p < 0.003$)	nd
	I year master	nd	Compared to GA for professional (MD = 0.389 $p < 0.006$) Compared to GA for complexity (MD = 0.361 $p < 0.018$)	nd
Education area	Kindergarten teacher 3–6/ Primary teacher	nd	nd	Compared to SA for professional (MD = 0.391 $p < 0.027$) Compared to PA for professional (MD = 0.309 $p < 0.012$)
	Social worker	nd	Compared to GA for complexity (MD = 0.443 $p < 0.000$)	nd
	Kindergarten educator 0–3	nd		nd
	Head of social services	nd	Compared to GA for complexity (MD = 0.361 $p < 0.018$) Compared to GA for professional (MD = 0.389 $p < 0.006$)	nd
Working/ Teaching experience	Never	nd	Compared to GA for complexity (MD = 0.290 $p < 0.007$)	nd

(Continued)

TABLE 6 (Continued)

B				
Variable		Between assessment methods for each factor		
		SA	PA	GA
	Few experiences (some days/weeks)	nd	nd	nd
	Many experiences (some months/1 year)	nd	Compared to GA for motivation (MD = 0.462 $p < 0.010$) Compared to GA for complexity (MD = 0.635 $p < 0.000$) Compared to SA for complexity (MD = 0.346 $p < 0.034$)	nd
	Full experience (2 or more years)	nd	nd	Compared to SA for professional (MD = 0.389 $p < 0.049$)

C				
Variable		Between factors within assessment methods		
		SA	PA	GA
Age	19–20	nd	nd	Compared to >26 for personal (MD = 0.985 $p < 0.000$) Compared to >26 for motivation (MD = 0.829 $p < 0.001$) Compared to >26 for complexity (MD = 0.709 $p < 0.004$)
	21–22	nd	nd	Compared to >26 for personal (MD = 0.635 $p < 0.018$)
	23–25	nd	nd	Compared to >26 for personal (MD = 0.817 $p < 0.002$) Compared to >26 for motivation (MD = 0.687 $p < 0.012$) Compared to >26 for complexity (MD = 0.647 $p < 0.021$)
	>26	nd	nd	nd
Year of attendance	I year bachelor	nd	nd	nd
	II year bachelor	nd	nd	Compared to I year master for professional (MD = 0.632 $p < 0.001$) Compared to III year bachelor for motivation (MD = 0.613 $p < 0.022$) Compared to I year master for motivation (MD = 0.530 $p < 0.010$) Compared to III year bachelor for complexity (MD = 0.693 $p < 0.006$)
	III year bachelor	nd	nd	nd
	I year master	nd	nd	nd
Education area	Kindergarten teacher 3–6/Primary teacher	nd	nd	Compared to heads of educational services for professional (MD = 0.517 $p < 0.021$)
	Social worker	nd	nd	nd
	Kindergarten educator 0–3	nd	nd	nd
	Head of social services	nd	nd	nd
Working/Teaching experience	Never	nd	Compared to full experience for learning (MD = 0.416 $p < 0.008$)	nd
	Few experiences (some days/weeks)	nd	nd	nd
	Many experiences (some months/1 year)	nd	nd	nd
	Full experience (2 or more years)	nd	nd	nd

TABLE 7 EFA factor loadings.

	Factor 1	Factor 2	Factor 3	Factor 4
GA motivation	0.813			
GA professional	0.793			
GA personal	0.773			
GA learning	0.692			
GA complexity	0.669			0.403
SA professional		0.769		
SA personal		0.701		
SA motivation		0.626		
SA learning		0.614		
PA professional			0.777	
PA personal			0.585	
PA motivation			0.547	
PA complexity			0.464	0.648
PA learning			0.439	
SA complexity				0.455

The EFA underscored four factors (see Table 7). Factors 1, 2 and 3 follow the structure of the instrument indicating, respectively, the group-assessment (38.86% of explained variance), the self-assessment (13.92%) and the peer-assessment (8.89%). The fourth factor indicated the importance of understanding the complexity of assessment procedures (7.22% of explained variance).

5.1.3 Next year

A specific questionnaire item asked students to indicate if, in their opinion, the formative activities should be repeated for students in the following year. This item was rated by the students on a four-point Likert scale, from 4 (Yes, absolutely) to 1 (Not at all). On this item, 97.51% of students chose options 4 and 3 for self-assessment, 94.12% for peer-assessment and 88.19% for group-assessment. The Friedman test did not show any statistically significant differences between the assessment methods ($\chi^2 = 0.036$ df = 2 $p < 0.982$).

5.2 Qualitative findings

The qualitative data analysis highlighted four common categories for each assessment method: ‘Positive aspects’; ‘Limits’; ‘Assessment issues’ and ‘Organizational issues’. The first category includes codes related to positive characteristics of the assessment method, and the second category (Limits) contains codes that indicate the main weaknesses of the assessment method. The ‘Assessment issues’ category comprises codes focused on the forms/modalities of assessment. Lastly, the ‘Organizational issues’ category focuses on the technical questions related to the assessment method. As shown in Figure 1, categories with the same name can include codes shared between two or three assessment methods or specific codes that emphasize exclusive features of one assessment method.

Tables 8, 9 show in detail the categories and the codes split into the assessment methods. Additionally, these tables display examples of sentences which illustrate the codes, the number of participants

who wrote sentences related to the code, and the frequency of the codes.

In addition to the information enclosed in Tables 8, 9, it is interesting to underline that 59 students considered the self-assessment activity ‘useful,’ whereas 51 students identified the peer-assessment activity and 44 students considered the group-assessment activity ‘very useful.’

6 Discussion

Both the quantitative and qualitative findings of this study provide important insights about the use of formative assessment in the higher education contexts and indicate significant implications from a practical point of view.

First of all, comparing the three formative assessment strategies, peer- and group-assessment proved to be more appreciated among the students involved in the study. Also, self-assessment was evaluated as a positive activity but peer- and group-assessment allowed students to reflect deeply on their own learning processes, create feedback and give them the opportunity to improve and modify their learning strategies.

Among the five factors that characterized this study (Learning; Complexity; Personal; Professional; Motivation), two of them revealed their importance. The scores for “Complexity” (understanding complexity of assessment procedures) and “Motivation” (increasing my motivation to use formative assessment strategies in the future) were significantly higher than the other factors mainly for peer- and group-assessment activities. Opportunities to understand the complexity of assessment procedures and to increase motivation for future use of formative assessment strategies represented crucial elements of the formative assessment activities. These reasons are explained in the qualitative comments. Firstly, formative assessment methods allowed students to recognize that an assessment procedure is not simple and linear but it is composed of many connected

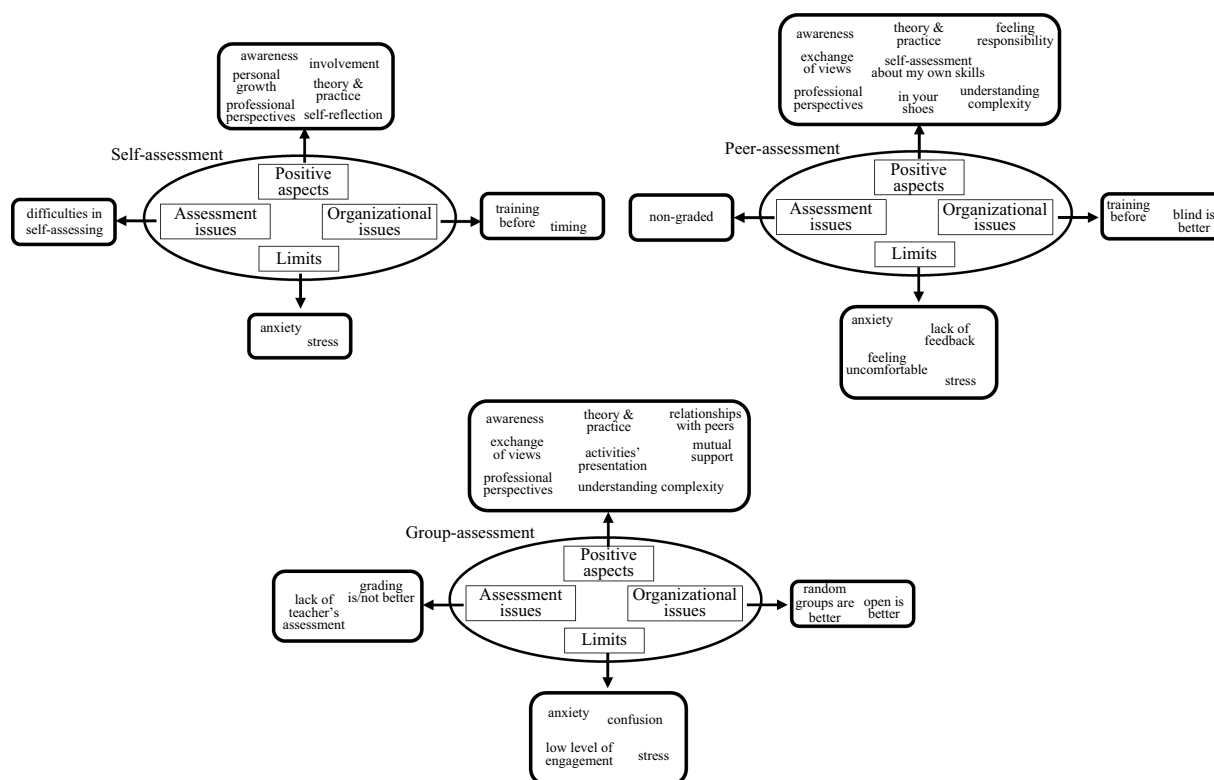


FIGURE 1
Map of categories and codes.

components, both summative and formative. Consequently, students must be able to learn and design several assessment methods linked in a coherent way, as indicated by [Ismail et al. \(2022\)](#). Secondly, the use of formative assessment strategies in higher education increased the probability of using formative assessment techniques in the future, confirming the study carried out by [Hamodi et al. \(2017\)](#).

Regarding differences among participants, peer- and group assessment were more appreciated by older students, indicating that these students rely on forms of professional relationships to develop their own professional perspectives, as denoted by [Biesma et al. \(2019\)](#) and [Montgomery et al. \(2023\)](#). Peer-assessment is, in general, more appreciated than group-assessment for developing “Complexity,” “Motivation” and, also, “Learning” but group-assessment was more valued than peer- and self- assessment for increasing the “Professional” factor. This observation means that group-assessment represents an effective simulation of a work experience ([Homayouni, 2022](#)). The younger students appreciated group-assessment more than peer-assessment for the “Personal,” “Complexity” and “Motivation” factors, showing that younger students needed formative assessment activities for developing awareness of their own skills and capacities. Finally, students with no work experience appreciated the self-assessment activities for the factor “Learning,” revealing that they valued moments for reflecting on themselves.

The last quantitative finding arises from the exploratory factor analysis which confirmed that the factor related to understanding the complexity of assessment procedures is particularly significant for the participants.

The findings highlighted by the qualitative analysis reveal essential results of the study. Regarding the positive aspects, all three strategies

helped students in growing their awareness in recognizing their learning processes, effectively connecting theory to practice, and consciously fostering their professional perspectives, as indicated by [Tillema \(2010\)](#). In particular, peer- and group-assessment boosted students’ capacity to exchange ideas and views with others; consequently, these strategies amplified peer relationships through responsibility to each other and deepened student understanding of the complexity of assessment procedures.

Additionally, the qualitative analysis identified three main limits in two aspects of formative assessment: lack of feedback in peer-assessment, and in group-assessment, low level of engagement and high level of confusion. Consequently, these two strategies must be carefully implemented. Even though they were highly appreciated by the students in general, the qualitative sections of the questionnaire indicate that peer- and group-assessment must be carried out in an adequate setting allowing students to express their ideas and motivations without pressure and confusion.

Lack of feedback, low level of engagement and high level of confusion represent three risks of formative assessment previously identified in the literature. The lack of feedback was already mentioned by [Brown \(2019\)](#) and the other two points were indicated by [Crossouard and Pryor \(2012\)](#).

The qualitative analysis also raised issues about the assessment quality and pinpointed crucial features related to the relationship between summative and formative assessment. Some students felt that group-assessment should be followed by a formal grade, while others indicated that the same activity should be carried out more informally. Regardless, all students felt the lack of an instructor-designed assessment to ensure that activities were well designed ([Homayouni,](#)

TABLE 8 Codes included into 'Positive aspects' category.

Category	Code	Assessment method	Sentences	Participants	Frequency
Positive aspects	Awareness	SA	This experience has been quite positive since I had the opportunity to develop my self-awareness	30	34
		PA	It's crucial to be able to assess a mate and being assessed	6	8
		GA	A reciprocal assessment allowed me to understand my strengths and weaknesses	8	10
	Involvement	SA	It was very useful to get involved myself	9	10
	Personal growth	SA	This activity allowed me to growth as a person	13	14
	Theory & Practice	SA	I could realize what I'm able to do actually	4	4
		PA	I could connect effectively contents and real-world situations	8	9
		GA	After this activity, I could study the contents very easily	17	19
	Professional perspectives	SA	It was important to assess themselves from a professional point of view	23	24
		PA	The qualitative comments helped me in understanding my professional skills	19	21
		GA	I could learn many assessment strategies and compare them with my mates	27	30
	Self-reflection	SA	It was crucial to reflect on your own skills and capacities	36	39
	Exchange of views	PA	Exchanging ideas with my mates allowed me to discover aspects of myself	29	33
		GA	I liked a lot to present an activity prepared by myself and I appreciated to listen to my mates' activities	38	46
	Feeling responsibility	PA	I felt the responsibility in assessing the task of a mate	12	13
	In your shoes	PA	I made many efforts to put myself in my mate's shoes	13	14
	Self-assessment about my own skills	PA	Assessing my mate, I assessed also myself, at the same time	17	19
	Understanding complexity	PA	This activity allowed me to understand how much is difficult to create good assessment practices	15	16
		GA	The activity was quite challenging because it's really difficult assessing somebody	14	15
	Relationships with peers	GA	Group-assessment allowed me to discuss with my mates and identify the right assessment	12	13
	Mutual support	GA	My mates were so supportive!	8	8

TABLE 9 Codes included into 'Limits', 'Assessment issues' and 'Organizational issue' categories.

Category	Code	Assessment method	Sentences	Participants	Frequency
Limits	Anxiety/stress	SA	I was not prepared and I was not able to assess myself	15	15
		PA	Assessing the task of another mate is always difficult	13	13
	Lack of feedback	PA	I needed also oral feedback	29	30
	Feeling uncomfortable	PA	I felt quite uncomfortable because I was afraid not to write the right things	11	13
	Low level of engagement	GA	My mates were not motivated so the activity was quite useless	15	17
	Confusion	GA	The activity must be well organized otherwise many mates were distracted	13	14
Assessment issues	Difficulties in self-assessing	SA	I was not aware of the values of this procedure	14	15
	Non-graded	PA	I felt the responsibility but I was calm because my assessment did not affect the final grade	4	5
	Grading is/not better	GA	Sometimes, it would be better to have a formal grade, sometimes not	6	8
	Lack of teacher's assessment	GA	I needed to receive an assessment also by the teacher	21	22
Organizational issues	Training before	SA	It was necessary to have a training before the activity	10	11
		PA	I've never done this activity; I needed to be trained before	3	4
	Blind is better	PA	It was better the blind procedure so I was totally free to express my ideas	13	14
	Random groups are better	GA	Random groups allow us to be more free and professional	9	11
	Open is better	GA	Group-assessment should be carried out face-to-face, not with an anonymous rubric	8	9

2022). These are fundamental points because the university teachers must decide if the formative assessment should affect and/or modify the final grade or not. This matter represents an open question that may also be influenced by the larger assessment culture (Doria et al., 2023). In this study, we decided that the formative assessment should not change the final grade because the formative strategies were focused on the metacognitive aspects.

From an organizational point of view, the students needed forms of training to effectively carry out both self- and peer-assessment. The organization of group-assessment was a particularly debated aspect because some students desired an open assessment discussion whilst other students wanted a completely blind assessment to avoid any conflict with their peers, as indicated by Hamodi et al. (2017).

7 Limitations of the study

This study presents some limitations. First, the participants are quite heterogeneous. They are all students enrolled in courses focused on educational contexts but the structure of the courses is different and this situation can affect the application of the assessment strategies. The second limitation is represented by the different topics faced in the different subjects. Even if the formative strategies can be considered valid for all subjects, it is necessary to reflect on the

specificities of each subject to verify potential differences in managing and arranging the formative techniques. Finally, the study procedure lasted the whole academic year. It is likely that the elapsed time can represent a bias for students who faced three formative strategies throughout the course.

8 Conclusions and implications for policy and practice

Instead of assuming that 'theory' is solely the domain of expert educational researchers, we reassert our alignment with other researchers who recognize the critical significance of educational practitioners' comprehension of theory in actualizing specific practices (Crossouard and Pryor, 2012). Our findings indicate that the participants in our study significantly enhanced their capacities in reflecting on their assessment competences.

Notably, the higher education students in this study understood the complexity of the assessment procedures (RQ #2) and were motivated to use multiple forms of assessment in their future work (RQ #4). In terms of learner-centred feedback on effectiveness (RQ #6), peer- and group-assessment strategies emerged as the most effective and productive methods from a formative point of view. As for RQ #5, we did not find specific differences in the use of formative

assessment in the programs for pre-service teachers, social workers, and heads of social services. Furthermore, despite the different opinions expressed by the participants, the majority declared that it would be important to repeat the formative assessment activities with students in the educational profession programs the following academic year.

These conclusions and the findings have implications for future work in this area. The value of integrating formative assessment and learner-identified feedback into instructional practices is evident and will inform future studies and formative assessments, with some recommendations.

For future studies, our first recommendation is an emphasis on peer-assessment strategies. From a practical and organizational perspective, ways to strengthen the peer-assessment components, such as mandatory peer comments are one recommendation for avoiding a reported lack of feedback. In the present study, qualitative comments were optional and some students did not write any comments to their peers.

Our second recommendation further derives from the students who did not receive peer comments and emphasized that the rubric scores and indicators were not enough to express meaningful feedback. Supporting students in their development of effective peer feedback practices would be an important next step in strengthening this component of instruction and developing student competencies in formative assessment for the education professions.

A third recommendation is related to group assessment. To avoid low levels of engagement and high levels of confusion during group-assessment, additional measures are suggested. One option includes arranging random groups of students with the presence of one instructor in each group to ensure sufficient guidance during the activity. Following the activity, the instructor should leave the group to promote open discussion and support students' assessment of the peer presentation, so that students can more freely and informally assess the learning. Then, the instructor will come back into the group to give a formal grade. We anticipate that this response could resolve issues of lack of teacher assessment and the question of assigning a formal evaluative grade to this activity, thereby effectively balancing summative and formative assessment methods (Lui and Andrade, 2022).

Typically, formative assessment is aimed at fostering developmental functions mainly related to metacognitive aspects of learning, whereas summative assessment is commonly targeted toward evaluative and administrative decisions (grades, reports, etc.) about performance (Baker, 2007). As the findings of this study reveal, we emphasize the main original contribution of this study: formative assessment is an effective combination of developmental and evaluative purposes.

We further recommend that future studies should investigate both (a) metacognitive steps to underline the formative values of assessment, and (b) evaluative steps to underscore the summative standards to be reached by a professional in education. This design will both align with the recent European Commission report (European Commission, Directorate-General for Education, Youth, Sport and Culture, et al.,

2023) and support the creation of a strong combination of formative and summative feedback to deeply assess learners' competences.

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 approval was not required for the studies involving humans because this study strictly followed the principles of ethics. The participants were informed about the aims, the activities and the instruments designed for this study. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

DP: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft. EN: Formal analysis, Methodology, Writing – original draft. EM: Formal analysis, Methodology, Writing – original draft. MI: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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

- Andrade, H. L. (2019). A critical review of research on student self-assessment. *Front. Educ.* 4:87. doi: 10.3389/feduc.2019.00087
- Andrade, H. G., and Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *J. Educ. Res.* 97, 21–30. doi: 10.1080/00220670309596625
- Asghar, A. (2010). Reciprocal peer coaching and its use as a formative assessment strategy for first-year students. *Assess. Eval. High. Educ.* 35, 403–417. doi: 10.1080/02602930902862834
- Baker, D. F. (2007). Peer assessment in small groups: a comparison of methods. *J. Manag. Educ.* 32, 183–209. doi: 10.1177/1052562907310489

- Bawane, J., and Sharma, R. (2020). Formative assessments and the continuity of learning during emergencies and crises. NEQMAP 2020 thematic review. UNESCO.
- Biesma, R., Kennedy, M.-C., Pawlikowska, T., Brugha, R., Conroy, R., and Doyle, F. (2019). Peer assessment to improve medical student's contributions to team-based projects: randomized controlled trial and qualitative follow-up. *BMC Med. Educ.* 19:371. doi: 10.1186/s12909-019-1783-8
- Bond, E., Woolcott, G., and Markopoulos, C. (2020). Why aren't teachers using formative assessment? What can be done about it? *Assess. Matters* 14, 112–136. doi: 10.18296/am.0046
- Brown, S. (1999). "Institutional strategies for assessment" in *Assessment matters in higher education: Choosing and using diverse approaches*. eds. S. Brown and A. Glasner (Buckingham: SRHE and Open University Press), 3–13.
- Brown, G. T. (2019). Is assessment for learning really assessment? *Front. Educ.* 4:64. doi: 10.3389/feduc.2019.00064
- Bruner, J. S. (1970). "Some theories on instruction" in *Readings in educational psychology*. ed. E. Stones (Methuen), 112–124.
- Cefai, C., Downes, P., and Cavioni, V. (2021). A formative, inclusive, whole school approach to the assessment of social and emotional education in the EU. NESET report. Publications Office of the European Union. doi: 10.2766/905981
- Charmaz, K. (2014). *Constructing grounded theory*. London: Sage.
- Clark, I. (2012). Formative assessment: assessment is for self-regulated learning. *Educ. Psychol. Rev.* 24, 205–249. doi: 10.1007/s10648-011-9191-6
- Corbin, J., and Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Los Angeles, CA: Sage.
- Creswell, J. W., and Clark, V. P. (2011). *Designing and conducting mixed methods research*. Thousand Oaks: Sage.
- Creswell, J. W., Clark, V. P., Gutmann, M., and Hanson, W. (2003). "Advanced mixed methods research designs" in *Handbook of mixed methods in social & behavioral research*. eds. A. Tashakkori and C. Teddlie (Thousand Oaks, CA: Sage), 209–240.
- Crossouard, B., and Pryor, J. (2012). How theory matters: formative assessment theory and practices and their different relations to education. *Stud. Philos. Educ.* 31, 251–263. doi: 10.1007/s11217-012-9296-5
- Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assess. Educ. Princ. Policy Pract.* 21, 149–166. doi: 10.1080/0969594x.2014.898128
- De Smet, M. M., Meganck, R., Van Nieuwenhove, K., Truijens, F. L., and Desmet, M. (2019). No change? A grounded theory analysis of depressed patients' perspectives on non-improvement in psychotherapy. *Front. Psychol.* 10:588. doi: 10.3389/fpsyg.2019.00588
- Dekker, H., Schönrock-Adema, J., Snoek, J. W., van der Molen, T., and Cohen-Schotanus, J. (2013). Which characteristics of written feedback are perceived as stimulating students' reflective competence: an exploratory study. *BMC Med. Educ.* 13, 1–7. doi: 10.1186/1472-6920-13-94
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., et al. (2007). A psychometric toolbox for testing validity and reliability. *J. Nurs. Scholarsh.* 39, 155–164. doi: 10.1111/j.1547-5069.2007.00161.x
- Doria, B., Grion, V., and Paccagnella, O. (2023). Assessment approaches and practices of university lecturers: a nationwide empirical research. *Ital. J. Educ. Res.* 30, 129–143. doi: 10.7346/sird-012023-p129
- European Commission, Directorate-General for Education, Youth, Sport and Culture. Looney, J., and Kelly, G. (2023). Assessing learners' competences – Policies and practices to support successful and inclusive education – Thematic report, Publications Office of the European Union, 2023. Available at: <https://data.europa.eu/doi/10.2766/221856>
- Evans, D. J., Zeun, P., and Stanier, R. A. (2013). Motivating student learning using a formative assessment journey. *J. Anat.* 224, 296–303. doi: 10.1111/joa.12117
- Hadrill, R. (1995). "The NCVQ model of assessment at higher levels" in *Assessment for learning in higher education*. ed. P. Knight (London: Kogan Page), 167–179.
- Hamodi, C., López-Pastor, V. M., and López-Pastor, A. T. (2017). If I experience formative assessment whilst studying at university, will I put it into practice later as a teacher? Formative and shared assessment in initial teacher education (ITE). *Eur. J. Teach. Educ.* 40, 171–190. doi: 10.1080/02619768.2017.1281909
- Higgins, M., Grant, F., and Thompson, P. (2010). Formative assessment: balancing educational effectiveness and resource efficiency. *J. Educ. Built Environ.* 5, 4–24. doi: 10.11120/jebe.2010.05020004
- Homayouni, M. (2022). Peer assessment in group-oriented classroom contexts: on the effectiveness of peer assessment coupled with scaffolding and group work on speaking skills and vocabulary learning. *Lang. Test. Asia* 12, 1–23. doi: 10.1186/s40468-022-00211-3
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., and Boud, D. (2020). The quality of assessment tasks as a determinant of learning. *Assess. Eval. High. Educ.* 46, 943–955. doi: 10.1080/02602938.2020.1828268
- Ismail, S. M., Rahul, D. R., Patra, I., and Rezvani, E. (2022). Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill. *Lang. Test. Asia* 12, 1–23. doi: 10.1186/s40468-022-00191-4
- Jensen, L. X., Bearman, M., and Boud, D. (2023). Characteristics of productive feedback encounters in online learning. *Teach. High. Educ.* 1–15, 1–15. doi: 10.1080/13562517.2023.2213168
- Kealey, E. (2010). Assessment and evaluation in social work education: formative and summative approaches. *J. Teach. Soc. Work.* 30, 64–74. doi: 10.1080/08841230903479557
- Koka, R., Jurâne-Brémene, A., and Koķe, T. (2017). Formative assessment in higher education: from theory to practice. *Eur. J. Soc. Sci. Educ. Res.* 9:28. doi: 10.26417/ejser.v9i1.p28-34
- Leach, L., Neutze, G., and Zepke, N. (1998). "Motivation in assessment" in *Motivating students*. eds. S. Brown, S. Armstrong and G. Thompson (London: Kogan Page), 201–209.
- Lui, A. M., and Andrade, H. L. (2022). The next black box of formative assessment: a model of the internal mechanisms of feedback processing. *Front. Educ.* 7:751801. doi: 10.3389/feduc.2022.751548
- Montgomery, L., MacDonald, M., and Houston, E. S. (2023). Developing and evaluating a social work assessment model based on co-production methods. *Br. J. Soc. Work* 53, 3665–3684. doi: 10.1093/bjsw/bcad154
- Morris, R., Perry, T., and Wardle, L. (2021). Formative assessment and feedback for learning in higher education: a systematic review. *Review of. Education* 9, 1–26. doi: 10.1002/rev3.3292
- Ng, E. M. W. (2016). Fostering pre-service teachers' self-regulated learning through self- and peer assessment of wiki projects. *Comput. Educ.* 98, 180–191. doi: 10.1016/j.compedu.2016.03.015
- Nicol, D. J., and Macfarlane-Dick, D. (2006). "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. High. Educ.* 31, 199–218. doi: 10.1080/03075070600572090
- OECD (2005). *Formative assessment: improving learning in secondary classrooms*. Paris: OECD
- Ozan, C., and Kincal, R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educ. Sci. Theory Pract.* 18, 85–118. doi: 10.12738/estp.2018.1.0216
- Panadero, E., Brown, G. T., and Srijbos, J.-W. (2015). The future of student self-assessment: a review of known unknowns and potential directions. *Educ. Psychol. Rev.* 28, 803–830. doi: 10.1007/s10648-015-9350-2
- Pereira, D., Flores, M. A., and Niklasson, L. (2015). Assessment revisited: a review of research in assessment and evaluation in higher education. *Assess. Eval. High. Educ.* 41, 1008–1032. doi: 10.1080/02602938.2015.1055233
- Petty, G. (2004). *Teaching today*. London: Nelson Thornes.
- Pintrich, P. R., and Zusho, A. (2002). "The development of academic self-regulation: the role of cognitive and motivational factors" in *Development of achievement motivation*. eds. A. Wigfield and J. S. Eccles (San Diego: Academic Press), 249–284.
- Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Aust. Occup. Ther. J.* 56, 150–155. doi: 10.1111/j.1440-1630.2009.00785.x
- Suhoyo, Y., Van Hell, E. A., Kerdijk, W., Emilia, O., Schönrock-Adema, J., Kuks, J. B., et al. (2017). Influence of feedback characteristics on perceived learning value of feedback in clerkships: does culture matter? *BMC Med. Educ.* 17, 1–7. doi: 10.1186/s12909-017-0904-5
- Tashakkori, A., and Teddlie, C. (2009). *Foundations of mixed methods research*. London: SAGE Publications.
- Tillema, H. (2010). "Formative Assessment in Teacher Education and Teacher Professional Development," in *International Encyclopedia of Education*. eds. P. Peterson, E. Baker, B. McGaw. 3rd edn (Oxford: Elsevier), 563–571.
- van Gennip, N. A. E., Segers, M. S. R., and Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: the role of interpersonal variables and conceptions. *Learn. Instr.* 20, 280–290. doi: 10.1016/j.learninstruc.2009.08.010
- Wanner, T., and Palmer, E. (2018). Formative self- and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assess. Eval. High. Educ.* 43, 1032–1047. doi: 10.1080/02602938.2018.1427698
- Watkins, M. W. (2018). Exploratory factor analysis: a guide to best practice. *J. Black Psychol.* 44, 219–246. doi: 10.1177/0095798418771807
- Watson, G. P. L., and Kenny, N. (2014). Teaching critical reflection to graduate students. *Collected Essays Learn. Teach.* 7:56. doi: 10.22329/celt.v7i1.3966
- Xiao, Y., and Yang, M. (2019). Formative assessment and self-regulated learning: how formative assessment supports students' self-regulation in English language learning. *System* 81, 39–49. doi: 10.1016/j.system.2019.01.004
- Yin, S., Chen, F., and Chang, H. (2022). Assessment as learning: how does peer assessment function in students' learning? *Front. Psychol.* 13:912568. doi: 10.3389/fpsyg.2022.912568
- Yorke, M. (2003). Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice. *High. Educ.* 45, 477–501. doi: 10.1023/a:1023967026413



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo,
Portugal

REVIEWED BY

Clara Amorim,
Polytechnic Institute of Viana do Castelo,
Portugal
Annemarie Hindman,
Temple University, United States

*CORRESPONDENCE

Andrew Biemiller
✉ andyb1009@gmail.com

RECEIVED 20 August 2023

ACCEPTED 12 February 2024

PUBLISHED 05 April 2024

CITATION

Biemiller A (2024) Word meaning types
acquired before vs. after age 5: implications
for education.

Front. Psychol. 15:1280568.

doi: 10.3389/fpsyg.2024.1280568

COPYRIGHT

© 2024 Biemiller. 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.

Word meaning types acquired before vs. after age 5: implications for education

Andrew Biemiller*

Dr. Eric Jackman Institute of Child Study, Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

This article concerns two types of word meanings: nonverbal meanings which appear to be associated with neurological representations and verbally-based meanings which appear to depend in part on other words to construct meanings. Using word use data from Hart and Risley's study of children aged 19 to 36 months, and word meaning knowledge data from Biemiller and Slonim's studies of children between aged 5 to 11, meanings were classified as nonverbal or verbally-based. Biemiller and Slonim used sampled word meanings reported known from grade levels 2 to 12 reported by Dale and O'Rourke in their Living Word Vocabulary. Virtually all meanings used at age 3 or known at age 5 (preschool) were classified nonverbal. By grade two, and even more by grade five, children had added many verbally-defined meanings, although by grade five the majority of the word meanings known were still nonverbal. Evidence for neurological meaning associates are cited. Implications for vocabulary support and instruction at various ages suggest that for children under 6, supporting larger nonverbal vocabularies while after age 6 should prioritize verbally-defined meanings.

KEYWORDS

children's vocabulary, concrete, nonverbal, symbolic, abstract, verbally-defined

Introduction

In this article, I will be discussing how types of word meanings change with age, and their implications for instruction. [Lawson-Adams and Dickinson \(2020\)](#) note that meaning types form a dimension of "representational modalities" including nonverbal meanings and verbally-based meanings.¹ No doubt, more distinctions can be made, particularly as children expand their vocabularies after first or second grade.

The distinction between nonverbal vs. verbally-defined meanings is based on the hypothesis that nonverbal meanings are associations between oral wordforms and neurological representations of perceived phenomena (mainly in the posterior cortex) and procedural organized action (mainly in the anterior cortex) plus some modifiers and functors. Verbally-defined meanings are at least in part constructed out of other meanings ([Deacon, 1997](#); [Nelson, 2007](#)). In my previous studies of vocabulary development ([Biemiller and Slonim, 2001](#);

¹ Other terms for nonverbal meanings are "indexical" (per [Deacon, 1997](#); [Nelson, 2007](#)), or "concrete" meanings. Verbally-based meanings are also described as "symbolic" per Deacon and Nelson, or "lexical" per [Lawson-Adams and Dickinson \(2020\)](#).

Biemiller, 2005, 2010a,b), I was not alert to the word type distinction between nonverbal vs. verbally-defined word meanings. The previous work showed that word meanings are acquired in a predictable order and that “priority” words could be identified using this sequence. I now think that the distinction between nonverbal and verbally-defined meanings is important for educators interested in fostering vocabulary development in schools. I will be presenting data from 3 sources that show that before age 5, most or all acquire word meanings have nonverbal neurological representations. These provide the basis for constructing verbally-based meanings after age 5. However, the data presented here indicates that children in the primary grades (K–2) continue to acquire more nonverbal than verbally-based meanings. Those with larger vocabularies can acquire more verbally-based meanings. Most of what are considered “academic vocabulary” are verbally-based meanings. I will argue that building adequate vocabularies will involve attention to acquiring enough nonverbal word meanings prior to age five, and continuing addition of nonverbal meanings but with much emphasis on verbally-defined word meanings in the primary grades.

The practical significance of different word meaning types is relevant for encouraging larger vocabularies, especially for children who acquire significantly below-average size vocabularies. There is clear evidence that children with smaller vocabularies as early as age 5 or younger tend to comprehend read texts less well in later years.

Vocabulary accumulation begins early. The size (breadth) of children’s vocabulary knowledge at kindergarten predicts their reading skills (including reading comprehension²) through fourth grade (Scarborough, 2001; Dickinson and Porche, 2011). Those predictions are based on the size of PPVT vocabulary assessments in kindergarten (picture-based), which are limited mainly to “concrete” meanings (what I refer to as “nonverbal meanings”). Cunningham and Stanovich (1997) showed that vocabulary size (PPVT) in grade one was predictive of reading comprehension as late as grade 11. In addition, Hart and Risley (1995) reported that the number of words used by children by age 3 was predictive of vocabulary in grade three, using both the PPVT-R and TOLD vocabulary tests with 29 of their original 42 children.

Neurological basis of nonverbal meanings

These meaning representations are long-evolved neurological organizations in animals. Neurologically there are six levels of perceptual and procedural processes in the human cortex (Luria, 1973; Pinel and Edwards, 2008). Vertebrates—including humans—acquire *perceptual* information from senses about *objects* and *agents* (spontaneously moving objects), *settings* (places where objects, events, and actions occur), *events* (sequences of actions on objects), *sounds*, etc. (Gibson, 1969, 1979). If the Gibsons are correct, perceptual processes create neurological representations for objects, settings, and events, among other sense organ inputs. Without such representations,

animals could not distinguish between “familiar” vs. “novel” objects, sound patterns, etc. These neurological “representations” are similar to digital “files” in computers. They can be stored, retrieved, and used in perception. Similarly, vertebrates acquire and inherit *procedural* information about actions and organized procedures. Procedural processes also create neurological representations (neurological files) that can be stored and accessed for complex instinctive actions. Lorenz (1981) describes details of instincts (Lorenz, “fixed action plans” or FAPs). For both neurological representations and digitally-organized information, “file” is a metaphorical term, an analog of a physical “file” in which we store written or graphic information that can be retrieved.

The perceptual components are mainly located in the posterior (rear) cortex, while the procedural components are largely located in the anterior (front) cortex of the brain. Luria (1973) has described layers of brain functions. The perceptual areas have inner layers where raw input from sense organs arrive (afferent neurons) to at outer levels where raw inputs are organized into perceptions of objects, settings, identifiable sounds, etc. The procedural areas where output impulses for actions at the lowest cortex level (efferent neurons) are directly connected with muscles to activate them. In the outer procedural layers, procedures are organized to produce goal-oriented actions to capture prey, find edible food, care for young, mate, etc. Luria’s book was written 50 years ago. However, much more recent neurological maps appear to be in agreement with Luria’s model (e.g., Pinel and Edwards, 2008, ch. 7). In addition, many *tasks* must include connections between procedures and perceptions (e.g., hunting a mouse.). Neurological programs integrating raw information and movement into ways of surviving and reproducing probably includes some “behavioral grammar,” as suggested by Lashley (1951).

Nonverbal meanings

Paivio (1986), Clark and Paivio (1991), and Lawson-Adams and Dickinson (2020) argue for *two* representational types for word meanings: nonverbal and verbally-based. I differ from them (and from Deacon, 1997; Nelson, 2007) in hypothesizing that most children below ages 4.5–5.5 associate *most* spoken word files (“wordforms”) with nonverbal meanings (neurological organizations or “mental representations” or files), including:

- *Perceptual* (nouns—objects, settings, events).³ what we classify as “nonverbal” *nouns* are either physically perceived objects (including agents), settings, or events, or other perceptually established neurological files (representations) from senses. These perceptual processes are mainly located in the posterior cortex.
- *Procedural* (verbs—actions, tasks). What we classify as “nonverbal” *verbs* are action components (e.g., *grasp*), or more complex actions (e.g., *pounce*) or task programs (e.g., beaver

² Reading “comprehension” assessment is usually done using standardized tests (e.g., Iowa Test of Basic Skills) with questions about short texts involving both *facts* which mostly involve nonverbal meanings and *inferences* which often include verbally-based meanings.

³ Note that some wordforms access both perceptual (noun) and procedural (verb) meanings that are related. For example, “John went for a *swim*,” vs. “Let us *swim* across the pond.” This suggests that some meanings share related perceptual (noun-like) and procedural (verb-like) meanings. In speech, which meaning depends on word order, functors, and affix cues.

constructing a lodge) procedurally established as neurological files (representations). These procedural processes are mainly located in the anterior cortex but must have connections with the perceptual processes in the posterior cortex.

- *Modification* (adjectives or adverbs—discriminating between similar perceptions or procedures). What we classify as nonverbal *adjectives* or *adverbs* refer to specific features of perceptions (e.g., *big* vs. *small* objects) or features of procedures (e.g., *quickly* vs. *slowly*).⁴ I do not know where modification neurological files are likely to be located in the brain.
- *Functional* (functors—prepositions, conjunctions, articles, demonstratives, pronouns, and probably interrogatives). Words generally activate parts of both the posterior, temporal, and anterior cortices. Function words differed by *less* activation in the anterior cortex.

Neurological evidence for nonverbal meanings

Tomasello et al. (2017) report evidence of neurological activity for “Brain connections of words, perceptions and actions.” R. Tomasello’s evidence for neurological organization of nonverbal word meanings:

... some additional cortical areas contribute to semantic processing in a more selective fashion, being particularly relevant for specific semantic categories, such as words typically used to speak about animals, tools, or actions and their related concepts. Some evidence also indicates that when recognizing a word such as *run*, activity in motor cortex, and even more specifically in leg-motor cortex, emerges, whereas, when hearing an object- and visually-related word such as *sun*, activity in visual areas is relatively more pronounced (Damasio et al., 1996; Hauk et al., 2004; Boulenger et al., 2009; Pulvermüller et al., 2009; Gainotti, 2010).

In other words, many word meanings involve neurological centers for procedural processes (action-meanings) or perceptual processes (sense-related meanings).

Diaz and McCarthy (2009) studied neurological activations to “content” words (nouns, verbs, and adjectives) vs. “function” words.

Both word types strongly activated temporal-parietal posterior cortex, middle and anterior temporal cortex, inferior frontal gyrus (lobe or fold in cortex AB), parahippocampal gyrus (cortex associated with memory, AB), and orbital frontal cortex. Activations were more extensive in the left hemisphere. Content words elicited greater activation than function words in middle and anterior temporal cortex, a sub-region of orbital frontal cortex, and the parahippocampal region. Words also evoked

extensive deactivation, most notably in brain regions previously associated with working memory and attention.

In other words, the neurological correlates of words activated parts of the posterior and temporal cortices and the front of anterior cortex. “Content” words differed from function words by greater activation in various parts of the temporal and anterior cortex regions.

Verbally-based meanings

Around age 4.5 to 5.0 or later, children become able to acquire *some* verbally-based word meanings. As discussed by Deacon (1997) and Nelson (2007), verbally-based word meanings are constructed in part from other, often nonverbal word meanings. Deacon and Nelson refer to these as “symbolic” words. For example, the meaning of *height* involves understanding a dimension that ranges from *short* to *tall*. Following Case (1985), acquiring these verbally-based meanings involve increased working memory around age 5 making possible for children to attend to relations between nonverbal meanings such as *short* and *tall*. These also appear to be the same as Lawson-Adams and Dickinson’s (2020) “lexical” meanings (also what I consider to be verbally-based meanings). These verbally-based “word meanings” go beyond nonverbal *perceptual*, *procedural*, *modification*, and *functional/grammatical* meanings. Other examples include, simple verbally-based “dimensional” words such as *weight* (vs. *light* and *heavy*), or *social class* (vs. *rich* and *poor*).

Reggin et al. (2021) discuss variables affecting acquisition of abstract words. They summarize: “Thus, (emotional) valence, interoceptive strength, and mouth action strength facilitate acquisition of abstract words, and are less important for acquisition of concrete words, consistent with the predictions of the embodiment account. These results provide further evidence to support the claim that abstract words are grounded in emotional and associated sensorimotor experiences.” The role of linguistic experience—range of other words encountered with a target word—was also a factor. Reggin et al. state, “the current findings suggest that all words, both concrete and abstract, are learned earlier when they have been experienced in diverse contexts, presumably because their meanings have greater opportunities to be linked to already known words.”

Many basic “dimensional” words such as *height*, *weight*, *length*, and *size*, are considered “known” by grade four in the readability scale of Chall and Dale (1996). [However, many of these meanings are really only learned by 67%–80% of grade four children in their samples as reported in the *Living Word Vocabulary* of Dale and O’Rourke (1981).⁶]

Other examples of words with verbally-based meanings are: *peace*, *astronomy*, *democracy*, or *liberty*. These more “abstract” meanings cannot be referred to simple things we perceive (perceptual files), or simple actions or combinations of actions into tasks that we carry out (procedural files). Note that these verbally-based meanings are very rarely assessed in picture vocabulary tests. None of these

⁴ Note that early-acquired (before age 5) “modification” meanings are often “binary” (e.g., *tall* vs. *short*), as Brown (1973) suggested. Many other modification terms learned early are binary, but not all are.

⁵ Relational words (e.g., *tall* or *short*) which may be known by age 3.

⁶ Or 49%–70%, after adjusting for guessing.

verbally-based words appear in [Hart and Risley's \(1999\)](#) sample of words used by age 3.

In addition, after age 5, children's ability to comprehend and produce longer narratives increases substantially. Case and his colleagues have documented growth in working memory correlated with cognitive changes (including length of sentences and narratives) around these ages as have others ([Case, 1985](#); [Case and Okamoto, 1996](#); [Gathercole et al., 1997, 1999](#)).

The grammatical classification of verbally-based meanings include nouns, verbs, modifiers (adjectives and adverbs), and functors. Note that this classification parallels the functional classifications based on neurological functions that I have used to describe nonverbal meanings. These grammatical classes of words continue to be necessary to construct meaningful sentences as children develop.

One theory of word meanings is that verbally-based meanings are constructed from co-occurrences of a specific words and its environment of other words in paragraphs in various texts ([Landauer, 2007](#); [Biemiller et al., 2014](#)). Landauer and associates refer to these meanings as "Latent Semantic Meanings" ([Landauer, 2007](#)). Note that after age 5, children can supply verbal meanings for nonverbal meanings as well as for verbally-defined meanings. [Landauer \(2007, p 23\)](#) suggests that latent semantic meanings are acquired "...not all or none, but grows gradually, not showing itself outwardly until good enough and in the right context with the right measuring instrument." However, [Carey \(1978\)](#) notes that some words can be "fast-mapped"—acquired from one or two experiences. I suggest that "fast-mapping" may be limited to nonverbal meanings.

[Hadley and Mendez \(2021\)](#) give a good illustration of the difference in meaning types between nonverbal meanings and symbolic or lexical meanings:

Consider the difference between learning a word like "gunwale" vs. a word like "analyze." For "gunwale," prior knowledge or morphological strategies would be of little help, but a quick reference to a picture would reveal the word's meaning (the upper edge of a boat). However, one would rarely, if ever, encounter the word "gunwale" again, and might quickly forget it. To learn the word "analyze," a picture or even a definition might not be helpful, but hearing multiple examples and uses in context could build a working knowledge of the word ([Bolger et al., 2008](#)). Since "analyze" is used frequently across domains, a child's knowledge of this word would continue to grow as they encountered it over time. Moreover, learning this word gives a child a valuable tool in making meaning from text.

[Hadley et al. \(2021\)](#) have studied concrete vs. abstract words included in a vocabulary intervention study with preschool children (average age 4.9 years). I believe these are similar to the distinction I draw between nonverbal meanings (concrete) and verbally-based meanings (abstract). In Hadley et al.'s article, they instructed 20 words that were above the "easy" level reported by [Biemiller \(2010a,b\)](#). Their results show that four of the five the most learned words were "nonverbal meanings" (could be files of perceived or procedural neurological stored information, e.g., *liquid*, *sway*). Their average score was 0.86 out of 6 possible. A score of 1 refers to one type of correct response (e.g., synonym or perceptual feature). Among the other 15 words are a mix of 6 non-verbal and 9 verbally-based meanings. The nonverbal meanings are rarely-encountered words in preschool (e.g.,

glade, *struggle*). Most of these words were not well learned. The verbally-based meanings were also rarely encountered meanings (e.g., *strategy*, *meander*). The average score for these 15 instructed words was 0.34 out of 6 possible. (The test words and data have been supplied by E. Hadley, personal communication.) It appears that teaching rare words to preschoolers was not very useful.

More fine-grained verbally-defined word meanings

[Uccelli et al. \(2014\)](#) provide a more fine-grained description of advanced verbally-defined meanings—that they describe as Core Academic Language. "In contrast, CALS refers to language skills that cut across content areas and are used to fulfill similar language goals, such as communicating or understanding precise meanings, concisely packed information, and explicitly marked conceptual relations." Their work makes clear that this more advanced language continues to develop between fourth and eighth grades and is a dimension rather than a binary distinction between nonverbal vs. verbally-defined meaning types.

Neurological evidence for verbally-based meanings

[Papagno \(2022\)](#) have summarized research on neurological studies contrasting "concrete" vs. "abstract" word meanings. Their summary states,

The anatomical implication is that abstract concepts are represented entirely verbally, in the left hemisphere, whereas the representations of concrete concepts have both left hemisphere, verbal components, and right hemisphere, visuo-perceptual components." Specifically, activation of neurological activity with abstract (verbally-defined) words occur in the left middle longitudinal part of the temporal lobe (middle temporal gyrus AB) and the lower left bottom part of the anterior cortex (left inferior frontal gyrus). These brain areas relate to perceiving and producing words. Concrete (nonverbal) meanings activate these left brain areas more strongly, and also activate parts of the posterior cortex on both sides of the brain.

Becoming able to explain (define) word meanings

At about the same age (5 years) when children begin to acquire verbally-based meanings, children also become increasingly able to supply verbal "meanings" for words—especially when the requested word has been "illustrated" with a context sentence. Prior to age 5, children are very limited in verbally explaining word meanings, although they can pass "picture vocabulary" test items and follow verbal instructions—when they "know" the words in an instruction ([Meichenbaum and Biemiller, 1998](#); [Biemiller and Slonim, 2001](#)). By 2012, [Wechsler \(2012\)](#) and others no longer requested word meanings from children under 6 as part of language or intelligence assessments.

This reflects the fact that children under 5 often cannot supply a word meaning despite “comprehending” words.

Once children become able to create and understand verbally-defined word meanings, they may create verbal meanings for what have been nonverbal meanings.

However, our data (Biemiller and Slonim, 2001) indicates that before the end of grade two, children were limited in supplying word meanings. They appeared to acquire a great many word meanings during grade two in Biemiller and Slonim’s cross-sectional data. Biemiller and Slonim have suggested that this reflected a limitation in acquiring verbal meanings before age 7 or 8. Cross-sectional data in Biemiller and Slonim (p 509) suggested that an estimated average of about 3,000 root word meanings were added around grade two, whereas for other grades, an estimated average of about 1,000 root word meanings were acquired. This “leap forward” in grade two probably reflects (a) increased formation of “latent semantic meanings” extracted from repeated receptive uses of a word form in various oral and written contexts, plus (b) increased ability to verbally describe information in nonverbal perceptual or procedural representations. [However, when verbal definitions are *taught*, they can be learned before age 5 (Hadley et al., 2016).]

I expect that after age 5, once children become able to create and understand verbally-defined word meanings, they may add verbal meanings for what have been nonverbal meanings. This can occur either by extracting meaning from verbal contexts or by describing information stored in visual perceptual files. They are also acquired when explained verbally by another person (Biemiller, 2005). In addition, more than one meaning can be associated with a word-form—both nonverbal and verbally-defined meaning types (e.g., *strike*—to hit; *strike*—a discovery).

Methods

Research questions

1. Study 1. Are most or all of word meanings known at age 3 nonverbal?
2. Study 2. Do proportions of word meaning types show an increase in verbally-based meanings between age 5 and age 8?
3. Study 3. Do proportions of word meaning types known increase from kindergarten to grade five with attention to growth of verbally-based meanings between grade two (end of primary grades, age 8) and grade five (near end of upper elementary grades, age 11)?
4. Study 3. Are there differences in acquisition of numbers of nonverbal vs. numbers of verbally-defined word meanings for at/below vocabulary children vs. above median children for grades kindergarten to grade five?

Classifying word meanings as nonverbal or verbally-based meanings

Quite simply, I classify word meanings as *nonverbal* if the meaning *could* be stored mentally as a nonverbal neurological mental file in the perceptual representations (object or setting seen, smelled,

tasted, felt, heard, etc.), procedural representations (actions, tasks), nonverbal modificational representations (perceivable or procedural distinctions), or functional representations (relations between objects, settings, actions, time, as well as pronouns and interrogative terms).

I classify word meanings as *verbally-based* if the meaning could not be stored nonverbally without some further verbal qualification. This includes quantitative dimensions, such as *numbers*, *height*, or *distance*—and qualitative dimensions such as *social class*, or biological *phyla*. Similarly, procedures such as *compute*, *analyze*, or *perceive* are verbally-based.

Readers may find some word meanings that I have listed in subsequent tables that they would classify differently. I have included all classified words and specific LWV meanings in several appendices in this article [excepting about 90% of the 1,361 words listed by Hart and Risley, 1999, where I assume each word’s most common meaning as reported by Dale and O’Rourke, 1981].

Three studies to support nonverbal vs. verbally-based word meanings that children know: Study 1—age 3; Study 2—ages 5–8; and Study 3—second and fifth grades and some information re kindergarten to grade five—I also include a vocabulary median split by grade

In this article I will be describing data from 3 studies drawn from a vocabulary book and two published studies concerning children’s word use or knowledge at various ages and adding a classification of meanings acquired as nonverbal vs. verbally-based.

Study 1: The purpose of Study 1 was to determine the types of word meanings *used* by children by age 3. These data are from children in Hart and Risley’s (1999) book: *The Social World of Children Learning to Talk*. Children were observed from 19 months to 36 months.

Study 2: The purpose of Study 2 was to examine the types of word meanings vocabulary acquired at the end of pre-kindergarten (age 5) to grade two (age 8). These word meanings were sampled from grade levels 2 to 12 in Dale and O’Rourke’s (1981) *Living Word Vocabulary*. These data on knowledge of 59 word meanings were classified as 30 nonverbal meanings and 29 verbally-based meanings in Biemiller and Slonim’s (2001) study. (At the time of the study, we were not attending to non-verbal vs. verbally-defined meanings.) The study population were from advantaged pre-kindergarten to grade two children in a laboratory school. By “advantaged,” I mean a population of college-graduate parents including many professionals, at a university laboratory school.

Study 3: The main purpose of Study 3 was to contrast word type acquisition at grades two and five. An additional analysis involved splitting samples in each grade into children at/below vocabulary median vs. above median. These data include knowledge of 117 word meanings known at grade level two to grade 12 for representative English-speaking children in kindergarten through grade five (Biemiller and Slonim, 2001). This study involved a population that included low-income, average, and advantaged elementary students. Again, these word

TABLE 1 Words used by age 3 by most, some, or few by types and frequency (Study 1).

Meaning type	Words used by			All words
	Most	Some	Few	
	(100%–80% chil.)	(79%–40% chil.)	(39%–5% chil.)	
Nonverbal				
Perceptual	32% (60)	54% (145)	58% (492)	53% (697)
Procedural	30% (56)	20% (54)	24% (206)	24% (316)
Modificational	6% (12)	15% (42)	11% (94)	11% (148)
Functional	32% (61)	11% (29)	7% (63)	12% (153)
Total Nonverbal	100% (189)	100% (269)	100% (855)	100% (1,314)
Verbally-based ^a				
Perceptual	(0)	(0)	26% (5)	25% (5)
Procedural	(0)	(0)	(0)	(0)
Modificational	(0)	100% (1)	74% (14)	75% (15)
Functional	(0)	(0)	(0)	(0)
Total Verbally-based	(0)	100% (1)	100% (19)	100% (20)

Taken from Hart and Risley's list of words used by age 3 in their 1999 book. Their sample included 42 children. Root words (mono-morphemic) used by 2 or more children are included in these data. ^aMost possible verbally-based words were numbers over five (e.g., *three, thirty*). I interpret these as modifiers. The remainder were: *year, world, temperature, week, and heaven*. (I do not know how these words were understood by these 3 year old children.) ^bThere were 23 words that I could not classify. Some were used by many. Examples are *bye, no, ouch*. Others were used by few. Examples are *welcome, thanks, giddyup*. In addition, the Hart and Risley also includes a number of irregular past tense forms (e.g., *threwed, falled*). I did not count these. Some but not all also appeared in present tense in their [Appendix](#).

meanings were sampled from Dale and O'Rourke's *Living Word Vocabulary*.

Study 1: words used at age 3 (semi-representative sample)

The purpose of Study 1 was to determine the type of word meanings *used* by children by age 3. My expectation was that all or most words used would have *nonverbal* meanings.

Hart and Risley (1999) have a list of 1,361 root words reported used by age 3.0 by 2 or more children in a semi-representative sample of 42 children. (The list of words appeared in their 1999 book.) I analyzed the root words reported *used* by 2 or more children in Hart and Risley's book. The word meanings were classified as nonverbal vs. verbally-based, and by grammatical function (noun, verb, modifier, or functor) and by how many words were used by their children: most—80% or more; many—79%–40%; or few—39% or fewer. (The “least used” words reported were used by at least 2 children by age 3.)

Methods: data on words at age 3

Sample

This was a semi-representative sample of disadvantaged, working-class, and advantaged children in Kansas. Children were recruited from 3 groups: 13 children from a university laboratory school—parents mostly had college education; 23 children from working class families whose parents mostly had non-college education; and 6 less advantaged children in financially assisted homes (all Black). Study 1's sample of “disadvantaged” children were recruited from attendees at sessions of the Special Supplemental

Nutrition Program for Women, Children, and Infants (U.S. Federal). This yielded the 6 Black welfare-supported families (Hart and Risley, 1995).

The children were roughly “representative” of children in Kansas City, KS around 1990. Their study probably underestimates the proportion of disadvantaged students in the area and ignores the many disadvantaged nonblack population in Kansas City or at large in the U.S.

Procedure

Hart and Risley (1999) recorded all child and parent/caretaker language during monthly one-hour home observations. They reported 1,361 root words used by 2 or more children between 19 months and 36 months. Hart and Risley's [Appendix A](#) notes when words were first used by two or more of the children. [These root words are listed by Hart and Risley without reference to grammatical affixes, i.e., inflections. Very few semantic affixes (e.g., re- or -able) appear among these words. (A few irregular past tense forms were used.) (Meaning types classified by the author.)]

Results: word meanings at age 3

In Study 1, most of the words reported used had clearly *nonverbal* meanings. A substantial number of functors were among the words used early by many or some children.

Table 1 describes most of the 1,361 root words reported used by *most* of Hart and Risley's sample (80% or higher), *some* (40%–79%), or *few* (39% to 5%—a minimum of 2/42 by age 3 in their 1999 study). Almost all of these words have nonverbal meanings. (There were 28 words that I could not classify.)

These data indicate that about a third of Hart and Risley's 119 most-used words were functors and most of the rest of the most-used

words had perceptual or procedural meanings. Among the less well-known words, meanings were about half perceptual, a quarter procedural, and the rest divided between modificational or functional meanings.

Appendix 1 shows a sample of words used by 2 or more of 42 children. In Appendix 1, I have included a sample of 9% of the whole list, yielding a table about the same size as Appendices 4, 5. Among words used by age 3 were *water*, *milk* (perceptual), *come*, *say* (procedural), *open*, *wet* (modificational), and *here*, *they*, *with* (functional). These words were used by *most* children (80% or more of the children) by age 3. The words used by *some* (40%–79%) were similar—very “concrete” every-day kinds of meanings. Words used by *few* (39% or fewer) were also very similar. The 4 functors used by few were: *either*, *or*, *almost*, *as*. Readers can see examples in Appendix 1.

Of the 19/1,361 apparently verbally-based meanings, 14 were numbers from 6 to 60. The remaining words were *world*, *year*, *heaven*, *week*, and *temperature*. We cannot tell how well these words were understood without the associated context.

Study 2: word meanings known at preschool, kindergarten, grade one, and grade two (advantaged sample)

The main purpose of Study 2 was to examine vocabulary acquired by the end of pre-kindergarten (age 5) to the end of grade two (age 8). My hypothesis was that at age 5 there would be few or no verbally-based meanings known. By age 8, I expected that some verbally-based meanings would be known by some or most children. The youngest children were roughly 5 at the time of testing (spring, 1999).

The word meanings tested were sampled from an estimated grade two or lower to grade 12 as determined by Dale and O'Rourke (1981). My samples of the word meanings included 30 meanings that were at least possibly nonverbal and 29 that were verbally-based. Sampling to construct this list of meanings to test was not constrained by nonverbal vs. verbally-based type—that's simply what they turned out to be. (At the time these word meaning samples were constructed in 1998, we were not aware of their type: nonverbal vs. verbally-based meanings.)

Methods: data on word meanings known by children from ages 5 to 8

The 59 word meanings used in this study were from Biemiller and Slonim's (2001) Study 2.

Sample

The children in this study attended a university laboratory school. Children were mainly from middle or upper middle-class families. Most parents were college educated. The school did charge tuition. These children could be considered educationally “advantaged.” All children in the study grades were included. This study was conducted in the spring of 1999. The children in the pre-kindergarten group averaged about 5 at the time of testing. Children in the older grades were about 6, 7, or 8 years old. There were 22 children in the pre-kindergarten class, 20 in kindergarten, 22 in grade one, and 22 in grade two.

Choosing words

For our analysis of word meaning knowledge among elementary school children, we drew samples of 10 root word meanings from each of grade levels 2, 4, 6, 8, 10, and 12 from Dale and O'Rourke's, 1981 *Living Word Vocabulary* (LWV; Biemiller and Slonim, 2001). We found roughly 2,500–4,000 root word meanings at each LWV grade level (Biemiller, 2005). (One word meaning was dropped from analysis when the test sentence led to 2 different plausible meanings from children.)

Procedure

Individual oral testing, words were presented in context sentences, responses written down by examiner. In all test sessions, we would read a test sentence to a child, then ask “What does (word) mean?” The child's response would be written down. For example, “Use a knife to spread the jam. What does *spread* mean?”

Coding responses: (1—correct—child understands word meaning in sentence to follow a narrative including this sentence; 0.5—not sure whether child knows this meaning; 0.1—wrong response; 0.0—no response. Agreement between raters in a blind test was 90%).

Statistical analysis

Analyses of variance and correlations conducted with *Systat 13* (Systat, 2009).

Results: word meanings by type from ages 5 to 8 (laboratory school sample)

To summarize, before grade one, virtually no *verbally-based* words were known. Only *subtract* was known by some children in kindergarten. By the end of first grade, 4 *verbally-based* word meanings were known by some children, and by grade two, 5 (of 29 tested) *verbally-based* word meanings were known including one known by most (*subtract*). There were significant gains across grades, and significantly more *nonverbal* than *verbally-based* meanings were known. The numbers of nonverbal and verbally-based meanings were substantially correlated for students in each grade.

Percent correct

In Table 2, percentages of the 59 word meanings known by meaning type sampled from Dale and O'Rourke's *Living Word Vocabulary* are shown for pre-kindergarten to grade two. Both types (nonverbal vs. verbally-based meanings) and grades varied

TABLE 2 Percentages of nonverbal and verbally-based meanings by grade and correlation by type (Study 2).

Grade	N	Type of meanings		Correlation nonverbal × verbally-based
		Nonverbal (30 meanings)	Verbally-based (29 meanings)	
Pre-Kdgn.	22	23%	2%	0.73
Kindergarten	20	40%	6%	0.56
Grade one	22	49%	13%	0.70
Grade Two	22	65%	24%	0.75
Total	86	44%	12%	0.80

TABLE 3 Word meaning types by grade and frequency known (sample of 59 word meanings from advantaged classes, Study 2).

Grade and age	Nonverbal meanings known in four grades			
	Most (100%–80%)	Some (79%–40%)	Few (39%–0%)	All words
Pre-kind. (ca 5 years)	0% (0)	27% (8)	73% (22)	100% (30)
Kindergarten (ca 6)	0% (0)	30% (9)	70% (21)	100% (30)
Grade one (ca 7)	20% (6)	43% (13)	37% (11)	100% (30)
Grade two (ca 8)	37% (11)	43% (13)	20% (6)	100% (30)

Word type	Verbally-based meanings known by four grades			
	Most (100%–80%)	Some (79%–40%)	Few (39%–0%)	All words
Pre-kind. (ca 5 years)	0% (0)	0% (0)	100% (29)	100% (29)
Kindergarten (ca 6)	0% (0)	3% (1)	97% (28)	100% (29)
Grade one (ca 7)	0% (0)	14% (4)	86% (25)	100% (29)
Grade two (ca 8)	7% (2)	17% (5)	76% (22)	100% (29)

significantly at $p < 0.001$ level (Type: $F = 688.2$, 1 and 82 df, repeated measure; Grade: $F = 36.4$, 3 and 82 df). More nonverbal meanings were known, and more meanings were known at older grades. Type and grade interacted significantly, reflecting the increasing percentage of nonverbal word meanings relative to verbally-based meanings in higher grades, as seen in Table 2 (type $\times$ grade: $F = 10.2$, 3 and 82 df, also $p < 0.001$).

Correlations for children’s percentages of nonverbal words and their percentages of verbally-based meanings within each grade

Within 3 of the 4 grades, the correlation was slightly above $r = 0.70$ or 50% of common variance (also shown in Table 2).

Details of how widely word meanings are known

The levels of nonverbal and verbally-based word meanings known by any children at these ages in Table 3 proved to be almost entirely nonverbal at ages 5 and 6, while some verbally-based meanings were known by some children at ages 7 and 8. Many fewer word meanings were known at ages 5 and 6 than at ages 7 and 8 (based on supplying verbal explanations of meanings).

Word meanings tested and passed are shown for pre-kindergarten (Appendices 2A,B) and grade two (Appendices 3A,B).

Study 3: word meanings known at kindergarten and grades one, two, four, and five (semi-representative sample)

The main purpose of Study 3 was to contrast word type acquisition at grades two and five. These data are taken from Biemiller and Slonim’s (2001) Studies 1 and 3, working with representative English-speaking children. Also included are percentages of meanings (both meaning types) known by children in kindergarten through grade five. I expected was that there would be substantial growth in verbally-based meanings by grade five. This implies greater student attention to the verbally-based meanings acquired between ages 8 and 11, and more ability to comprehend language with verbally-based meanings.

I also analyzed word acquisition of different word types for children with lower-sized vocabularies vs. larger sized vocabularies (median split).

Methods: data on word meanings known from kindergarten to grade five

Sample

For Study 3, two samples of representative English-speaking children in 3 public schools ranging from assisted housing to upper middle-class neighborhoods in a small Canadian city. All children were Caucasian. Combining these 2001 studies, there were 47 kindergarten, 43 grade one, 51 grade two, 43 grade four, and 44 grade five students. (Grade three was omitted because of Provincial tests at that grade level.) Being limited to English-speaking Caucasian children, this study is only representative of a limited population.

Choosing words

117 words—includes the previous 59 words used in Biemiller and Slonim’s Study 2 (which used the same selection method used in their Study 1) and an additional 58 words in their Study 3. Choosing words are as described in the Study 2 above. Two additional words were omitted from analysis because students gave two plausible definitions for the test sentences for these words.

Procedure

Same as Study 2.

Statistical analysis

Analyses of variance and correlations conducted with *Systat 13* (Systat, 2009).

Results: word meanings known by children by type from kindergarten to grade five (representative sample)

In Study 3 there were the same significant gains in word meanings at higher grades, and more nonverbal than verbally-based meanings.

TABLE 4 Percentages of nonverbal and verbally-based meanings by grade and correlation of types (sample of 117 word meanings from representative classes, Study 3).

Grade	N	Type of meanings		
		Nonverbal (62 meanings)	Verbally-based (55 meanings)	Correlation nonverbal x verbally-based
Kindergarten	47	26%	4%	0.73
Grade one	43	30%	7%	0.47
Grade two	51	52%	19%	0.69
Grade four ^a	43	59%	30%	0.70
Grade five	44	68%	37%	0.63
Total	228	47%	19%	0.85

^aGrade three omitted due to provincial testing.

TABLE 5 Word meaning types by grades two and five and frequency known (sample of 117 word meanings from representative classes, Study 3).

Grade	Nonverbal meanings known in grades two and five					All words
	(Age)	N	Most (100%–80%)	Some (79%–40%)	Few (39%–0%)	
Grade two	(ca 8)	51	18% (11)	50% (31)	32% (20)	100% (62)
Grade five	(ca 11)	44	50% (31)	31% (19)	19% (12)	100% (62)

Word type	Verbally-based meanings known in grades two and five					All words
	(Age)	N	Most (100%–80%)	Some (79%–40%)	Few (39%–0%)	
Grade two	(ca 8)	51	2% (1)	20% (11)	78% (43)	100% (55)
Grade five	(ca 11)	44	16% (9)	31% (17)	53% (29)	100% (55)

In this representative population, the number of nonverbal meanings known was again substantially correlated with the number of verbally-based meanings known.

Percentage correct

Not surprisingly, more words were known at older grades, and more nonverbal meanings were known than verbally-based meanings (Table 4). ANOVA indicated significant difference between nonverbal vs. verbally-based meanings ($F=1,644.8$; 1 and 223 df, repeated measure) and Grades ($F=118.3$; 4 and 223 df). The interaction of Type and Grades was also significant ($F=10.2$; 4 and 223 df). This interaction reflects the fact that nonverbal meanings Type difference were larger (around 30% from grade two up, while the difference was 23% in kindergarten and grade one).

Correlations between nonverbal and verbally-based meanings

As in Study 2, there was a high correlation in Table 4 between the percentages of nonverbal and verbally-based meanings across grades ($r=0.85$, $N=228$). Within grades, correlations between nonverbal and verbally-based meanings for 3 of the 5 grades were around half of common variance (Table 4). The other grades were a bit lower. Thus, children who have acquired more nonverbal meanings apparently also tended to acquire more verbally-based meanings.

Details of vocabulary at grades 2 and 5

The most important finding is that there was the predicted increase in knowledge of verbally-based meanings between grade two and grade five as shown in Table 5. The percentage of verbally-based

meanings known by many or some students rose from 22% in grade two to 47% of 55 verbally-based meanings in grade five.

Grade two children

Table 5 describes the grade two distribution of the types of words and how well known of the 117 words sampled from LWV grade levels 2 to 12 (Table 5 data and Appendices 4A,B—words and LWV meanings by word frequency and grammatical categories).

In grade two, of the sample of meanings known well (80% or more), all were nonverbal except one (11 of 12). The majority of the meanings known by some (79%–40%) of the grade two students of the total pool of 117 meanings were also nonverbal (31) vs. verbally-based (11). However, among the 63 meanings known by few (less than 40%) were 20 nonverbal and 43 verbally-based meanings.

Appendices 4A,B (nonverbal and verbally-based) list all 117 words and meanings with LWV levels by grammatical category and frequency for grade two.

By grade two, about 77% of all nonverbal word meanings in our sample of LWV grade level 2–12 meanings were known by most or some children. Among sample words known well continued to be mostly nonverbal (Appendices 4A,B). Perceptual meanings (nouns) known well included objects (*fish*) but also events (*flood*) and settings (*café*). There were few verbs (procedural meanings) known well—only *spread*, *stab*, and *listen* in our study. There was 1 grade two well-known functional meaning—*near*. The only grade two well-known verbally-based word meaning was *math*.

The relatively few verbally-based word meanings (22% of 55 words) were known by some or most. There some of each grammatical category.

TABLE 6 Percentage of word meanings known by at/below median vs. above median and by grade (sample of 117 word meanings, Study 3; SDs in parentheses).

Grade median split	N	Nonverbal	Verbally-defined	All
Kindergarten				
Low	23	17% (6)	0% (1)	7% (3)
Hi	24	35% (8)	7% (4)	22% (6)
Grade one				
Low	23	21% (9)	3% (4)	12% (4)
Hi	21	39% (7)	11% (4)	25% (4)
Grade two				
Low	27	42% (13)	13% (6)	27% (9)
Hi	24	64% (8)	26% (8)	45% (6)
Grade four				
Low	23	49% (9)	23% (10)	36% (8)
Hi	20	71% (7)	39% (9)	55% (7)
Grade five				
Low	22	61% (7)	29% (9)	45% (6)
Hi	22	75% (8)	45% (9)	60% (6)

Grade five children

Table 5 describes the grade five distribution of the types of known words and how well known are the 117 words sampled from LWV grade levels 2 to 12. Overall, there were 62 nonverbal meanings and 55 verbally-based meanings (Table 5 data and Appendices 5A,B—words and LWV meanings by word frequency and grammatical categories).

In grade five, the majority of the word meanings known by some or most were still nonverbal (50). There were 26 verbally-based meanings known by some or most. There were still 12 nonverbal meanings known by few, and 29 verbally-based meanings.

While the majority of new “well-known” words were nonverbal meanings, about a quarter were verbally-based meanings, including *justice*, *subtract*, and *secure* (Appendices 5A,B). All but one of these “well-known” words had been known by “some” in grade two. The additions to verbally-based word meanings “known by some” included *period* (a time in history), *nation*, and *former* (the first of two). These had been “known by few” in grade two. However, in grade five there remained more than half of our sample of verbally-based word meanings known by few or none. Almost all of these little-known word meanings were assigned to Dale and O’Roarke’s LWV grade levels 8, 10, or 12.

Differences between those with larger vocabularies vs smaller vocabularies for nonverbal and verbally-defined words

For this analysis, I identified children with above-median vocabulary vs. those at/below median vocabulary. Results are shown in Table 6.

The main effects of Word Type and Grade were significant. Type: $F=2,041.76$ (d.f. 1, 219, $p<0.001$; repeated measure). Grade: $F=317.12$ (d.f. 4, 219, $p<0.001$). The main effect of Vocabulary median split was constructed so not a meaningful effect. However, interactions with other variables are relevant.

There are differences between the high vocabulary groups and at/below vocabulary groups in nonverbal vs. verbally-defined words in different grades. 3-way interaction: $F=4.01$ (d.f. 4, 219). The two-way interactions were also significant.

The practical effect was that the difference in *gains* between grade two vs. grade five the at/below median children vs. above median children is not large. In grade five, the gains from grade two were 19 percent point gains for nonverbal meanings and 16 verbally-defined point gains for at/below-median grade fives vs. 11 nonverbal and 19 point verbally-defined gains for above-median grade fives. In other words, if the low-vocabulary children had learned as many words as the high-vocabulary children by grade two, it is possible that they could have been able to build similar vocabularies between grades three and five.

Discussion

To summarize

In Study 1, words “used” by age 3 had mostly nonverbal meanings—perceivable, procedural, modificational, or functional (98.5% of 1,331 words classifiable reported used by at least 2/42 children in Hart and Risley, 1999).

In Study 2, we wanted to look at word meanings acquired from age 5 up to age 8 (between pre-kindergarten to grade two). Before grade one, virtually none of the 29 verbally-based tested words were known. Only *subtract* was known by some children in kindergarten. By the end of first grade, 4 verbally-based word meanings were known by some children, and by grade two, 5 of 29 tested verbally-based word meanings were known by some children including one known by most (*subtract*). There were significant gains across grades, and significantly more nonverbal than verbally-based meanings were known. The number of nonverbal and verbally-based meanings were substantially correlated for students in each grade.

In Study 3, Biemiller and Slonim used a larger sample of word meanings from the *Living Word Vocabulary*—117 word meanings—conducted with representative English-speaking students. In grade two, most or some knew 42 of 62 nonverbal meanings, and 12 of 55 verbally-based meanings. By grade five, most or some knew 50 of 62 nonverbal meanings, and 26 of 55 verbally-based meanings—large increases. In short, relatively few verbally-based meanings were acquired by grade two, but many by grade five.

Contrasting children at/below median vocabulary vs. those above median vocabulary showed that size of verbally-defined vocabulary acquired by grade two was larger for above-median children. The size of verbally-defined vocabulary gains between grades two and five were similar (16 points for at/below-median vs. 19 points for above-median). By grade five, the at/below median children had acquired similar proportions of vocabulary as the above-median group in grade two.

Studies 1 and 3 probably had under-represented disadvantaged samples. Study 2 had all advantaged children.

Study 3's sample was reflective of a small city in Ontario, which at the time (1998) had very few non-White families.

What early-acquired vocabulary is important for later comprehension?

From Study 1

To build a larger vocabulary by age 3, I hypothesize that parents and caretakers should probably use more words with nonverbal meanings (illustrated concretely). Introducing modificational words probably requires two or more examples that illustrate differences in the modificational characteristics (e.g., contrasting a *small* ball vs. a *large* ball.). A wider range of functors could be introduced and used. Hart and Risley (1995) reported that vocabulary size “used” by age 3 was predictive of vocabulary size by grade three.

From Study 2

In Study 2, I hypothesize that relatively few verbally-based meanings should be addressed by teachers in kindergarten, having found that few verbally-based meanings are learned at age 5. Hadley et al. (2021) also reported difficulty in teaching advanced word meanings to young children. Teaching simple quantitative dimensions like *height* and *weight*, *numbers* to 100, and functors could be helpful. However, by grade one, more verbally-based meanings could be introduced and used. In addition, it is likely that less advantaged children acquire fewer functors by age 4 or 5. (Study 2 only included advantaged children.) These functors may be necessary to facilitate acquiring other verbally-based word meanings.

From Study 3

These verbally-based meanings known by *some* children by the end of grade two may be priority meanings for the primary grades—to be addressed as they would be helpful to classify occur in texts or instruction during the primary grades from kindergarten to grade two. In the future, it would be helpful to classify all the primary priority word meanings in *Words Worth Teaching* by type (Biemiller, 2010a). This should yield a smaller group of mostly verbally-defined priority words for the primary years.

Findings from the analysis of at/below median-*vs* above-median vocabulary children within grades indicated that the size of *gains* of verbally-defined meanings from grades two to five were similar.

Implications of the correlation between nonverbal and verbally-based meanings

The strong correlation between percentages of nonverbal and verbally-based meanings is why the PPVT and other picture-based vocabulary assessments in kindergarten or grade one could predict subsequent vocabulary including verbally-based meanings. Several studies show that a larger PPVT vocabulary in kindergarten is predictive of reading comprehension—especially at grade three and later. Would simply expanding the number of nonverbal meanings both prior to and after age 5 ensure the development of more advanced verbally-based language? Probably, but I recommend concentrating on the words that have been acquired by those with larger vocabularies.

Fostering verbally-based meanings from grade two on

McKeown (2019) has summarized methods of fostering abstract (verbally-defined) meanings in elementary and middle school. She lists:

1. There are many aspects to know about a word, including features of its meaning, situations in which it is used, associations with other words, and how it behaves syntactically in context.
2. Words are polysemous; their meanings are not static but shift according to context. These shifts may be large or subtle; for example, *accommodate* can mean physically providing room for someone and providing for someone's need or request, or it can take a more metaphorical sense of being able to understand a new idea that may challenge your perspective.
3. Word knowledge is incremental, gradually developing over multiple encounters.

McKeown (2019) notes that many abstract words are polysemous—have several related meanings. Teacher should use multiple examples from which teachers and students could extract the underlying common meaning. (There are also words with completely unrelated meanings, e.g., *fast* as in “speedy” vs. *fast* in “not eating for a period.” McKeown suggests not teaching these at the same time.)

Some early-acquired *verbally-based* meanings are “basic dimensional meanings” (Case, 1985)—e.g., *height*, *weight*, *speed*, *time*, *color*, *size*, *number*, *social class*, etc. I hypothesize that instruction can include brief teacher explanations offered during reading texts, teacher answers to children's questions (also peers' answers), in-text explanations, and more formal vocabulary instruction. This method proved effective with primary grade children in Biemiller and Boote (2006). In addition to quantitative dimensions, examples of verbally-based word meanings could include words like: *justice*, *react*, *secure* (free from fear), and *through*.

Acquiring verbally-defined meanings can occur through teacher explanation or through children's questions or self-constructed word meanings (Biemiller, 2005). Children's questions or self-meaning construction requires “word consciousness” (Graves,

2006) or word awareness. (A child cannot address an unknown meaning if they do not attend to unknown words.) In preliminary research, I found that sensitivity to unknown words could be fostered. [This preliminary work is summarized in Biemiller (2010a).]

Priority words

In *Words Worth Teaching*, I defined “priority words” for instruction as those known by children with large vocabularies but often not by those with smaller vocabularies at the same age (Biemiller, 2010a). I now think that from grade two on that the verbally-based “priority” word meanings may be more important for instruction than “priority” words with nonverbal meanings. Nonverbal meanings can be learned easily (if objects, pictures or actions are physically present or in texts), while verbally-based meanings probably require more direct instruction.

The finding that low-vocabulary students have similar *gains* between grades two and five, suggests that if the low-vocabulary children before grade two might avoid “low-vocabulary” by grade five. But this depends on whether they could learn as many word meanings (mostly nonverbal) by grade two. If so, it is possible that they could attain grade five vocabulary levels now reached by the high-vocabulary grade fives. This remains to be proven—both that it is possible to increase sufficient vocabulary before grade two, and that if that could be done, that most children could reach grade five with adequate levels for continued academic success.

Both acquiring a larger nonverbal vocabulary by age 5—and continuing exposure to “priority” nonverbal vocabulary after age 5 are necessary, in addition to fostering verbally-based meanings from kindergarten on.

Re word priorities

Hadley and Mendez (2021) wrote:

Researchers prioritized unknown words and words described as Tier 2 when selecting words. The most frequently selected word type was concrete nouns; more than half of all words selected were concrete in nature (nonverbal meanings—AB). A small percentage of target words appeared on published lists of words recommended for instruction. Word selection varied by context, with different types of words selected for science content vs. narrative/fiction. Finally, there were few relationships between child characteristics and word characteristics.

As Hadley and Mendez comment, these lists of words for instruction focus too much on nonverbal meanings—that are probably learned easily. They need to be introduced and may require brief explanation. However, clarifying or explaining verbally-based meanings will require more instruction. Furthermore, it may be necessary for curriculum developers and teachers to find or construct texts that allow context and attention for needed verbally-based meanings. This should be researched in future curriculum studies.

Lawson-Adams and Dickinson (2020) emphasize teaching word meanings in the elementary years, noting that nonverbal referents (e.g., pictures, short videos, action, and events) may be needed for nonverbal meanings, and possibly also with some verbally-based meanings.

Hadley et al. (2021, p 9) also share the hypothesis that adding nonverbal word meanings is “easier” (imageable) since “abstract” words are very hard to learn for low-vocabulary children at age 6.0 and below. Concentrating on nonverbal meanings for younger students may be a better use of instruction time. Hadley et al. observe that “In other words, the Matthew effect might be due to inadequate explicit instruction, but it may also be attributable to teaching words that are too abstract and disconnected from the conceptual knowledge of beginning language learners.”

I have similar hypotheses about vocabulary interventions with children age 6.0 and younger. The difference is that I hypothesize that these meanings involve direct developmental neurological representation as perceptual, procedural, or functional files (representations). As reported in the present research, acquisition of verbally-based files (representations) only begin to appear at age 6–7 (grade 1) and more at age 7–8 (grade 2).

Additional observations regarding neurological evidence for word meaning modalities

As noted previously, Tomasello et al. (2017) review neurological research on activation of cortex regions used when word meanings activating procedural areas and perceptual areas in conjunction with auditory words. In addition, the R. Tomasello et al. study reports computer simulations of processing perceptual and procedural words.

The word learning simulations documented the spontaneous emergence of word/symbol-specific, tightly interconnected cell assemblies within the larger networks, each binding articulatory-acoustic wordforms to sensorimotor semantic information. Due to network structure, connectivity, and Hebbian associative learning, which maps neuronal correlations, the emerging ‘semantic circuits’ for object-and action related words (Italics added, AB) exhibited category-specificity primarily in modality preferential areas; the “higher” multimodal connector hub areas central to the network architecture showed only moderate category-specificity.

I look forward to seeing further research on direct neurological evidence on the relationship between both between nonverbal meanings and neurological representations, and between verbally-based meanings and neurological meanings. Verbally-based meanings are not as neurologically localized as the nonverbal meanings per Papagno (2022).

Limitations

The samples used for these studies are only moderately representative of English-speaking children in N. America. Both Hart and Risley (1999) and Biemiller and Slonim (2001) are not representative of diverse populations in N. America, although both researchers attempted to incorporate children at varying levels of socio-economic status.

I have not attempted to create a more refined dimension of word meaning type, as Uccelli et al. (2014) suggest. Just as the present study

has demonstrated developmental changes in word meanings acquired, a more refined system would probably provide additional educationally-relevant developmental trends.

Most of the findings in this research is correlational—observations of associations between age (studied cross-sectionally) and word types acquired. These findings are not *proof* of what should be taught at various ages. Future research should include research on instruction of verbally-defined word meanings, and research on the longitudinal effects of expanding knowledge of nonverbal word meanings for children age 6 and under for their language and literacy in later grades.

Conclusion

I hypothesize that the type of words to be taught or fostered for children under age 6 should be mostly what I call nonverbal word meanings. The median-split analysis in Study 3 suggests that building stronger nonverbal vocabularies before grade two could be very beneficial. After age 6, and especially by grade two, the word meanings to be taught or fostered would be mainly verbally-based meanings. I expect that some new nonverbal meanings should also be present and addressed in and after grade two but will require less instructional efforts. These predicted beneficial effects of instruction on vocabulary remain to be proven. Note that most of these data were obtained with White children from a range of socio-economic families.

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

I am referring to the studies conducted in the 1980's and 1990's. I have recoded vocabulary data from these studies. I assume that the Hart and Risley research was conducted with approval at the University of Kansas. My research with Naomi Slonim was approved by the University of Toronto, Laboratory School at Toronto, and the

participating school board and parents in a southern Ontario city. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

AB: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1280568/full#supplementary-material>

References

- Biemiller, A. (2005). "Size and sequence in vocabulary development: implications for choosing words for primary grade vocabulary instruction" in *Teaching and learning vocabulary: Bringing research to practice*. eds. E. H. Hiebert and M. L. Kamil (Mahwah, NJ: Lawrence Erlbaum Associates), 223–242.
- Biemiller, A. (2010a). *Words worth teaching*. Columbus, OH: McGraw-Hill Book Co.
- Biemiller, A. (2010b). "Vocabulary development and implications for reading problems" in *Handbook of research on Reading disabilities*. eds. A. McGill-Franzen and R. Allington (New York, NY: Routledge), 208–218.
- Biemiller, A., and Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *J. Educ. Psychol.* 98, 44–62. Not cited. doi: 10.1037/0022-0663.98.1.44
- Biemiller, A., Rosenstein, M., Sparks, R., Landauer, T. K., and Foltz, P. (2014). Models of vocabulary acquisition: direct tests and text-derived simulations of vocabulary growth. *Sci. Stud. Read.* 18, 130–154. doi: 10.1080/10888438.2013.821992
- Biemiller, A., and Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: evidence for a common sequence of vocabulary acquisition. *J. Educ. Psychol.* 93, 498–520. doi: 10.1037/0022-0663.93.3.498
- Bolger, D. J., Balass, M., Landen, E., and Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: an instance-based learning approach. *Discourse Process.* 45, 122–159. doi: 10.1080/01638530701792826
- Boulenger, V., Hauk, O., and Pulvermüller, F. (2009). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cereb. Cortex* 19, 1905–1914. doi: 10.1093/cercor/bhn217
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Carey, S. (1978). "The child as word learner" in *Linguistic theory and psychological reality*. eds. M. Halle, J. Bresnan and G. Miller (Cambridge, MA: MIT Press), 264–293.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. New York, NY: Academic Press.
- Case, R., and Okamoto, Y. (1996). The role of central conceptual structures in the development of Children's thought. *Monogr. Soc. Res. Child Dev.* 61:i. doi: 10.2307/1166077
- Chall, J. E., and Dale, E. (1996). *Readability revisited: The new Dale-Chall readability scale*. Cambridge, MA: Brookline Books.

- Clark, J. M., and Paivio, A. (1991). Dual coding theory and education. *Educ. Psychol. Rev.* 3, 149–210. doi: 10.1007/BF01320076
- Cunningham, A. E., and Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Dev. Psychol.* 33, 934–945. doi: 10.1037/0012-1649.33.6.934
- Dale, E., and O'Rourke, J. (1981). *The living word vocabulary*. Chicago, Ill.: World book/Childcraft International.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., and Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature* 380:1996.
- Deacon, T. R. (1997). *The symbolic species: The co-evolution of language and the brain*. New York, NY: W. W. Norton.
- Diaz, M. T., and McCarthy, J. (2009). A comparison of brain activity evoked by single content and function words: an fMRI investigation of implicit word processing. *Brain Res.* 1282, 38–49. doi: 10.1016/j.brainres.2009.05.043
- Dickinson, D. K., and Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Dev.* 82, 870–886. doi: 10.1111/j.1467-8624.2011.01576.x
- Gainotti, G. (2010). The influence of anatomical locus of lesion and of gender-related familiarity factors in category-specific semantic disorders for animals, fruits and vegetables: a review of single-case studies. *Cortex* 46, 1072–1087. doi: 10.1016/j.cortex.2010.04.002
- Gathercole, S. E., Hitch, G. J., Service, E., and Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Dev. Psychol.* 33, 966–979. doi: 10.1037/0012-1649.33.6.966
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., and Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Appl. Cogn. Psychol.* 13, 65–77. doi: 10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York, NY: Appleton-Century-Crofts.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Graves, M. F. (2006). *The vocabulary book*. New York, NY: Teachers College, Columbia University.
- Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Eunsook, K., Hirsch-Pasek, R., and Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Appl. Dev. Psychol.* 77:101332. doi: 10.1016/j.appdev.2021.101332
- Hadley, E. B., Dickinson, D. K., Hirsch-Pasek, K., Golinkoff, R. M., and Nesbitt, K. T. (2016). Examining the Acquisition of Vocabulary Knowledge Depth among Preschool Students. *Read. Res. Q.* 51, 181–198. doi: 10.1002/rrq.130
- Hadley, E., and Mendez, K. (2021). A systematic review of word selection in early childhood vocabulary instruction. *Early Child Res. Q.* 54, 44–59. doi: 10.1016/j.ecresq.2020.07.010
- Hart, B., and Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes, Publisher.
- Hart, B., and Risley, T. R. (1999). *The social world of children learning to talk*. Baltimore, MD: Paul H. Brookes, Publisher.
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/s0896-6273(03)00838-9
- Landauer, T. K. (2007). "LSA as a theory of meaning" in *Handbook of latent semantic analysis*. eds. T. K. Landauer, D. S. McNamara, S. Dennis and W. Kintsch (Mahwah, NJ: Lawrence Erlbaum Associates, Publishers), 3–34.
- Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral mechanisms in behavior*, Ed. L. Jeffress. New York, NY: Wiley, 112–137.
- Lawson-Adams, J., and Dickinson, D. K. (2020). Building lexical representations with nonverbal supports. *Read. Res. Q.* 56:603. doi: 10.1002/rrq.326
- Lorenz, K. Z. (1981). *The foundations of ethology*. New York, NY: Springer-Verlag.
- Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books Inc.
- McKeown, M. G. (2019). Effective vocabulary instruction fosters knowing words, using words, and understanding how words work. *Lang. Speech Hear. Serv. Sch.* 50, 466–476. doi: 10.1044/2019_LSHSS-VOIA-18-0126
- Meichenbaum, D., and Biemiller, A. (1998). *Nurturing Independent Learners*. Cambridge, MA: Brookline Books.
- Nelson, K. (2007). *Young minds in social worlds: Experience, meaning, and memory*. Cambridge, MA: Harvard University Press.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York, NY: Oxford University Press.
- Papagno, C. (2022). "The neural correlates of abstract and concrete words" in *Handbook of clinical neurology | the temporal lobe*. eds. I. G. Micelli, P. Bartolomeo and V. Navarro, vol. 187 (Amsterdam, Netherlands: Elsevier/ScienceDirect), 263–275.
- Pinel, J. P., and Edwards, M. (2008). *A colorful introduction to the anatomy of the human brain*. Boston, MA: Pearson.
- Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., and Nimmo-Smith, I. (2009). Distributed cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Hum. Brain Mapp.* 30, 3837–3850. doi: 10.1002/hbm.20811
- Reggin, L. D., Muraki, E. J., and Pexman, P. M. (2021). Development of abstract word knowledge. *Front. Psychol.* 12:478. doi: 10.3389/fpsyg.2021.686478
- Scarborough, H. S. (2001). "Connecting early language and literacy to reading (dis) abilities: evidence, theory, and practice" in *Handbook of early literacy research*. eds. S. B. Neuman and D. Dickinson (New York, NY: Guilford), 97–110.
- Systat (2009). *Systat 13*. Chicago, IL: Systat, Inc.
- Tomasello, R., Garagnania, M., Wennekers, T., and Pulvermüller, F. (2017). Brain connections of words, perceptions and actions: a neurobiological model of spatio-temporal semantic activation in the human cortex. *Neuropsychologia* 98, 111–129. doi: 10.1016/j.neuropsychologia.2016.07.004
- Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., and Sa'nchez, E. (2014). Core academic language skills: an expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Appl. Psycholinguist.* 33:1400006X. doi: 10.1017/S014271641400006X
- Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence 4th Edn*. London, UK: Pearson.



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Wahyu Widada,
University of Bengkulu, Indonesia
Yousef Wardat,
Higher Colleges of Technology,
United Arab Emirates

*CORRESPONDENCE

Oscar Vivanco-Galván
✉ oavivanco@utpl.edu.ec

RECEIVED 02 January 2024

ACCEPTED 03 April 2024

PUBLISHED 17 April 2024

CITATION

Jiménez-Gaona Y and
Vivanco-Galván O (2024) Biotechnology
project-based learning encourages learning
and mathematics application.
Front. Educ. 9:1364640.
doi: 10.3389/feduc.2024.1364640

COPYRIGHT

© 2024 Jiménez-Gaona and Vivanco-Galván.
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.

Biotechnology project-based learning encourages learning and mathematics application

Yuliana Jiménez-Gaona¹ and Oscar Vivanco-Galván^{2*}

¹Departamento de Química, Facultad de Ciencias Exactas y Naturales, Universidad Técnica Particular de Loja, San Cayetano, Loja, Ecuador, ²Departamento de Ciencias Biológicas y Agropecuarias, Facultad de Ciencias Exactas y Naturales, Universidad Técnica Particular de Loja, San Cayetano, Loja, Ecuador

Project-based learning (PBL) is a promising approach to enhance mathematics learning concepts in higher education. Here, teachers provide guidance and support to PBL implementation. The objective of this study was to develop PBL-based biotechnological projects as a strategy for mathematics learning. The methodology design was applied to 111 university students from Biochemical, Chemical Engineering and Business Administration careers. Knowledge, skills, perceptions, and engagement were measured through questionnaires, workshops, rubrics, and survey instruments. As a result, the paired comparison between tests, questionnaires and project shows significant differences ($p < 0.001$) between the experimental group and the control group. It is concluded that the teaching of mathematics should be oriented to the development of competencies, abilities, and skills that allow students to generate real solutions and broaden their vision of the applicability of their knowledge using new learning strategies.

KEYWORDS

mathematical models, biotechnology, project based learning (PBL), learning, learning mathematics activities

1 Introduction

Science, technology, engineering, and mathematics (STEM) education has become a crucial topic both inside and outside of school (Han et al., 2015). Currently, mathematics learning tends to be oriented towards textbooks, and students can only work on math problems based on what the teacher exemplifies; however, if given different context-based problems, they will have difficulty solving them (Fisher et al., 2020). Likewise, the traditional classroom model does not encourage student's interest in STEM (Sahin, 2009).

The research gap is between what students learn at the university and what they really need in the workplace (Holmes et al., 2015). Higher education institutions have been trying to provide students with both (i) hard skills, such as cognitive knowledge and professional skills (Vogler et al., 2018), and (ii) soft skills, such as problem-solving and teamwork (Lennox and Roos, 2017). However, these skills are difficult to achieve through traditional learning.

One learning that creates an active, collaborative atmosphere, and can increase self-confidence in students is Project-based learning (PBL) (Cruz et al., 2022; Guo et al., 2022; Markula and Aksela, 2022). The PBL method is applied as a teaching model that involves deep study and research on real problems (Asija, 2011). PBL is a model applied in most learning processes, and biomathematics is now considered one of the modern applications of mathematics in an interdisciplinary research field, with numerous applications in biology,

biotechnology, and biomedical sciences (Hofmeyr, 2017; Bodine et al., 2020), which allows the modeling of biological processes using mathematical techniques and concepts (Hoppensteadt, 1995; Lombardero, 2014).

For instance, biotechnology applies a set of techniques and tools to modify living organisms to produce new knowledge to develop products and services. Consequently, to promote biotechnology projects, it is important to consider the creation of mathematical models to better understand patterns and phenomena within the world of biology, medicine, ecology, and environmental sciences, according to North Carolina State University (Campbell, 2013).

Several studies have focused on collaborative strategies, such as case studies (Cazes et al., 2006), and simulations (Jiménez et al., 2018), design-based thinking, problem-based learning (PBL) (Zorrilla-Pacheco et al., 2022) based on real problem-solving and knowledge construction in authentic professional contexts (Davidson and Worsham, 1992; Cuseo, 1996; Bature and Atweh, 2019). According to Vivanco-Galván et al. (2018), PBL is an academic strategy that allows the analysis and proposal of new enterprises based on research and innovation. Likewise, to maximize the participation and involvement of students in the content and activities to be developed, collaborative strategies (Chan and Idris, 2017; Oviedo and Zhuma Mera, 2019), which have a positive influence on student learning, were used. Jiménez and Castillo (2018) applied the design thinking methodology as a project-based strategy that allows students to strengthen their mathematical learning.

Nani (2015) proposed a PBL strategy that involves real-life challenges, focusing on real questions or problems with the potential to be applied in real fields. Likewise, Fisher et al. (2020) proposed a descriptive qualitative literature review on the application of PBL in mathematics, where the results indicate that PBL engages students in problem-solving and reasoning activities, given students opportunities to work autonomously to construct their own learning, and produces prototypes, products, or services from the students' work.

Guo et al. (2020) also proposed a literature review of PBL in higher education and concluded that more studies should be conducted to evaluate student learning processes and students' skills and competences. In addition, further experimental research should be conducted to determine the effects of PBL on student learning.

Research suggests learning that emphasizes open-ended problems may be more effective than direct instruction in boosting students' academic achievement (Yeh, 2009; Bartholomew and Strimel, 2018). To successfully implement PBL, we start by training students to answer the open-ended questions typically found in projects of their environment.

To transform this scenario, it is recommended that students engage in authentic professional contexts, where they can actively participate in problem-solving and construct knowledge. To achieve this, a robust learning model is essential—one that enables students to propose solutions for the challenges they encounter using an innovative approach. This model aims to enhance mathematical skills and foster a deeper understanding of the subject matter.

One important field for apply in education is the Biomathematics that finds wide-ranging applications across various fields, including epidemiology, ecology, microbiology, chemical reactions, and intracellular dynamics. In this study, our specific focus lies within the agri-food sector, encompassing wine companies and the cheese

industry. Additionally, we explore innovations in crop and seed improvement, such as functional foods and nutraceuticals, as well as their impact on novel food development. Furthermore, we delve into the realm of environmental applications, particularly bioremediation.

1.1 Related work

In this section, we will briefly introduce detailed information about the specific nature of the PBL-based projects developed and implemented, as well as the more recent approaches using PBL, Technologies and ChatGPT.

Lazic et al. (2021) indicate the influence of the PBL model in elementary mathematics education experiments with two groups on a sample ($N=147$) in order to compare a PBL-based education (experimental) with the usual way of learning mathematics (control). The results determined that there was a statistically significant effect of the interaction between teaching methodology in mathematics education and the mark in mathematics on student performance ($F=8.040$, $p=0.000$). Results shown that students from the experimental group have achieved better scores on the test compared to students from the control group.

Cruz et al. (2022) present a review on project-based learning (PBL) in learning math concepts. The basic idea is to do a data analysis focused on math, PBL, and TICs. The potential results suggest that PBL engages students in math learning environments and must be understood as a learning strategy to develop complex and soft mathematics-related skills.

Gningue et al. (2022), also suggest that a positive school climate and teacher leadership have been shown to benefit student achievement. Tashtoush et al. (2023) present a study whose aim is to investigate the impact of Information and Communication Technologies (ICT)-based education on the academic enthusiasm of 61 female students divided into two groups (31 experimental and 30 control) in mathematics. The study utilized the Mathematics Academic Enthusiasm (MAE) questionnaire, which included three subscales: cognitive, behavioral, and emotional. The reliability of the (MAE) questionnaire was evaluated by Cronbach's Alpha coefficient (0.89, 0.92, 0.91, and 0.91) for behavioral, cognitive, emotional, and overall scales, respectively. The study's findings revealed that the ICT-based education method had a more significant impact on students' MAE for the cognitive and behavioral subscales than traditional education methods.

Alneyadi and Wardat (2024) analyze the impact of Chatbot applications on student achievement in Quantum Theory courses. The experimental group (55 participants) received instruction enhanced by ChatGPT, while a control group (57 participants) experienced traditional teaching. The experimental group showed significant improvements in post-test scores for the Knowing, Applying, and Reasoning sub-skills. The overall test score for the experimental group was (15.55, 25.49, 17.51, 68.55) for each respective subscale, while the control group's scores were (10.63, 16.39, 11.73, 61.98). The use of ChatGPT in the experimental group resulted in a notable improvement in the "impact of ChatGPT on student achievement" compared to the control group.

In this sense, the principal aims were (i) to assess the comprehension of mathematical functions, (ii) to develop biotechnology projects PBL-based as a strategy for teaching-learning

TABLE 1 The total score averaging over 10 points to mathematical functions.

Topic	Control group	Score	Experimental group	Score
Mathematical functions	Questionnaire I	4	Questionnaire I	4
	Workshop	2	Workshop	2
	Test	4	Project (PBL)	4
	Total score	10	Total score	10

mathematical functions to university students, and (iii) to know its incidence in the knowledge acquisition, skills, perceptions, and motivation about the PBL application in a real work environment.

2 Materials and methods

2.1 Participants

The project was developed at the Universidad Técnica Particular de Loja (UTPL) with 111 students belonging to Business Administration (Group A), Biochemistry and Pharmacy and Chemical Engineering (Group B) as experimental group; and the control group corresponds to other group of Biochemistry and Pharmacy (Group C). Groups A, B and C were randomized before the start of the project.

The inclusion criteria are specifically targeted university students in their first year, focusing on mathematical functions and their relevance to biotechnology and bioenterprise. The exclusion criteria include academic years, non-mathematical subjects and technical, socio-humanistic, language, and religion careers.

The strategy was applied as a mathematical functions teaching-learning strategy in the second semester of the academic period from October 2018 to February 2019. Each group of analysis has had 37 students.

2.2 Instruments

To achieve our broad goals of equipping students with the ability to apply mathematical functions in their future careers, we employ various instruments in this study. Specifically:

2.2.1 Assessment

We utilize tests, questionnaires, and workshops to evaluate achievement.

To assess the understanding of mathematical functions, several types of instruments were used, such as the level of verbal and mathematical abilities of the students and comprehension assessment tools, such as questionnaires, test and workshops (see [Appendix A](#)). The Socratic tool¹ was used to answer the questionnaires and evaluate

the teaching-learning process. These instruments were applied to control and experimental groups.

2.2.2 Project-based learning

To develop projects, we implement a project-based learning (PBL) approach.

The project was used to assess the cognitive abilities involved in thinking (e.g., reasoning, teamwork, and problem-solving). To rate the students' project presentation, the quality of the scientific report and prototype a grading rubric with several criteria (see [Appendix B](#)) was used. Applied only to experimental groups.

2.2.3 Perceptions and motivation

We assess perceptions and motivation through a survey.

A student survey opinion was used as a PBL project perception assessment instrument with 16 questions of 5-point Likert scale (1: Strongly disagree; to 5: strongly agree), (see [Appendix C](#)) applied to experimental groups.

In the table below show details on the scores for each activity in experimental and control group ([Table 1](#)).

2.3 Reliability of the instrument

A principal component analysis (PCA) was conducted to examine the factor structure of the survey. The PCA used Varimax rotation and Kaiser normalization. The threshold for the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy was 0.6 ([Tabachnick and Fidell, 2000](#)). Cronbach's α coefficient values were assessed ([Cronbach, 1951](#)) in order to measure the internal reliability of a survey.

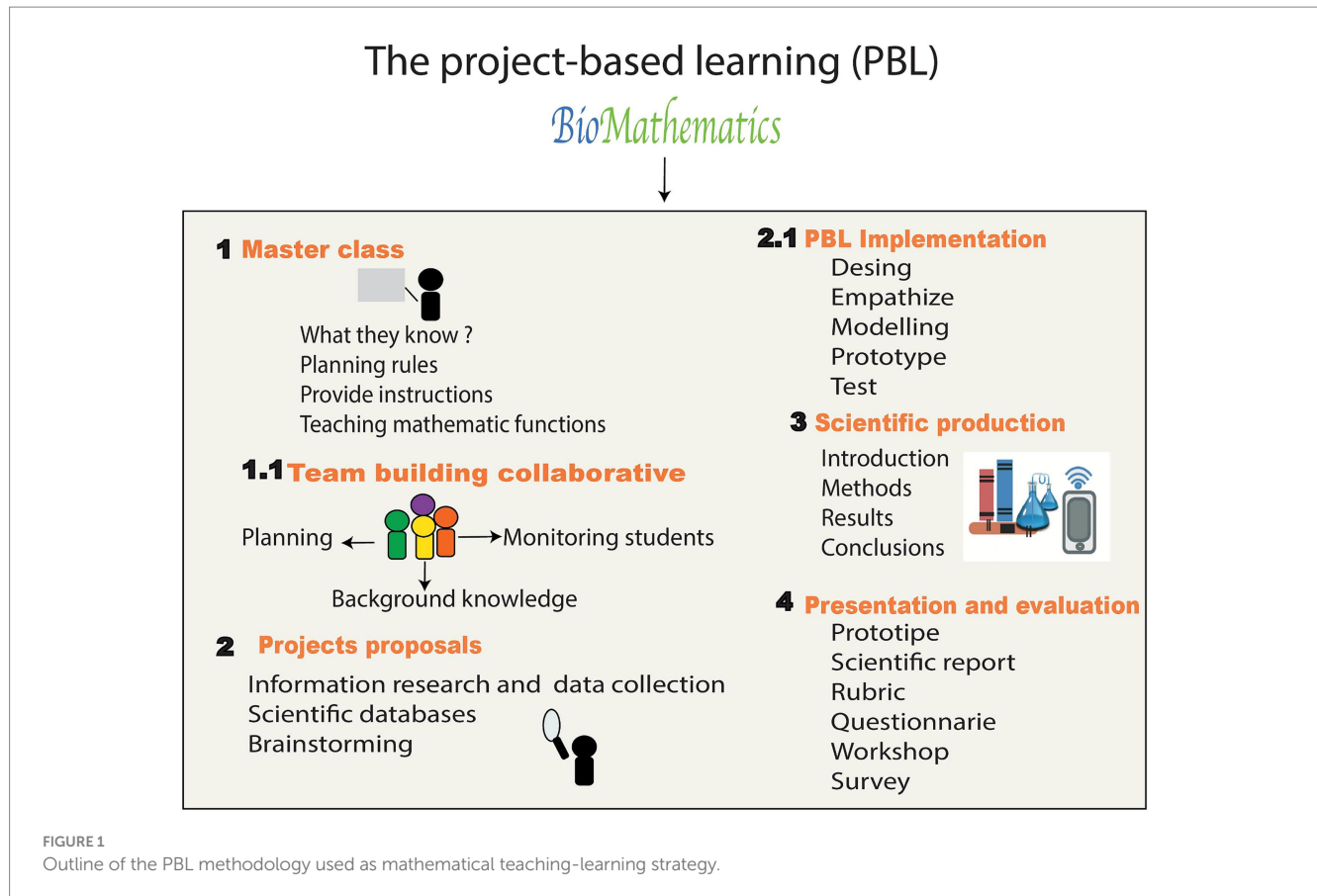
2.4 PBL methodology

This methodology was applied to experimental groups (A and B). Our proposal as Math and Biology teachers is to implement the PBL strategy in our lectures, we used the methodology applied by [Vivanco-Galván et al. \(2018\)](#), looking for a better way of teaching and learning mathematical function concepts, trying to facilitate connections between students' learning experiences and their future work environments. To achieve this, we focus on solving real-world biotechnological problems. By engaging with authentic challenges, students can bridge theory and practice, preparing them for their professional journeys.

This methodology includes the mathematical function master classes looking for adapting this knowledge to understand the problem definition, PBL implementation, prototypes, and scientific reporting and project presentation. Finally, the workflow includes the final test ([Figure 1](#)).

Here our role as the course teacher was to (i) plan the rules of the project; (ii) form a random work team; (iii) provide direct instructions before starting the project; (iv) monitor student projects during development to ensure that they have enough background knowledge, and (v) learning experience evaluation. The student's role is to propose PBL projects and to present a prototype and scientific report about the research.

¹ www.socratic.com



2.4.1 PBL strategy development

The main steps to development the PBL strategy were: (1) master class; (2) project proposals based on innovation, entrepreneurship, and technology topics (3) solution implementation by PBL strategy; (4) prototype and scientific report presentation and (5) results evaluation based on several instruments.

2.4.1.1 Phase 1: master class

According to the teaching plan contents to strengthen students' skills and achieve meaningful learning, the teacher explains the mathematical function concepts with emphasis on the development of practical exercises and workshops on polynomial, rational, exponential, and logarithmic functions. In addition, during the class, the teacher gives the rules of the project, conforms to random work teams, and provides direct instruction before starting a project.

This instruction can range from a short explanation of the quantitative and qualitative experimental design to strengthening scientific knowledge (open-ended questions, problem selection, planning, solution search, prototype design, and evaluation) to reinforcing mathematical topics about the functions describe above.

2.4.1.2 Phase 2: project proposals based on innovation, entrepreneurship, and technology topics

The proposed solutions focused on biotechnology contributions to the sustainable development goals (SDGs), most of which were related to health and well-being, underwater life and terrestrial ecosystems, climate action, affordable and clean energy, and quality education.

2.4.1.3 Phase 3: solution implementation by PBL strategy

The PBL projects were developed through the application of mathematical models, involving especially exponential and logarithmic functions to model bacterial population behavior, growth or decay of a radioactive substance, acidity or pH level of a substance, depreciation of equipment, profitability of a company, and capitalized interest among other biological and administrative applications implemented within the execution of the projects.

2.4.1.4 Phase 4: prototype and scientific report presentation

The evaluation of the projects was carried out by a committee of experts conformed by UTPL teachers-researchers, who used a rubric as an evaluation instrument ([Appendix B1](#)) to choose the topics more appropriate to the real needs of different categories. In addition, the scientific report, final presentation, and prototype of the project were considered in the rubric.

2.4.1.5 Phase 5: evaluation of results

Finally, as cognitive skill evaluation the Miller Pyramide ([Miller, 1990](#)) was adopted, and Bloom's Taxonomy ([Wu et al., 2013](#)) to evaluate the mathematical knowledge.

Students' competencies were evaluated based on the four levels established by George Miller's Pyramid ([Miller, 1990](#)). Starting from an initial level (*knows*) to a higher level (*Does*). The initial level was evaluated using workshops, questionnaires, and tests. A higher level was evaluated through the implementation of projects, resulting in prototypes or products ([Bueno and Fernández, 2016](#)). Throughout

this phase, the student is in the process of learning, so that he/she is acquiring knowledge and increasing his/her formative quality and pre-professional training.

2.5 Data analysis

Statistical analysis was performed using the software R version 4.1 (R Core Team, 2021). The nonparametric Kruskal-Wallis analysis was carried out to test the statistical significance between the experimental and control groups obtained from the questionnaire, workshop, project and final score, followed by the Mann-Whitney test. The descriptive analysis data were expressed as mean $\pm$ standard deviation (SD), standard error (SE). The survey was analyzed using principal components analysis (PCA).

3 Results

A total of 111 students participated in the project, 74 students in PBL process from Biochemistry, Chemical Engineering and Business Administration degrees, as the experimental groups and another Biochemistry class as the control group with 37 students. The most relevant results that characterize the PBL strategy and its application in learning mathematical functions are presented below.

In our study, the mean final scores present difference between experimental group (A and B) with an average score of 8.9/10 and 8.82/10 respectively, and the control group I with an average score of 6.66/10 (Table 2).

The results show that Kruskal-Wallis test revealed significant differences in the trueness values between the 3 groups (A-B-C) ($p < 0.001$) for each variable analyzed (Figure 2).

To determine the differences between groups we applied pairwise comparison tests showing significant differences ($p < 0.001$) between the experimental group (A and B) and the control group (C). The variable that showed significant differences ($p < 0.001$) between

experimental and control group were Project/test, questionnaire, total score of mathematical functions and workshop (Table 3).

The p -values of groups A and B correspond to PBL project, while the p value of C group correspond to the test. The table above show significant differences ($p < 0.001$) between the experimental group and the control group.

Thus, using PBL as mathematical function teaching-learning strategy through the application of workshops based on scientific cases, questionnaires, and projects, it was possible to increase academic performance in the experimental group. The application of the project indicated significant statistical differences (Figure 2) between the experimental group and the control group giving rise to scenarios in which students generate solutions to different problems within their professional environment, increasing their motivation for the study of mathematics.

3.1 Projects proposed by students

The proposals on bioentrepreneurship projects were divided into categories and subcategories as shown in Appendix D.

A mathematical model is an approach to reality, for this, the students modeled different scenarios, for example, bacterial growth, radioactive decay of a chemical substance, depreciation of the value of equipment, acidity level of a substance and concentration of hydrogen ions, all of them applied to a mathematical language. Each of the models shows different equations and mathematical functions to predict these processes (Table 4).

Thus, according to the total score obtained from the evaluation rubric scale (9–10 Excellent, 7–8 Good, 5–6 Fair, 4–5 Poor, 1–3 Very poor) the projects order which achieved higher scores under the concept of clarity, rationale, and internal feasibility of the proposal, considered relevant and feasible for its development at the local level are as shown in Appendix E.

3.2 Survey Cronbach's analysis

The opinion survey was applied through Google Forms to identify students' opinions about the PBL methodology applied in the classroom. Cronbach's analysis of the instrument resulted in a coefficient of 0.96; thus, the instrument was considered highly reliable.

The summary survey item results showed that the majority agreed that they obtained significant learning during the application of the PBL strategy, since the 53.5% is in agreement that learned to apply functions as mathematical models for the project development; 48.8% indicated that they learned how to plan a research project and write a scientific report based on the results obtained; 51.2% agreed that the development of the project helps there to understand the applicability of mathematics in their careers. In general, 53.5% say that the PBL strategy had a positive impact on their academic training.

4 Discussion

PBL has been shown to boost academic scores in both experimental and control groups. Additionally, it enhances skills and

TABLE 2 Descriptive statistical analysis between experimental and control groups, over each activity.

	Group	Mean	SE	SD
Workshop	A	2.00	0.00	0.00
	B	2.00	0.00	0.00
	C	1.81	0.05	0.36
Questionnaire	A	3.35	0.11	0.71
	B	3.14	0.10	0.62
	C	2.54	0.13	0.79
Project/Test	A	3.55	0.05	0.36
	B	3.68	0.05	0.32
	C	2.31	0.10	0.62
Total score mathematical functions	A	8.90	0.13	0.82
	B	8.82	0.11	0.69
	C	6.66	0.16	0.99

The overall score includes the score of workshops, questionnaire, project to experimental groups (A and B), and workshops, questionnaire and test for control group (C) over a total of 10 points. (A: Business Administration; B: Biochemistry and Pharmacy, Chemical Engineering and C: Biochemistry and Pharmacy). The samples size for each group was 37.

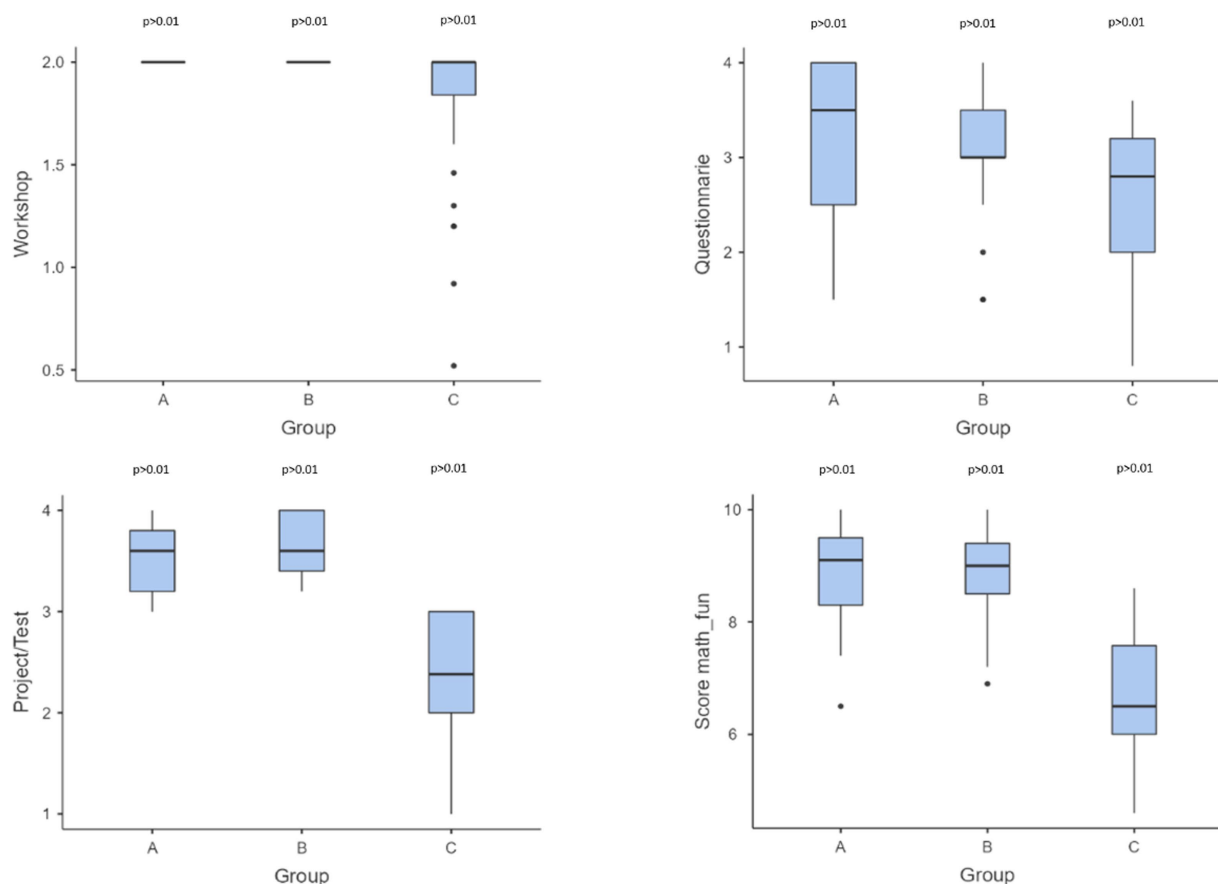


FIGURE 2

Boxplot showing the mean values of experimental and control groups for each variable analyzed. The y axis values correspond to the value of each activity. Questionnaire I (4 points), Workshop (2 points), and Test (4 points) in the control group, while in experimental groups we evaluate the Project (4 points) in replacement of the test.

enriches learning experiences by emphasizing the resolution of real-world problems.

The achievement gains, beyond attitude changes, PBL positively impacts student achievement. By applying math to real-world scenarios, students deepen their understanding and retain knowledge. PBL fosters critical thinking, communication, and teamwork—essential skills for success in any field. As a result, student performance improves, and they achieve better outcomes in assessments (Gay, 2022). In a study by Chapman et al. (2011) and Bender (2012) it was noted that the application of Project-Based Learning (PBL) leads to an increase in academic subject-passing scores. Our study reveals that the mean final scores for assessment indicator did not significantly differ between experimental groups A and B. Both groups achieved an average score of 8.9/10 and 8.82/10, respectively. However, a significant difference exists between the experimental and control groups. Specifically, the average score for mathematical functions in the control group was 6.66/10 (as shown in Table 2). Notably, paired comparison tests (Table 2) demonstrate significant differences ($p < 0.001$) for the project between the experimental group (A and B) and the control group (C) across all analyzed variables. This interpretation of the results should be approached with caution due to the limited number of samples. It represents a sample applied at Universidad Técnica Particular de Loja specifically for university

TABLE 3 Pairwise comparisons of variable Project/Test between experimental and control group.

		<i>W</i>	<i>p</i>
A	B	2.42	0.202
A	C	−10.08	<0.001
B	C	−10.56	<0.001

careers related to biotechnology and bioenterprise, excluding technical, socio-humanistic, language, and religion careers.

Globally, there has been an increased need to introduce advanced statistical-mathematical models and tools in research (Rodríguez and Bermúdez, 1995) where the central goal in science education is to teach students to think critically about scientific models and data (Holmes et al., 2015).

Thus, Rais and Lamada (2010) emphasize that PBL provides an ideal environment for applying skills and enhancing the quality of a student's learning process, reaching a high cognitive level remains a key objective of the learning process. In this way, the good use and application of these tools allow analysis of the problem, proposing ideas, and executing a project, as a real simulation applied in several possible scenarios (refer to Appendices D, E for further details).

TABLE 4 Mathematical models applied to biotechnological processes.

Function	Application	Mathematical equation
Exponential growth	Bacterial growth <i>Recursive function</i>	$P(t) = P_0 (1 + r)^t$
	Fermentation <i>Logistic function</i>	$P(t) = \frac{a}{1 + Ce^{-kt}}$
	Compound interest	$A = P \left(1 + \frac{r}{n} \right)^{nt}$
Exponential decrease	Radioactive decay over time	$N = N_0 2^{\frac{-t}{x}}$
	Exponential decay model	$N(t) = N_0 e^{-kt}$
	Depreciation	$A = P_0 (1 - r)^t$
Logarithmic	pH acidity level	$pH = -\log_{10} aH^+$
	pOH acidity level	$pOH = -\log_{10} aOH^-$
	Hydrogen ion concentration	$\left[H^+ \right] = 10^{-pH}$ $y = \log_a x$

As well as, PBL provides students with orientation to correlate the mathematics subject; with real life applications (Asija, 2011). The PBL development generated some proposals that could be part of the creation of bioenterprises that self-supply local and national markets, thus favoring the development of our country’s productive system. At the national level, companies contributing to the national economy operate in the agri-food sector, specifically in food production using biotechnological methods, food safety (quality, conservation, etc.), seed production and crop improvement processes (CEIM Confederación Empresarial de Madrid-CEOE, 2002).

Likewise, as shown above, mathematics can contribute to the development of various projects, such as projects in the agricultural sector, as explained by Chávez Esponda et al. (2013), through the calculation of soil pH knowing the concentration of hydronium, which leads to the equation of a function of great utility for the study of soil characteristics, as well as the inverse process. This information is highly valuable for improving the soil conditions prior to crop establishment. Other widely used functions are the response curves, which allow the establishment of a relationship between crop yield and nutrients in a plot of land. The calculation of irregularly shaped land areas from field measurements and the calculation of the growth rate of a plant uses differential and integral calculus tools. However, the current study could be constrained to mathematical functions. It would be intriguing to conduct a fresh investigation that encompasses additional mathematical subjects, all while embracing the Project-Based Learning (PBL) approach using Artificial Intelligence (AI) education tools. Such an endeavor could yield benefits for high-need students across various global contexts, provided that classrooms incorporate suitable PBL project activities.

Finally, according to the results obtained from the opinion survey (Appendix F), we consider that the formative experience allowed the teacher to improve the mathematics teaching-learning and increase the percentage of approved students, while they increased their degree of motivation using experimental scientific designs PBL-based.

Improving the skills and abilities developed by students to generate biotechnological solutions to problems in their environment and achieving a better potential to solve problems through logical reasoning.

5 Conclusion

The study was conducted with only 111 university-participating students, represented by three courses: Biochemistry and Pharmacy, Chemical Engineering and Business Administration carrers, representing about 15% of all first academic cycles at UTPL. It suggests that the findings are based on a small representative sample of the first cycle. However, it could be argued that as future work, we might increase the representative sample to all Biological Sciences and Medicine Faculties.

In this experience, the experimental and control groups showed significant differences ($p < 0.001$), demonstrating that PBL as a mathematical teaching-learning strategy provided promising results in the first academic cycles at UTPL. Moreover, teaching strategies strengthen scientific knowledge in students’ critical thinking and problem-solving skills. It suggests that PBL can be applied in another classroom and provide opportunities for students to build self-confidence and demonstrate skills they cannot demonstrate on a traditional assessment.

5.1 Limitations and future work

The main limitations of this work are the sample size and potential bias. These limitations must be addressed in future work, such as increasing the sample size to increase the validity of the assessment of PBL activities, implementing blinding and randomization assignments to help ensure that participants are assigned to groups in an unbiased manner, and reducing the risk of selection bias.

Also, incorporating new methodologies focused on artificial intelligence (AI) models like ChatGPT into Project-Based Learning (PBL) can offer exciting opportunities to enhance student learning experiences. For example, encouraging students to utilize ChatGPT as a research tool to gather information, brainstorm potential solutions, analyze data, simulate scenarios and generate new ideas for their projects, to generate project documentation such as written reports, or develop interactive presentations to communicate their project findings and insights effectively, taking into account the ethical considerations of using AI.

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

YJ-G: Writing – original draft, Writing – review & editing. OV-G: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Alneyadi, S., and Wardat, Y. (2024). Integrating ChatGPT in grade 12 quantum theory education: an exploratory study at emirate school (UAE). *Intelligence* 2:4.
- Asija, H. (2011). Developing project-based learning beyond the curriculum. Presentation at the National Initiative on Mathematics Education, Northern Region Conference, November 21–23 (2011).
- Bartholomew, S. R., and Strimel, G. J. (2018). Factors influencing student success on open-ended design problems. *Int. J. Technol. Des. Educ.* 28, 753–770. doi: 10.1007/s10798-017-9415-2
- Bature, I. J., and Atweh, B. (2019). Collaboration: a collective bargain for achieving quality mathematics classroom practice. *Int. J. Educ. Methodol.* 5, 347–361. doi: 10.12973/ijem.5.3.347
- Bender, W. N. (2012). *Project-based learning: Differentiating instruction for the 21st century*. California: Corwin Press.
- Bodine, E. N., Panoff, R. M., Voit, E. O., and Weisstein, A. E. (2020). Agent-based modeling and simulation in mathematics and biology education. *Bull. Math. Biol.* 82, 1–19. doi: 10.1007/s11538-020-00778-z
- Bueno, C. R., and Fernández, J. T. (2016). Evaluación de competencias profesionales en educación superior: retos e implicaciones. *Educ XX1* 19, 17–37. doi: 10.5944/educXX1.12175
- Campbell, P. N. (2013). *Biology in profile: a guide to the many branches of biology*. London, England: Elsevier. 3–5.
- Cazes, C., Gueudet, G., Hersant, M., and Vandebrouck, F. (2006). Using E-exercise bases in mathematics: case studies at university. *Int. J. Comput. Math. Learn.* 11, 327–350. doi: 10.1007/s10758-006-0005-8
- CEIM Confederación Empresarial de Madrid-CEOE. (2002). Guía de creación de bioempresas [Archivo PDF]. Available at: <https://www.madridmasd.org/sites/default/files/informacionidi/biblioteca/publicacion/doc/Libro14madrimasd.pdf>
- Chan, L. L., and Idris, N. (2017). Cooperative learning in mathematics education. *Int. J. Acad. Res. Bus. Soc. Sci.* 7, 539–553. doi: 10.6007/IJARBS/v7-i3/2757
- Chapman, C., Laird, J., Ifill, N., and KewalRamani, A. (2011). Trends in High School Dropout and Completion Rates in the United States: 1972–2009. U.S. Department of Education.
- Chávez Esponda, D., Sabín Rendón, Y., Toledo Dieppa, V., and Jiménez Álvarez, Y. (2013). La Matemática: una herramienta aplicable a la Ingeniería Agrícola. *Revista Ciencias Técnicas Agropecuarias* 22, 81–84.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1007/BF02310555
- Cruz, S. M. A., Viseu, F., and Lencastre, J. A. (2022). Project-based learning methodology as a promoter of learning math concepts: a scoping review. *Front. Educ.* 7:953390. doi: 10.3389/feduc.2022.953390
- Cuseo, J. B. (1996). *Cooperative learning: A pedagogy for addressing contemporary challenges & critical issues in higher education*. Stillwater, OK: New Forums Press.
- Davidson, N., and Worsham, T. (Eds.). (1992). *Enhancing thinking through cooperative learning*. New York: Teachers College Press.
- Fisher, D., Kusumah, Y. S., and Dahlan, J. A. (2020). Project-based learning in mathematics: a literatur review. *J. Phys. Conference Series*. 1657.
- Gay, E. (2022). Project-based learning in mathematics classroom. Thesis, Assumption University. Digital commons @ Assumption University.
- Gningue, S. M., Peach, R., Jarrah, A. M., and Wardat, Y. (2022). The relationship between teacher leadership and school climate: findings from a teacher-leadership project. *Educ. Sci.* 12:749. doi: 10.3390/educsci12110749
- Guo, Y., Liu, B., and Zhao, D. (2022). Online continual learning through mutual information maximization. *Inter. Conference Mach. Learn.* 8109–8126.
- Guo, P., Saab, N., Post, L. S., and Admiraal, W. (2020). A review of project-based learning in higher education: student outcomes and measures. *Int. J. Educ. Res.* 102:101586. doi: 10.1016/j.ijer.2020.101586
- Han, S., Capraro, R., and Capraro, M. M. (2015). How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: the impact of student factors on achievement. *Int. J. Sci. Math. Educ.* 13, 1089–1113. doi: 10.1007/s10763-014-9526-0
- Hofmeyr, J. H. S. (2017). Mathematics and biology. *S. Afr. J. Sci.* 113, 1–3. doi: 10.17159/sajs.2017/a0203
- Holmes, N. G., Wieman, C. E., and Bonn, D. (2015). Teaching critical thinking. *Proc. Natl. Acad. Sci.* 112, 11199–11204. doi: 10.1073/pnas.1505329112

Acknowledgments

The authors give thanks to the Vicerrectorado Académico of the Universidad Técnica Particular de Loja.

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/feduc.2024.1364640/full#supplementary-material>

- Hoppensteadt, F. (1995). Getting started in mathematical biology. *AMS* 42, 969–975.
- Jiménez, Y., and Castillo, D. (2018). Educación de calidad mediante la estrategia Design Thinking. In Conference proceedings EDUNOVATIC 2017: 2nd virtual international conference on education, innovation and ICT
- Jiménez, Y., Valarezo, M. E. B., Carrillo, I., Castillo, D., Feijoo, D., and Vivanco, O. (2018). "Simulation technologies to strengthen teaching-learning skills in biochemistry, nursing and medical students" in *Perspectivas educativas*. eds. J. A. Nicoletti and M. C. Perissé (Universidad Nacional de La Matanza), 81–98.
- Lazic, B., Knežević, J., and Maričić, S. (2021). The influence of project-based learning on student achievement in elementary mathematics education. *S. Afr. J. Educ.* 41, 1–10. doi: 10.15700/saje.v41n3a1909
- Lennox, A., and Roos, F. (2017). "Preparing students for the workplace—development of soft skills" in *Changing business environment: gamechangers, opportunities and risks*. eds. N. Delener and C. Schweikert (New York: The Guilford Press) 354–360. Available at: https://www.researchgate.net/profile/Zhenning_Xu2/publication/320064694_MARKET_SEGMENTATION_USING_CLUSTER_ANALYSIS_AND_COMBINATORIAL_FUSION/links/59cbb1e20f7e9bbf3b6bb2/MARKET_SEGMENTATION-USING-CLUSTER-ANALYSIS-AND-COMBINATORIAL-FUSION.pdf#page=373
- Lombardero, O. A. (2014). Un vistazo a la Biomatemática. *Números* 86, 29–38.
- Markula, A., and Aksela, M. (2022). The key characteristics of project-based learning: how teachers implement projects in K-12 science education. *Discip. Interdiscip. Sci. Educ. Res.* 4, 1–17. doi: 10.1186/s43031-021-00042-x
- Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Acad. Med.* 65, S63–S67. doi: 10.1097/00001888-199009000-00045
- Nani, K. L. (2015). Kemampuan Penalaran Statistis, Komunikasi Statistis, dan Academic Help-Seeking Mahasiswa Melalui Pembelajaran Berbasis Proyek Berbantuan ICT. Doctoral dissertation, Universitas Pendidikan Indonesia.
- Oviedo, B., and Zhuma Mera, E. (2019). Estrategias de trabajo colaborativo ABP-TPA. *Revista Universidad Sociedad* 11, 153–158.
- Rais, M., and Lamada, M. S. (2010). Model project based-learning sebagai upaya meningkatkan prestasi akademik mahasiswa. *Jurnal Pendidikan Dan Pengajaran*. 43. doi: 10.23887/jppundiksha.v43i3.129
- R Core Team. (2021). *R: A language and environment for statistical computing [computer software]*. Vienna, Austria: R Foundation for Statistical Computing.
- Rodríguez, L. F., and Bermúdez, L. T. (1995). Usos y aplicaciones de la simulación en la investigación agropecuaria. *Agronom. Colomb.* 12, 198–204.
- Sahin, M. (2009). Correlations of Students' grades, expectations, epistemological beliefs and demographics in a problem-based introductory physics course. *Int. J. Environ. Sci. Educ.* 4, 169–184.
- Tabachnick, BG, and Fidell, LS. (2000) *Computer-assisted research design and analysis. 1st ed.* Needham Heights: Allyn & Bacon, Inc.
- Tashtoush, M. A., AlAli, R., Wardat, Y., Alshrafin, N., and Toubat, H. (2023). The impact of information and communication technologies (ICT)-based education on the mathematics academic enthusiasm. *J. Educ. Soc. Res.* 13:284. doi: 10.36941/jesr-2023-0077
- Vivanco-Galván, O. A., Castillo-Malla, D., and Jiménez-Gaona, Y. (2018). HACKATHON multidisciplinario: fortalecimiento del aprendizaje basado en proyectos. *Revista Electrón. Calidad Educ. Superior* 9, 119–135. doi: 10.22458/caes.v9i1.1893
- Vogler, J. S., Thompson, P., Davis, D. W., Mayfield, B. E., Finley, P. M., and Yasseri, D. (2018). The hard work of soft skills: augmenting the project-based learning experience with interdisciplinary teamwork. *Instr. Sci.* 46, 457–488. doi: 10.1007/s11251-017-9438-9
- Wu, S. Y., Hou, H. T., Hwang, W. Y., and Liu, E. Z. F. (2013). Analysis of learning behavior in problem-solving-based and project-based discussion activities within the seamless online learning integrated discussion (SOLID) system. *J. Educ. Comput. Res.* 49, 61–82. doi: 10.2190/EC.49.1.c
- Yeh, Y. C. (2009). Integrating e-learning into the direct-instruction model to enhance the effectiveness of critical-thinking instruction. *Instr. Sci.* 37, 185–203. doi: 10.1007/s11251-007-9048-z
- Zorrilla-Pacheco, S. C., Flores-Samaniego, Á. H., and Jiménez-Gaona, Y. C. (2022). El Aprendizaje Basado en Proyectos y su aplicación didáctica en la enseñanza de las medidas de localización. *Revista Electrón. Calidad Educ. Superior* 13, 226–249. doi: 10.22458/caes.v13i1.4043



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo,
Portugal

REVIEWED BY

Inês Peixoto Silva,
Instituto Politécnico de Viana do Castelo,
Portugal
Zhuoying Wang,
The University of Texas at Austin,
United States

*CORRESPONDENCE

Arne Sørensen
✉ arne.sorensen@nord.no

RECEIVED 05 January 2024

ACCEPTED 19 March 2024

PUBLISHED 17 April 2024

CITATION

Sørensen A and Lagestad P (2024) Physical education students' reflections about the learning outcomes of different teaching methods: a mixed methods study. *Front. Educ.* 9:1365916. doi: 10.3389/feduc.2024.1365916

COPYRIGHT

© 2024 Sørensen and Lagestad. 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.

Physical education students' reflections about the learning outcomes of different teaching methods: a mixed methods study

Arne Sørensen* and Pål Lagestad

Department of Teacher Education and Arts, Nord University, Bodø, Norway

Introduction: Teaching in higher education is still mainly executed as lectures, even though research about student-active instruction methods points to more motivated students, higher enjoyment, and more optimal learning outcomes. The purpose of this study was to obtain better insight into how physical education (PE) students assessed their learning outcomes in relation to the use of different pedagogical approaches.

Methods: A master's course in PE was planned and implemented using the following eight different learning approaches: lectures; practical exercises about themes in lectures; discussions during lectures; discussions outside of lectures; planning and exercises for peer students; individual work preparing to write an academic text; individual work writing the academic text; and reading for an exam. The study constituted a mixed methods study, which used quantitative data from students' evaluation of eight different learning approaches on a Likert-type scale, and in-depth qualitative data from follow-up interviews with some of the same students, with the aim of explaining the main findings. Quantitative data about the students' reflections on the learning outcomes of the different learning approaches were collected among 59 different students at three different times (2021, 2022, and 2023), after finishing a course in the fifth semester in a master's program in PE.

Results: The findings showed that the students reported achieving the highest learning outcomes from practical exercises and attaining the lowest learning outcomes from lectures. In depth interviews among seven randomly selected students were also used to obtain reflections from the students about the different learning approaches. Quantitative analyses again revealed that practical exercises produced the highest learning outcomes, while lectures resulted in the lowest learning outcomes. Qualitative analyses of the in-depth interviews indicated that practical activities enabled students to relate theory to practice, make them active, and are associated with future work, while the quality of lectures depended on characteristics of the teacher and were often experienced as long and unstimulating.

Discussion: According to the results, we recommend that student teachers in higher education acquire the ability to plan and execute practical lessons in relation to themes focused upon in lectures and involve students more in discussions during lectures.

KEYWORDS

student teachers, learning outcomes, lectures, practical activities, pedagogical methods

Introduction

According to the Norwegian government's report on recommended strategies to ensure excellent quality in higher education (Meld. St. 16, 2016–2017), teachers in higher education should evaluate the efficacy of their educational work. Because high research quality seems to be relatively more rewarded than teaching in higher education, and most of the lecturers in Norway are more focused on research than teaching, Vabø and Ramberg (2009) find that such evaluation studies may be lacking.

In Norway, teachers at universities and university colleges are often simultaneously performing teaching and conducting research. Although they are supposed to practice “research-based teaching,” precisely what this expression means and how best to implement it remain unclear (Børte et al., 2020).

Previous research

Traditionally, teaching in universities is executed as lectures that are delivered by highly qualified professors. Indeed, this is still the most common method in higher education, despite the fact that it has been widely criticized for causing students to be passive listeners and diminishing learning outcomes (Young et al., 2009; Wieman, 2019). Ideally, education provided to students should both constitute a presentation of updated research on the actual subject in the curriculum, as well as use the learning approaches that create the highest learning outcomes and are most efficient (Griffiths, 2005). For example, if teachers in higher education plan and execute their lessons with small “breaks” in the lecture and use different methods for involving the students, such as problem sets, brainstorming or open discussions, research has demonstrated improved exam performance, and students reported better perceived effectiveness and decreased distractions (Miller et al., 2013). Moreover, the students stated that during long and unstimulating lectures, they are more likely to send e-mails, send text-messages, or engage in other activities on their electronic devices than during engaging lectures (Miller et al., 2013). For students who are familiar with lectures that are devoid of interaction with the lecturer or who believe that breaks during in the lecture produce low learning, these activities feel markedly unproductive (Knight and Wood, 2005).

Brown and Bakhtar (1988) report that the lecturer saying too much, speaking too quickly, and not including sufficient summaries constitute a major challenge for students. They suggest that the effects of lectures can be substantially improved by focusing on presenting the lecture in more succinct manner and keeping the focus on the major points of the lecture. Extant literature has identified several barriers that prevent lectures from including active learning methods in educational plans, including a lack of requisite class time and insufficient time to prepare active learning methods. Furthermore, it is argued that teachers' familiarization with holding traditional lectures makes continuation of such lectures comfortable for them and enjoyable (Brown and Bakhtar, 1988; McCabe and O'Connor, 2014; Miller and Metz, 2014; Trinidad, 2019). Students believe that the reason for the lecturers decision not to incorporate active learning methods in their lessons could be either that they want to control the education through the use of lectures or they do not possess the needed knowledge to implement active learning methods (Miller and

Metz, 2014). Furthermore, when curricula are relatively large and information-dense, lecturers regard lectures that do not include any interference from the students to be the only viable option (Lujan and DiCarlo, 2006).

Regmi (2012) contends that the debate should not concern being for or against using lectures as a method for learning, but rather how to organize teaching with several methods to enhance learning outcomes. In fact, this argument is almost identical to that of Penson (2012), who acknowledges the advantages of continuing to have lectures in education at universities, as well as combining them with relevant practical activities to produce optimal education. Medical students stated that lectures were important, as they assisted them to obtain key knowledge in a subject, and that learning was effective when combined with other learning methods. It was also reported that learning in the lectures depended on the characteristics and skills of the lecturer (Bates et al., 2018). A study among nursing students identified the following five essential factors that improved learning in lectures: (1) effective and clear speaking; (2) an emphasis on exam topics; (3) taking students' interests into account; (4) connecting theory to practice; and (5) providing a well-structured presentation (Al-Modhefer and Roe, 2009). A prerequisite for lectures to be educational is for the lecturer to be actively engaged (French and Kennedy, 2017). Ideally, the lecturer should present the latest updated theories and be able to discuss different perspectives about the theories. If a lecturer succeeds in this, it will positively affect students' motivation and challenge them academically (French and Kennedy, 2017).

It is well documented that by activating students, learning outcomes and motivation increase (Freeman et al., 2014; Damsa et al., 2015; Goodman et al., 2018). According to Bernstein (2018), active learning methods include posing questions for in-class or online discussions, using problem-based/case-based teaching, including peer reviews of writing, giving homework assignments, performing laboratory experiments, and “underteaching,” amongst others (Bernstein, 2018). The main purpose of active learning methods is to put the students in situations in which they engage in activities that are relevant to the subject, and ensure that the situation/task leads to reflection, communication, and analysis (Dewey, 2008; Smith and Cardaciotto, 2011). Furthermore, it is recommended that the teacher facilitate and create environments in education in which students' learning occurs in the specific activity (Goodman et al., 2018). In fact, students appreciate active learning methods and highlight several positive effects that are derived from it, including markedly increased autonomy and engagement (McCabe and O'Connor, 2014; Sørensen et al., 2023). Active learning methods also increase involvement and motivation for the subject (Lea et al., 2003), improve in-depth learning, and create a greater acceptance of failure in the subject (Lea et al., 2003; McCabe and O'Connor, 2014). Previous literature has also reported improved performance in exams and a decrease in failure in exams because of the use of active learning methods instead of lectures (Freeman et al., 2014; Goodman et al., 2018).

Theoretical framework

From a theoretical perspective, Dewey (1916) describes the need for societies to pass on the skills and knowledge that they accrue through a formal system of education. However, transferring learning

through a formal context is problematic, according to Dewey. Dewey offers a general critique of the general trend of academization, and censures teaching for not sufficiently considering pragmatic aspects of life. For Dewey (1916, p. 307), vocation is “a continuous activity having a purpose.” In other words, when teachers relate theory to practice, a meaningful purpose is obtained. Indeed, Dewey states that we learn through vocations, rather than for them—a theory closely related to Aristotle’s concept of “*phronesis*,” or practical wisdom. Drawing upon Dewey, Kolb (1984) defines experiential learning as a process that links education, work, and personal development. It is again asserted that theory must focus on practice. Without an interested student, without a clear starting point in the student’s former experience/expectations, obtaining a good learning result is highly challenging. Furthermore, Dewey (1916) points out that an interaction must occur between the student and what is learned, i.e., the student must learn theory that he or she finds both useful and interesting. However, students report that the theory “easily becomes remote and dead—abstract and bookish” Dewey (1916, p. 8). Dewey contends that with a clear starting point in pragmatic examples, achieving excellent learning outcomes will be much easier.

It is asserted that schools should change their educational practice from lecture-focused to providing students with experiences that apply to real-life situations (Dewey, 2008). Consequently, students could improve their learning outcomes through the process of reflecting upon their experiences. Furthermore, by giving more responsibility to the students for the learning process itself, they will be more capable of constructing their own learning (Trinidad, 2019). Researchers have demonstrated that Dewey’s progressive pedagogy and philosophy still play a role in schools’ daily practice (Berg, 2013; Little, 2013; Lagestad, 2014; Sadovnik et al., 2017).

Research questions

Considering the previous discussion about learning from different pedagogical approaches, this study will examine the following two research questions: (1) Are there differences between students’ reflections about their learning outcomes using different pedagogical approaches? and (2) What are the students’ reflections about the pedagogical approach that creates the best and the worst learning outcomes?

Methods

Design

This study uses a mixed-methods design that comprises both questionnaire data to examine students’ reflections on their learning outcomes in relation to the use of eight different pedagogical approaches, and in-depth interviews to obtain more nuanced knowledge about their reflections about this research area. Quantitative analyses assess differences between students’ reflections of their learning outcomes in eight pedagogical approaches, while the interviews investigate why students prefer certain pedagogical approaches over others.

The research was carried out in a master’s program in physical education (PE) in the fifth semester. In terms of the curriculum, the two main learning outcomes of this subject were: (1) to determine

how sport teachers at college should facilitate their education to develop skills and performance in sport at college; and (2) to identify how sport teachers at college should use research-based education to increase their students’ skills and performance. Education in the actual subject was similar in the three following years, with the same teacher and the same procedures. This research was executed in this actual 15-credit course, but no other courses of the master’s degree had been changed. This research was executed at a small university with relatively small groups of around 20 students in each year, and with only one class of PE each year. The course was organized according to the curriculum, and teaching in the subject was planned and carried out using varied teaching methods, including lectures, and practical exercises/activities that dealt with the topics in lectures and discussions in the class. In addition, the students had long-term work in which they attempted to participate in a research-based exercise program for six weeks. This work was largely autonomic, in which the students could choose training methods and sport based on extant scientific knowledge in the actual sport. This academic task included execution of pre- and post-tests, and exercises concerning relevant research.

The students were also given responsibility in groups to complete certain practical lessons. In addition, one question was added to the questionnaire which did not directly concern teaching, but rather dealt with academic discussions with fellow students or others. The intention here was to determine whether this was regarded as an area that could provide positive learning effects for the students.

Participants

The quantitative part of the study (questionnaire) included students from three different classes in a master’s program in PE in three consecutive years (2021, 2022, and 2023). At the time of the data collection, the students were in the middle of their five-year program. The participants received a verbal orientation that explained the purpose of the study. They were also informed that participation in the study was voluntary, and that they could withdraw from the study at any time without consequence. All the students that participated provided written consent in accordance with regulations of the Norwegian Center for Research Data (ref. code no 383620).

Of the 17 students of the class of 2021, five females and nine males filled out the questionnaire in November 2021 (77.8% response rate). Of the 23 students of the class of 2022, eight females and 15 males (100% of the class) filled out the questionnaire in November 2022. Of the 22 students of the class of 2023, all students (seven females and 15 males) filled out the questionnaire in November 2023. In total, 59 of the 62 students (20 females and 39 males) filled out the questionnaire, which yielded a response rate of 93.6%. The qualitative part of the study (interviews) included seven randomly selected students from the 2023 group. The reason for low participation in 2021, where that some of the students did not meet at the university the day of data collection.

Procedures

The selection of the eight different learning methods were based on constructivist pedagogy and the researchers experiences with the

efficacy of these methods. Quantitative data about the students' reflections on the learning outcomes of eight different methods were collected at three different times (2021, 2022, and 2023) after completing a course in their fifth semester in a master's program in PE. The purpose of the questionnaire was to attain knowledge of the students' experience of their learning in eight different pedagogical approaches used in the course, which were: (1) lectures; (2) practical exercises about themes in lectures; (3) discussions during lectures; (4) discussions outside of lectures; (5) planning and exercises for peer students; (6) individual work preparing to write an academic text; (7) individual work writing the academic text; and (8) reading for an exam. According to the specific research question, the questionnaire was self-developed and not based upon a pre-validated instrument. However, the questions and the answers options also have a high face validity (Holden, 2010), and should not lead to different interpretations of the questions. Furthermore, we will argue that both the questions and the answer options led to high reliability.

The questionnaire was designed with a scale between one and six, where six was the highest score:

1. I achieved no learning outcomes from this method.
2. I achieved little learning outcomes from this method.
3. I achieved some learning outcomes from this method.
4. I achieved good learning outcomes from this method.
5. I achieved very good learning outcomes from this method.
6. I achieved excellent learning outcomes from this method.

The survey was carried out in the classroom at the same time of year (November) in all three times (2021, 2002, and 2023), and the questionnaire was administered in paper format. A researcher was present during all the data collection and was available to assist with any questions or difficulties that the students may have had filling out the questionnaire; however, no student had any such questions or difficulties. When the students replied to the survey, they had used these methods for the entire semester (August–November). However, they had experienced some of these methods during their 2.5 years of master study.

An interview guide was created with the aim to examine the reasons for the high scores on practical exercises and low scores on lectures. The main purpose of the questions was to gain access to the students' reflections on their experiences of eight different pedagogical methods used during the course, as well as what they considered to constitute the advantages and disadvantages of these learning outcomes. Open-ended questions that were included were: "Can you explain an episode in your own schooling where you felt you learned a lot?," "Can you describe the way you learned and why you remember this episode well?," "Can you describe how you feel you learn during practical activities?," and "Can you describe how you feel you learn in lectures?" At the end of the interview guide, the following two questions were included about the quantitative findings: "The findings showed that the students identified lectures as the method that created the least learning. Were you surprised by this result, and what are your reflections about it?," and "The findings showed that the students identified practical activities as the method that created the most learning. Were you surprised by this result, and what are your reflections about it?"

After completing the interview guide, a pilot interview was carried out with a randomly selected student in the same study as the

participants in the present study, with the intention to test the effectiveness of the questions in the interview guide. The interview guide obtained many reflections from the student regarding her learning in practical activities and in lectures, and thus was not changed. The interviews of the seven randomly selected students took place during a two-week period in November and December 2023. A voice recorder (Olympus DS-50) was used during the interviews. All interviews were conducted by the same researcher, at the researcher's offices. The individual interviews lasted between 45 and 60 min.

Data analysis

The student's responses on the questionnaires were analyzed in SPSS, Version 29 (IBM, Armonk, NY, United States). There was a non-normal distribution of the variables (seven out of eight variables achieved significance on a Kolmogorov–Smirnov test). Moreover, since the scale used to measure the dependent variables did not meet the assumptions of parametric tests, nonparametric tests were performed. Mann–Whitney U-tests were conducted to test for differences between year of participation. No differences were found for the eight pedagogical methods regarding year of participation ($p > 0.05$). A Friedman nonparametric test was performed to examine differences in learning outcomes within the eight pedagogical approaches, followed by Wilcoxon nonparametric tests (pairwise comparison).

The interview data were transcribed and entered in the qualitative analysis program NVivo 12 Plus. The interviews were transcribed verbatim, and coded so that the participants were anonymized and given pseudonyms. The analyses were based on transcribed answers focusing on meanings, as described by Johannessen et al. (2016). Opinions and statements were assessed for themes, and then condensed, coded, and categorized in units of analysis (Brinkmann and Kvale, 2009). In this process, the participants' statements were assigned codes that were classified into categories (Hastie and Glotova, 2012). The data were sorted based on these categories to elucidate patterns, similarities, relationships, and/or differences between the statements. The analysis and the interpretation followed hermeneutical principles, in that the interpretation process led to an increasingly deeper understanding of the statements, in parts and in aggregate, in the interviews (Kvale, 1983). The transcribed text was read several times.

Reading the text led to the creation of 12 categories from the students' statements. However, five of these categories were related to the pedagogical methods "lectures" and "practical exercises about themes in lectures," and only these were selected into this study. As shown in the Results section, the main findings of the quantitative analyses were that the pedagogical approach "practical exercises about themes presented in lectures" had the highest learning outcomes score and significantly higher scores than six of the seven other pedagogical approaches. Furthermore, "lectures" had the lowest learning outcomes score, and it was significantly lower than two of the other pedagogical approaches. Therefore, the interview data included in this study included only the three categories about practical exercises and the two categories about lectures, with the aim to obtain more nuanced knowledge regarding the students' reflections about these two pedagogical approaches. The three categories about "practical exercises about themes presented in lectures" were: (1) practical

activities—to relate theory to practice; (2) practical activities about future work; and (3) practical activities—the importance of being active. The two categories about “lectures” were: (1) learning in lectures depends on the teacher; and (2) the lectures are often long and unstimulating.

Results

Quantitative results

Friedman analyses showed that there was a significant difference between the eight pedagogical approaches ($\chi^2_{59} = 49.3, p < 0.001$), in which practical exercises exhibited the highest learning outcome, while lectures showed the lowest learning outcome. Follow-up analyses using the Wilcoxon test with Bonferroni corrections revealed that practical exercises produced significantly higher learning outcomes than lectures ($Z = -5.2, p < 0.001$), discussions during lectures ($Z = -3.7, p < 0.001$), discussions outside of lectures ($Z = -3.4, p < 0.001$), planning and exercises for peer students ($Z = -4.4, p < 0.001$), individual work writing an academic text ($Z = -3.3, p < 0.001$), and reading for an exam ($Z = -3.7, p < 0.001$). Furthermore, lectures were reported to yield significantly lower learning outcomes than practical exercises and preparing to write an academic text ($Z = -5.2, p < 0.001$ and $Z = 3.8, p < 0.001$, respectively).

Qualitative results

The qualitative analyses of lectures and practical exercises pointed towards the following five categories: (1) learning in lectures depends on the teacher; (2) lectures are often long and unstimulating; (3) practical activities—to relate theory to practice; (4) practical activities about future work; and (5) practical activities—the importance of being active.

Learning in lectures depends on the teacher

The statistical analyses of the quantitative data showed a significantly lower experienced learning outcome of lectures compared to the pedagogical learning methods of practical exercises and preparing to write an academic text. The first main finding from the analyses of the interview data was that learning in lectures depended on the characteristics and skills of the teacher. Indeed, Susan reported that “the teachers are so different,” and Deborah stated that it was “a great advantage if it’s a person who knows a lot and is good at lecturing, and who is interactive with us listeners.” The students agreed that some lectures were more interesting to listen to than others, and that this was due to the quality of teaching from the individual teacher. When Sebastian was asked how he felt about sitting and listening to a lecture, he stated: “I feel it’s easy to drift off when it’s just a lecture, but it depends a bit on the teacher we have.” Oliver echoed Sebastian’s response to the same question:

I feel it’s easy to drift off when it’s just a lecture, but it depends a bit on the teacher we have. How the teacher engages the students, and

how they manage to make what they are talking about interesting. I feel that sometimes it’s just a lot of lecturing, and then it’s easy to zone out.

The students also reported that lectures could be an arena for effective learning under given conditions. Learning during lectures occurred when the teachers presented new theories and described issues in a concise and understandable way. Furthermore, the importance of including students in discussions in the lectures was a crucial element, which not only assisted the students to be able to remain focused, but also increased their motivation and enjoyment at school. Monica describes this as follows:

If the lecture is a one-way communication, then it works poorly, I think. When the lecturer involves the students with questions, or lets us discuss issues with our peer students, it enhances our concentration and motivation.

This experience was supported by George, who emphasized that discussions during lectures increased learning:

When there are discussions in the lectures, I feel it is easier to understand the themes. Often, students in the class have questions for the lecturer that are interesting and make me think differently than I had without those questions. When we have discussions, I believe that I remember better than if it had been a one-way communication.

In addition, Kevin highlights the importance of having discussions during lectures to create more learning:

I think discussions in lectures have the potential to create very good learning. It could give us an opportunity to evaluate if the aspects the lecturer talks about are facts or part of the lecturer’s opinions. I think it is very favorable when it is possible to exchange knowledge and experiences which leads to better understanding.

The analyses of the students’ reflections about why there is low motivation in lectures among some students revealed that the lecturer must take more responsibility for motivating the students through creating more interesting lessons. All the students described student involvement in the lectures as positive for learning outcomes and stated that such involvement was critical to increase the student’s motivation in lectures. In addition, some of the students responded that it was essential for the lecture to be at an appropriate academic level and concern practice to enhance their learning and remain focused. This took place when the students felt that they could relate the presented theory in lectures to something that they already knew, and that they could understand the relevance of the themes to their future employment as teachers.

Lectures are often long and unstimulating

The second main finding from the analyses of the interview data was that the students often considered lectures to be long and unstimulating. This is exemplified by a statement by Susan, who describes her experience of lectures as follows:

Some of the lectures are very long, slow, and monotonous. Where a PowerPoint presentation with white backgrounds with a lot of text was used, and some of the lecturers just read what is written on the PowerPoint. These factors make it difficult to maintain concentration.

Indeed, the ability to stay focused during long and unstimulating lectures was identified as a substantial challenge among all the students. This challenge was especially present if the lecture was at a high scientific level, which made it difficult to understand, or when the students felt that the lecture had little relevance to their future employment as a teacher. In long and unstimulating lectures with little student involvement, Deborah identified some elements that could lead to low concentration:

In lectures, we have our computers right in front of us, and one touch could take us everywhere. If teachers had been behind us in the classroom, they would have been surprised about what was happening on the students' computers during the lectures. Furthermore, sometimes peer students talk, and I do not hear what the teacher is saying.

The students also stated that lectures should be just one of several pedagogical methods used, and when it was combined with other pedagogical methods, learning outcomes became markedly better. Oliver explains this:

In my opinion, it is okay to have some lectures if they are combined with some practical activities at least once a week. If there are three weeks with only lectures and one practical activity, and then three weeks with lectures, I do not think we remember so much of the themes.

Students' explanations of their relatively low learning outcomes in lectures concerned more than only long and unstimulating lectures. Some of the students highlighted that they were more interested in being active than sitting still in an auditorium; after sitting in a classroom and listening to their teachers for 13 years of education prior to starting university, they were highly frustrated with this activity. Other students indicated low motivation and effort for studying as possible reasons for low learning outcomes in lectures. Confronting the low learning score in lectures, Sebastian pointed to a traditional gender-related view:

Since there are more male than female PE students in our course, and males are more interested in being active than females, this could be an explanation of the low score for learning outcomes in lectures.

Practical activities – to relate theory to practice

The statistical analyses of the quantitative data revealed that the students perceived significantly higher learning outcomes from practical activities related to themes in the lectures compared to most other pedagogical methods, except for preparing to write an academic

text. The third main finding was that students felt that it was important to be able to relate theory to practice. When presented with the results showing that practical activities about themes created the most learning, all students used statements as “of course,” and explained that this method made them relate the theory to practice. Sebastian described how he learned from practical activities presented at the university: “When I read about theory, it is just a theory, and the practical activities give us the opportunity to see how they work.” Monica concurred that theoretical themes must be focused on in practical activities:

When we have practical activities, it is vital that the lecturer focus on the theoretical themes, asking the students relevant questions, giving feedback and planning for reflections amongst the students on relevant aspects of the theme.

When Susan was asked about her thoughts on her own learning through practical activities, she replied:

I think that's great, because it's very nice to get a hands-on approach to something you have just had a theoretical approach to, and it's also having a lot to do with our approach to the practical aspect being not just activity, but it's much better to pause, talk and discuss.

Several students emphasized the need for the lecturer to allocate ample time for feedback and discussions from practical activities to improve learning outcomes. Additionally, students noted that the knowledge applied in practical activities was more memorable than that merely discussed in lectures. All the students agreed on the critical importance of a clear link between the theoretical concepts covered in lectures and the practical activities within the same course. They clearly communicated that without this connection, little learning would occur. This is well highlighted in Kevin's statement:

I learn more from the practical aspect, both here when you have hands-on experience and in high school when we had practical classes. When you sit down and do things, you notice what goes wrong more than when you only write about it theoretically. It's difficult to see what's going wrong if you are just supposed to write a plan and explain. Then, it can seem quite good in your head, but be poor in practice. I like to do things, so it's generally about getting to do it, the practical aspect.

When Kevin was asked to discuss a time when he felt that he received good learning from what occurred at school, he said: “It's mostly the lessons where we either have practical teaching first and then we link it to theory afterwards, or the other way around.” Deborah also stated: “Physical education is a practical subject, so it does not matter how much you have read about basketball or volleyball if you have not tried to play it”.

Practical activities are about future employment

The fourth main finding from the interview data analysis revealed that students identified practical activities as yielding the highest

learning outcomes and ascribed this to their relevance to future work. Practical activities were a part of the schedule in this actual course. Often, the students received the responsibility to plan and execute lessons for their peer students based on relevant theory. Susan explains her experience in the following way:

When we plan and execute teaching for our peer students, it is a very good and relevant training for us who are future teachers. I experience good learning from this task, where it helps me to evaluate and understand the theory better. We have had several practical activities about the lectures this semester, and it has helped my learning a lot.

Other students confirmed that when they received actual training on aspects that were directly related to their future employment, the learning outcomes were excellent. They argued that, in fact, there could not be too much practice of the teacher-role within their education, and that this strategy assisted them to successfully combine theory and practice. Kevin exemplifies this:

An advantage is that then you see the totality, the lectures will not be too long, and you will be told quite thoroughly what we are going to go through, and then you will do the practical part. If we create a teaching plan or something else in a group or alone, and then go into the hall and carry it out, it will be exactly what we will be doing when we get out as teachers. I really like this strategy. I learn a lot and am very much in favor of it being very practical [...]. You simulate exactly what you will do at school as teachers.

When Deborah was informed that the students reported the most learning through practical activities, she replied: “Of course. It is about doing what you will do as future teachers at school.” When Oliver was asked to discuss a time when he felt that he received good learning from school, he stated:

It is most often when you are out in practice. In our practice, we visited someone in the city with a disability, and it was very educational. When you get out and actually experience what is being talked about, instead of just talking about things.

Practical activities – the importance of being active

The fifth and final major insight from the interview data analysis was that students viewed practical activities as a platform for active participation, rather than merely sitting passively and listening to the teacher. Confronted with the statistical analyses showing that practical activities were reported to produce the highest learning outcome, none of the students were surprised by this result because it provided them with the opportunity to be active. George explained:

I think when we have practical exercises about issues we have discussed in lecture, we remember the principles better. Furthermore, it is motivating. We have started in this field because

we enjoy physical activity. So, when we get the opportunity to do something actively, we are happy.

Sebastian echoed George’s sentiment, and stated, “I think I learn more from practical activities because I enjoy being active, while I do not learn just as much sitting in a lecture because I want to do things.” Kevin remarked, “It is quite clear that it is motivating to be active yourself versus sitting for 4–6 h. So, it pays off well, I like it. First, I hear it [the theory], and then I feel it in my body.” When Susan was asked what advantages and disadvantages, she sees in practical activities compared to lectures, she answered:

The advantages are that you feel it in your body after hearing about it. So, it is more motivating to be in motion than to sit still, and there is an opportunity for interaction with others. I do not know if there are so many disadvantages [...]. For example, if it’s dancing and you hate dancing, then I’d say it’s a disadvantage. But that sounds like a rarity for me, who likes all-round activity, so it would take a lot for there to be some practical disadvantages.

The interview data revealed that numerous students faced challenges with dyslexia and maintaining concentration during reading and listening to lectures. Therefore, the efficacy of practical activities was very high for them. Additionally, some students noted that incorporating practical activities into the schedule enhanced social interactions and fostered greater engagement between students and the lecturer.

Discussion

The main findings of the statistical analyses were that practical exercises related to lecture themes resulted in the highest learning outcomes according to student reflections, whereas traditional lectures were found to yield the lowest learning outcomes. These findings are in accordance with other literature that evaluates the effectiveness of lectures as a pedagogical approach in university subjects (Freeman et al., 2014; Damsa et al., 2015; Goodman et al., 2018).

We argue, however, that even if the results showed that there were significant differences in students’ reflections about their learning outcomes across various pedagogical methods, the descriptive data revealed a high average rating of learning outcomes for all methods experienced during the semester in a master’s-level PE course (Figure 1). In other words, in general, the students reported very good learning outcomes from practical exercises related to lecture themes and identified good to very good outcomes from all other methods, with the notable exception of lectures. These findings show that the students generally perceived positive learning outcomes from most pedagogical methods.

Lectures as a pedagogical approach

The results indicated that, among the eight pedagogical methods, lectures were identified by students to produce the worst learning outcomes. However, the analyses of students’ reflections revealed that lectures could be led to positive learning outcomes if

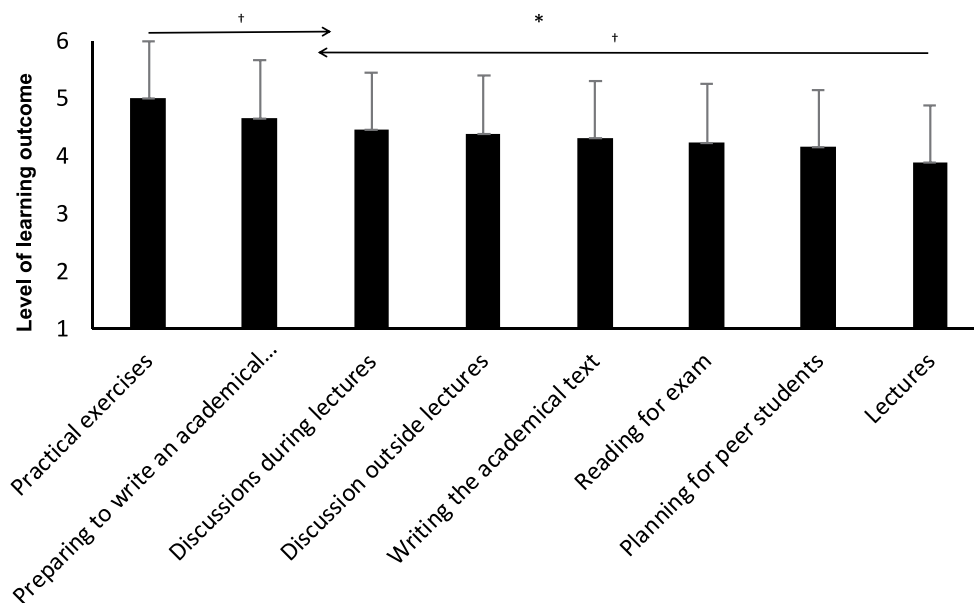


FIGURE 1

Students' reflections of their learning outcomes in eight different pedagogical approaches on a scale between one and six, where six was the highest score. *Significant difference between the groups ($p < 0.001$). † Practical exercises that were significantly different from the approaches are to the right of the arrow ($p < 0.001$). Lectures that were significantly different from the approaches are to the left of the arrow ($p < 0.001$).

certain conditions were fulfilled. These findings are in accordance with Miller et al. (2013), who reported that when the lecturer organized small “breaks” in the lectures by involving the students in different activities, this resulted in several positive outcomes (Miller et al., 2013). Furthermore, all of the students agreed that the efficacy of lectures depends on the characteristics and skills of the lecturer, which is consistent with the results of a previous study (Bates et al., 2018). They stated that some of the lectures were more interesting to listen to than others, which could be ascribed to the theme and the student's interest in it, as well as to how the presentation was made. The students also expressed appreciation for lectures if new theories were presented in an understandable manner, and when the relevance to their future teaching careers was evident. Indeed, the PE students in the present study shared the opinion that lectures play a valuable role in universities (Bates et al., 2018). However, they also highlighted the importance to their learning outcomes of how the presentations in lectures were organized. All the students, however, agreed that greater student involvement, engendered by the lecturer, leads to improved motivation and learning. These findings are similar to those of other investigations focused on how the effectiveness of lectures can be increased (Brown and Bakhtar, 1988; Al-Modhefer and Roe, 2009).

The analyses of in-depth interview data revealed that the students often found lectures to be long and unstimulating. This finding is supported by other studies that assess the efficacy of lecturers and find that lectures often make students into passive listeners (Young et al., 2009; Wieman, 2019). One explanation for this may be that the lecturers, who are often relatively more focused on research, may not always be adequately prepared for teaching (Vabø and Ramberg, 2009). Another factor could be the increasing diversity among

students at universities currently (Damsa et al., 2015), in which the heterogeneity in their motivation and prior knowledge levels could influence their learning outcomes in lectures. In fact, this argument was expressed by Kevin in the Results section. In accordance with Benneworth et al. (2016), our findings identify an important challenge with teaching at universities: there are no methods that suit all students. For example, lectures conducted at a high academic level cater to certain students, but not all students. This might stem from teachers perceiving an overwhelming number of outcomes, so that they choose to present continuous lectures that are replete with advanced academic content, as suggested by Lujan and Dicarlo (2006).

In the Results section, Deborah pointed towards the possibility of students to divert their attention away from lectures, instead focusing on other activities on their computers. A similar finding was reported by Miller (2013), who found that students were often unengaged in lectures, and chose instead to send emails, send text messages, or engage in other activities on their computers.

During the interviews, the students suggested several solutions to prevent lectures from being long and unstimulating, emphasizing the importance of discussions and integrating lectures with practical activities. Furthermore, the students reported that when lectures facilitate interactions between the lecturer and students, the learning will increase and certain issues, such as low motivation and boredom, will decrease. In aggregate, we argue that while lectures can constitute an effective pedagogical strategy that offers substantial learning benefits, it often fails to do so due to its protracted nature. This is supported by another study in which a significant number of university students reported experiencing boredom in lectures, with 59% feeling bored at least half of the time and 30% feeling bored most or all of the time (Mann and Robinson, 2009).

Practical activities as a pedagogical approach

The statistical analyses showed that among the eight pedagogical methods, practical exercises were rated as having the highest learning outcomes according to student reflections. The qualitative analyses of in-depth interview data suggested three reasons for this: (1) the ability to relate theory to practice; (2) the relevance of practical activities to future work; and (3) the active engagement that these activities promote among students. Many students emphasized the importance of applying theory and practical context and seeing firsthand how theoretical concepts function in real-life scenarios. We contend that these findings underscore one of the major challenges in education: the necessity of bridging the gap between academic learning and its practical application in real life. Without this connection, the true value of education could be dubious. This finding closely aligns with Dewey (1916), who asserts that linking theory to practice imbues the learning process with purpose. Dewey emphasizes the necessity of theory being grounded in practice, and posits that effective learning is challenging without engaging the student and building upon his or her prior experiences and expectations. In accordance with Dewey's philosophy, students should engage with theories that they find both relevant and interesting to achieve optimal learning outcomes. Even if the theory of Dewey was created long time ago, previous research have shown that Dewey's pedagogy and philosophy still play a role in schools' daily practice (Berg, 2013; Little, 2013; Lagestad, 2014; Sadovnik et al., 2017).

In relation to connecting theory to practice, Biggs and Tang (2011) identify two crucial factors for effective student learning. The first factor is the perceived value of the learning material to the learner. A purely theoretical approach can diminish the perceived value, as explained by Sebastian. In this study, all students reported positive learning outcomes of a combination of lectures and practical activities related to the lecture topic. They emphasized the necessity of a direct link between the two pedagogical approaches to facilitate meaningful learning experiences. Students also noted that focusing on the same theme made the material both theoretically and practically memorable. In addition, the students underscored the importance of an active role for the lecturer in determining the learning outcome. It is crucial for the lecturer to design practical activities with sufficient time allocated for posing relevant questions to students, providing feedback, and facilitating reflections on the connection between theory and real-life applications. Indeed, this approach is in line with the educational strategies supported by Dewey (2008).

Our discovery that students value diverse pedagogical approaches for the same theoretical concept is in accordance with findings from extant literature (Penson, 2012; Regmi, 2012). Based on student responses in interviews, incorporating practical activities as part of an active learning strategy enhances their engagement, involvement, and depth of learning. Similar findings are also reported in several previous research papers (Lea et al., 2003; McCabe and O'Connor, 2014; Sørensen et al., 2023).

In the context of this study, students were tasked with planning and delivering lessons to their peers on pertinent theoretical subjects. The students viewed this as valuable preparation for their future

teaching careers. Moreover, it provided them with an opportunity to critically assess and gain a deeper understanding of the theories involved (Dewey, 2008). The participants in this study were PE students, suggesting a natural inclination toward practical activities and sports. As one of the students responded, "When I enjoy being active, it helps me to learn better from being active." The findings that students appreciate practical activities related to future work, is also reported in other studies (Lagestad, 2014).

Final reflections

According to Biggs and Tang (2011), the main goal for teachers and lecturers globally is to assist students to achieve their full learning potential. Unfortunately, no definitive formula for effective teaching currently exists. In fact, the discussion on enhancing student activity through varied teaching methods to improve learning outcomes has been ongoing for decades, as evidenced by studies from Young et al. (2009) and Wieman (2019). Research indicates that lectures that lack depth can result in passive and unmotivated students, while overburdening students with responsibility in education can also produce adverse effects (Bremner, 2019). Consequently, pedagogical experts in higher education recommend employing a combination of different approaches to achieve an effective balance in teaching (Lea et al., 2003; Bremner, 2019).

The main findings of our study indicate that while students acknowledge that lectures have the potential to facilitate good learning, they often fall short of achieving this. In contrast, practical activities that are theory-based were found to be effective in learning, relevant for future teaching careers, and enjoyable. We recommend that lecturers in other subjects perform similar experimental research to better understand student reflections on learning through diverse teaching methods and approaches. Indeed, this advice is in accordance with Biesta (2013), who describes learning as a risky endeavor without a fail-safe "recipe" for success. We also argue that our findings resonate with the ideas proposed by Regmi (2012), who emphasizes that the debate should not center on whether to use lectures as a teaching method, but rather on how to structure teaching using a range of methods to optimize learning outcomes.

Strength and limitations of the study

The study has a mixed-methods design, in which quantitative data about eight traditional pedagogical methods are followed-up by in-depth qualitative data about these methods, which is a notable strength. Furthermore, the quantitative data encompass students from three distinct year groups, which adds diversity to the findings. However, this study is not without certain limitations. From a critical perspective, it is possible that the PE students may have overrated practical exercises due to the enjoyment that they derived from these activities. On the other hand, considering that the participants were in their third year of teacher education, they likely possess a more informed perspective on their learning development, supported by both their teaching experiences and their role as educators in high schools and colleges.

Moreover, we suggest that students' capacity to accurately assess their learning outcomes in relation to different pedagogical methods might be challenging. For instance, in a study in which actual learning outcomes were compared with perceived learning, students felt that they learned more from lectures than student-active methods. However, when assessing the actual learning outcomes, it was found that student-active methods were more effective than lectures (Deslauriers et al., 2019). This discrepancy highlights the complexity of measuring and understanding learning outcomes, and the importance of considering both perceived and actual learning in pedagogical strategies.

Conclusion

This study investigated students' perceptions of their learning outcomes from eight traditional teaching methods in a masters-level PE course, utilizing a mixed-methods approach. The quantitative analysis revealed that practical exercises resulted in the highest learning outcomes, while lectures were found to yield the lowest. The qualitative analysis, which involved in-depth interviews focusing on these two teaching methods, showed the importance of students connecting theory with practice. Additionally, these activities were perceived as more relevant to future professional roles compared to other methods, and they encouraged active student participation, which was viewed positively by PE students. Lectures were frequently perceived as long and unstimulating, largely due to students feeling that many instructors failed to engage them in discussions. Considering the findings, it is suggested that student teachers obtain the ability to plan and implement practice lessons that are related to themes in lectures, and actively involve students in discussions during lecture sessions, in order substantially enhance engagement and learning.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found in the article/supplementary material.

References

- Al-Modhefer, A.-K., and Roe, S. (2009). Nursing students' attitudes to biomedical science lectures. *Nurs. Stand.* 24, 42–48. doi: 10.7748/ns.24.14.42.s47
- Bates, M., Curtis, S., and Dismore, H. (2018). Learning approaches and lecture attendance of medical students. *J. Further High. Educ.* 42, 248–258. doi: 10.1080/0309877X.2016.1261089
- Benneworth, P., Pinheiro, R., and Sánchez-Barrioluengo, M. (2016). One size does not fit all New perspectives on the university in the social knowledge economy. *Sci. Public Policy.* 43, 731–735. doi: 10.1093/scipol/scw018
- Berg, P. (2013). Exploring activities and pedagogy reflective of progressive education (order no. 3592449). Available from ProQuest one academic. (1437003890). Available at: <https://login.ezproxy.nord.no/login?url=https://www.proquest.com/dissertations-theses/exploring-activities-pedagogy-reflective/docview/1437003890/se-2>
- Bernstein, D. A. (2018). Does active learning work? A good question, but not the right one. *Scholarsh. Teach. Learn. Psychol.* 4, 290–307. doi: 10.1037/stl0000124
- Biesta, G. J. J. (2013). Beautiful risk of education. Taylor & Francis Group. Available at: <http://ebookcentral.proquest.com/lib/nord/detail.action?docID=4185950>
- Biggs, J., and Tang, C. (2011). *Teaching for Quality Learning at University*. McGraw-Hill Education. Available at: <http://ebookcentral.proquest.com/lib/nord/detail.action?docID=798265>
- Børte, K., Nesje, K., and Lillejord, S. (2020). Barriers to student active learning in higher education. *Teach. High. Educ.*, 28, 1–19. doi: 10.1080/13562517.2020.1839746
- Bremner, N. (2019). From learner-centred to learning-centred: becoming a 'hybrid' practitioner. *Int. J. Educ. Res.* 97, 53–64. doi: 10.1016/j.ijer.2019.06.012
- Brinkmann, S., and Kvale, S. (2009). *Interviews: learning the craft of qualitative research interviewing*. Los Angeles, LA, USA.
- Brown, G., and Bakhtar, M. (1988). Styles of lecturing: a study and its implications. *Res. Pap. Educ.* 3, 131–153. doi: 10.1080/0267152880030204
- Damsa, C., De Lange, T., Elken, M., Esterhazy, R., Fosslund, T., and Al, E. (2015). *Quality in Norwegian higher education: a review of research on aspects affecting student learning*. Available at: <http://hdl.handle.net/11250/2360199>
- Deslauriers, L., McCarty, L., Miller, K., Callaghan, K., and Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proc. Natl. Acad. Sci.* 116, 19251–19257. doi: 10.1073/pnas.1821936116
- Dewey, J. (1916). *Democracy and education*. Free Press.
- Dewey, J. (2008). *Democracy and education: an introduction to the philosophy of education*. Champaign, Ill: Book Jungle.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci U S A* 111, 8410–8415. doi: 10.1073/pnas.1319030111
- French, S., and Kennedy, G. (2017). Reassessing the value of university lectures. *Teach. High. Educ.* 22, 639–654. doi: 10.1080/13562517.2016.1273213

Ethics statement

The studies involving humans were approved by Norwegian Agency for Shared Services in Education and Research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. PL: Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

- Goodman, B. E., Barker, M. K., and Cooke, J. E. (2018). Best practices in active and student-centered learning in physiology classes. *Adv. Physiol. Educ.* 42, 417–423. doi: 10.1152/advan.00064.2018
- Griffiths, R. (2005). Knowledge production and the research-teaching Nexus: the case of the built environment disciplines. *Stud. High. Educ.* 29, 709–726. doi: 10.1080/0307507042000287212
- Hastie, P., and Glotova, O. (2012). "Analyzing qualitative data" in *Research methods in physical education and youth sport*. In The Corsini Encyclopedia of Psychology (4th ed). eds. K. Armour, D. Macdonald, R. B. Holden, I. B. Weiner and W. E. Craighead (London, UK: Routledge).
- Holden, R. B. (2010). "Face Validity" in *The Corsini Encyclopedia of Psychology*. (4th ed). eds. I. B. Weiner and W. E. Craighead (USA, Wiley: Hoboken, NJ). doi: 10.1002/9780470479216.corpsy0341
- Johannessen, A., Tufte, P. A., and Christoffersen, L. (2016). *Introduksjon til samfunnsvitenskapelig metode*. Oslo, Norge: Abstrakt forlag A/S.
- Knight, J. K., and Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biol. Educ.* 4, 298–310. doi: 10.1187/05-06-0082
- Kolb, D. A. (1984). *Experiential Learning: Experience as the source of learning and development*. Englewood Cliffs, NJ.
- Kvale, S. (1983). The qualitative research interview. *J. Phenomenol. Psychol.* 14, 171–196. doi: 10.1163/156916283X00090
- Lea, S. J., Stephenson, D., and Troy, J. (2003). Higher education Students' attitudes to student-centred learning: beyond 'educational bulimia'? *Stud. High. Educ.* 28, 321–334. doi: 10.1080/03075070309293
- Lagestad, S. (2014). The importance of relating theory and practice when teaching police students. *J. Police Studies/Cahiers Politiques*. 1.
- Little, T. (2013). 21st century learning and progressive education: an intersection. *Int. J. Progress. Educ.* 9, 84–96.
- Lujan, H. L., and Dicarolo, S. E. (2006). Too much teaching, not enough learning: what is the solution? *Adv. Physiol. Educ.* 30, 17–22. doi: 10.1152/advan.00061.2005
- Mann, S., and Robinson, A. (2009). Boredom in the lecture theatre: an investigation into the contributors, moderators and outcomes of boredom amongst university students. *Br. Educ. Res. J.* 35, 243–258. doi: 10.1080/01411920802042911
- McCabe, A., and O'Connor, U. (2014). Student-centred learning: the role and responsibility of the lecturer. *Teach. High. Educ.* 19, 350–359. doi: 10.1080/13562517.2013.860111
- Meld. St. 16 (2016–2017) Kultur for høyere utdanning. Kunnskapsdepartementet. Available at: <https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/>
- Miller, C. J., McNear, J., and Metz, M. J. (2013). A comparison of traditional and engaging lecture methods in a large, professional-level course. *Adv. Physiol. Educ.* 37, 347–355. doi: 10.1152/advan.00050.2013
- Miller, C. J., and Metz, M. J. (2014). A comparison of professional-level faculty and student perceptions of active learning: its current use, effectiveness, and barriers. *Adv. Physiol. Educ.* 38, 246–252. doi: 10.1152/advan.00014.2014
- Penson, P. E. (2012). Lecturing: a lost art. *Curr. Pharm. Teach. Learn.* 4, 72–76. doi: 10.1016/j.cptl.2011.10.010
- Regmi, K. (2012). A review of teaching methods--lecturing and facilitation in higher education (HE): a summary of the published evidence. *J. Effect. Teach.* 12, 61–76.
- Sadovnik, A. R., Semel, S. F., Coughlan, R. W., Kanze, B., and Tyner-Mullings, A. R. (2017). Progressive education in the 21st century: the enduring influence of John Dewey. *J. Gilded Age Progress. Era* 16, 515–530. doi: 10.1017/S1537781417000378
- Smith, C., and Cardaciotto, L. (2011). Is active learning like broccoli? Student perceptions of active learning in large lecture classes. *J. Scholar. Teach. Learn.* 11, 53–61.
- Sørensen, A., Lagestad, P., and Mikalsen, H. K. (2023). Student teacher experiences of learning and pedagogical involvement using a student-centered learning approach. *Educ. Sci.* 13:965. doi: 10.3390/educsci13090965
- Trinidad, J. (2019). Understanding student-centred learning in higher education: students' and teachers' perceptions, challenges, and cognitive gaps. *J. Furth. High. Educ.* 44, 1–11. doi: 10.1080/0309877X.2019.1636214
- Vabø, A., and Ramberg, I. (2009). "Arbeidsvilkår i norsk forskning". (Oslo, Norway: NIFU).
- Wieman, C. E. (2019). Expertise in university teaching & the implications for teaching effectiveness, Evaluation & Training. *Daedalus* 148, 47–78. doi: 10.1162/daed_a_01760
- Young, M. S., Robinson, S., and Alberts, P. (2009). Students pay attention!: combating the vigilance decrement to improve learning during lectures. *Act. Learn. High. Educ.* 10, 41–55. doi: 10.1177/1469787408100194



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

María Inés Martín-García,
University of Castilla-La Mancha, Spain
Gustavo Cunha de Araujo,
Federal University of North Tocantins (UFNT),
Brazil

*CORRESPONDENCE

José Gijón Puerta
✉ josegp@ugr.es

RECEIVED 20 December 2023

ACCEPTED 19 March 2024

PUBLISHED 18 April 2024

CITATION

Lizarte Simón EJ, Khaled Gijón M,
Galván Malagón MC and Gijón Puerta J
(2024) Challenge-obstacle stressors and
cyberloafing among higher vocational
education students: the moderating role of
smartphone addiction and Maladaptive.
Front. Psychol. 15:1358634.
doi: 10.3389/fpsyg.2024.1358634

COPYRIGHT

© 2024 Lizarte Simón, Khaled Gijón, Galván
Malagón and Gijón Puerta. 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.

Challenge-obstacle stressors and cyberloafing among higher vocational education students: the moderating role of smartphone addiction and Maladaptive

Emilio Jesús Lizarte Simón¹, Meriem Khaled Gijón²,
María Carmen Galván Malagón³ and José Gijón Puerta^{1*}

¹Department of Didactics and School Organization, University of Granada, Granada, Spain,

²Laboratory for Cognition, Health, Training and Interaction Among Humans, Animals and Machines, University of Granada, Granada, Spain, ³Department of English Philology, University of Extremadura, Badajoz, Spain

The start of higher vocational education and training is a new stage for students with a challenge between theoretical classes and the new expectations and demands of companies during the internship period. To understand some of the implications of stress on cyberloafing, we can distinguish between stress perceived as an obstacle that can be overcome – challenge stress – or as a threat that can block work performance – obstacle stress – and stress perceived as an obstacle that can be overcome – challenge stress – or as a threat that can block work performance – obstacle stress-. The aim of this research is to find out the relationships between challenge-obstacle stress in Cyberloafing, as well as the moderating effect of Smartphone Addiction and Maladaptive. In this study, the Challenge-Hindrance Stressors, Smartphone addiction scale-short version (SAS-SV) instrument, the Maladaptive subscale of the Cognitive Emotional Regulation Questionnaire (CERQ) and Cyberloafing were applied to 403 upper-level vocational training students from different secondary schools in all provinces of the autonomous community of Andalusia, Spain, distributed throughout the provinces that make up this autonomous community. The findings show that students' challenge stressors do not increase Cyberloafing, enabling them to cope with the academic demands and work challenges during the theory and internship period. On the contrary, obstacle stressors generate stressful situations that undermine the acquisition of objectives and development of academic competences. In our research we observe that challenge-obstacle stressors have a disparate influence on cyberloafing. Challenge stressors are negatively related to Maladaptive. The same is not true for obstacle stressors.

KEYWORDS

cyberloafing, challenge, hindrance, stress, smartphone addiction

1 Introduction

In recent decades, vocational education and training has been seen as a type of education that, through training for the performance of a professional activity, can help to meet the challenges of digitalization and globalization (Bürge and Gonon, 2021). According to the OECD (2018) *Getting Skills Right: Spain reports on the State of the Education System of the Spanish Ministry of Education and Vocational Training*, the number of jobs generated by digitalization and the ecological transition will be the two major transforming elements of the economic model (Ministerio de Educación y Formación Profesional (MEFP), 2021; Ley Orgánica, 2022). These jobs will need to be filled by competent and professionally qualified people, through training scenarios that connect theoretical knowledge with practical skills and vice versa. Forecasts for Spain in 2025 indicate that 49% of jobs will require intermediate skills and only 14% of jobs will require low skills.

The vocational education and training system has had to be transformed in order to train students through the development of professional competences. These competences facilitate entry into the world of work, which today is characterized by continuous change (Virgós et al., 2022). For this changing world, vocational education is shaped by the rationality of school and the rationality of workplace-based learning (Schaap et al., 2012). The former is related to theoretical knowledge, while the latter is related to practical knowledge (Ferm, 2021). Both knowledge is essential for the development of professional competences (Bouw et al., 2021) and, for this development, a link between theory and practice is required. Theoretical knowledge carries more weight in the field of education, while practical knowledge is more important in the work context (Billett, 2014). In both contexts, however, the use of ICT-supported communication platforms appears in relation to productivity and the challenges of digitalization. This incorporation has changed contemporary organizations (Tandon et al., 2021) and, since its beginning, there has been debate about the benefits and drawbacks that new technologies can bring in relation to their productive or non-productive use during working hours. Thus, we enter into the issue of time spent using technologies for activities that bring personal benefit unrelated to work performance (Batabyal and Bhal, 2020; Tandon et al., 2021), even in the case of students in academic training. This behavior is referred to as Cyberloafing (Lim, 2002; Reizer et al., 2022) and is generally considered to be a counterproductive behavior that can occur during school hours. Taking into consideration the influence of stressors during theoretical and practical classes in vocational training, which can lead to lower academic performance and dropout (Lizarte and Gijón, 2022a,b), we will focus on how stressors are related to cyberloafing and smartphone addiction and maladaptation as mediating variables.

The relevance of the article presented is based on the scarcity of previous studies that focus on higher vocational training students, due to the fragmentation of these teachings in very varied professional areas (professional families in their denomination in Spain), which has been highlighted by studies such as the one carried out by the Bankia Dual Foundation for the period 2005–2017 or the recent review by Toepper et al. (2022), which points to the probable causes of this scarce research in higher VET, which may also be due to the difficulty

of transferring their results to general recommendations to be applied in education systems, due to the high contextualization of these studies.

1.1 Cyberloafing in the classroom

It is a proven fact that technological innovations (multimedia applications, web technologies and next-generation educational applications) enable courses to be delivered more effectively and interactively (Ersoy et al., 2016; Johnson et al., 2023). However, misuse of the internet in the classroom can lead to distraction and disconnection from the lesson the teacher is teaching (Wu and Xie, 2018; Flanigan and Babchuk, 2022). In relation to the above statement, accessing applications via the internet in the classroom can lead to cyberloafing in students (Baturay and Toker, 2015; Akgün, 2020), leading in some cases to lower academic performance, decreased motivation and discipline problems (Junco and Cotten, 2012; Farivar and Richardson, 2018; Gökçearsan et al., 2018). However, some research considers cyberloafing to have certain positive effects, such as stress reduction and relaxation during work sessions and the completion of large amounts of work tasks in the workplace (Yildiz and Saritepeci, 2019). The problem that arises in practice is to be able to quantify the negative or positive balance of cyberloafing for learning, both on a general and individualized level, as well as to develop teaching strategies for the management of cyberloafing as a didactic tool, although this will probably be closely related to the teaching style of the teacher and the context of each subject and educational stage.

1.2 Stress in der Berufsausbildung

In relation to stress, the change involved in the transition from secondary school to vocational training has been studied by a number of authors (Steinmann, 2005; Pusceddu, 2016; Hösli et al., 2018; Böhn and Deutscher, 2022). It has become clear that for many students, this new stage can be more sensitive to stress, due to the cognitive, affective, biological and social changes that occur at this stage of life. Vocational trainees have to fulfil the role of a student during theoretical lessons and the new expectations and demands of the company during the traineeship (Berweger et al., 2013; Leu, 2014). For this reason, it is often more time-consuming for vocational students to regulate their stress in certain academic situations during the internship period (Schulten and Wussler, 2013). The use of different coping strategies will explain individual differences in relation to adapting to the stress required by these new educational demands (Hernández et al., 2023). See Figure 1 to visualise the hypothetical model of relationships that we will describe below.

1.3 Challenge-hindrance stressors and cyberloafing

To understand some implications of stress on cyberloafing, we can distinguish between stress perceived as an obstacle that can be overcome -challenge stress- or as a threat that can block job

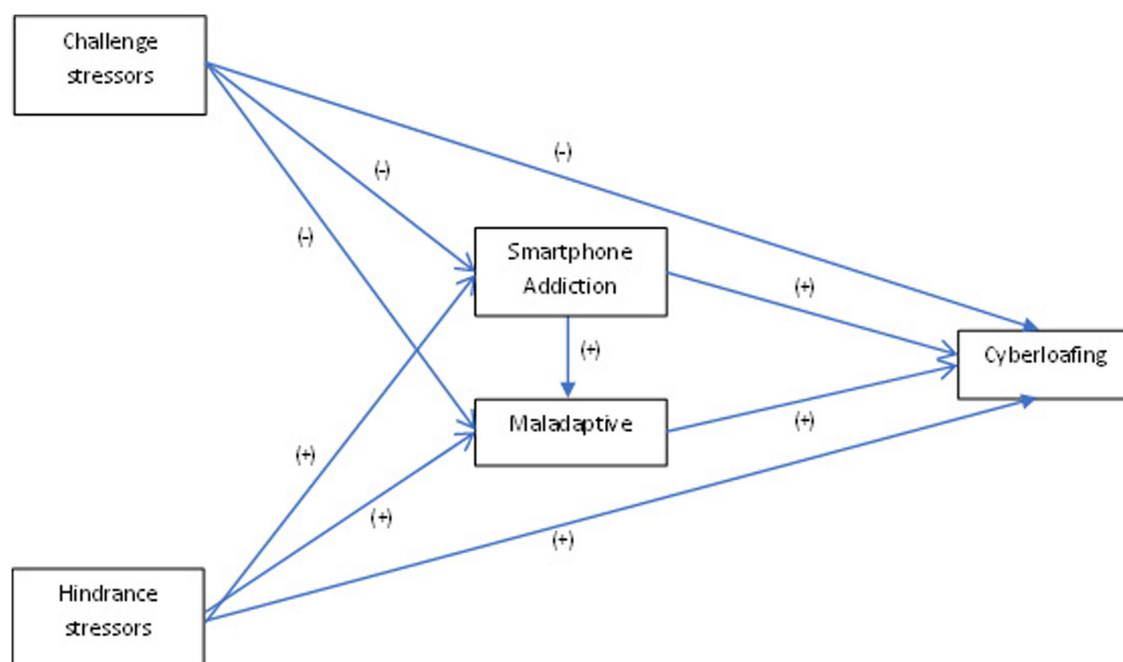


FIGURE 1
Hypothesized model of relationships among stressors, smartphone addiction, Maladaptive, and cyberloafing.

performance -hindrance stress- (LePine et al., 2005; Podsakoff et al., 2007; Karatepe et al., 2018; Afshari et al., 2022). In the first case -challenge-, the effects of this type of stress can help to reconstitute and boost individual growth, including high demands of work, responsibility and time pressure.

Indeed, the achievement of work objectives is related, in part, to the type of stressors that exist in the workplace and their perception. Some studies indicate that challenge stressors are associated with positive employee behaviors, with improved team performance, increased motivation and effort boosting (Kassam et al., 2009; Pearsall et al., 2009). However, obstacle stressors hinder goal achievement, role ambiguity and excessive concern for job security (Kassam et al., 2009; Pearsall et al., 2009; Mazzola and Disselhorst, 2019). Thus, challenging and hindering stressors influence learning, affecting the level of development of professional competences during studies (Lizarte, 2021; Lizarte and Gijón, 2022a). From a general perspective, stress has an adaptive and mainly psychological character that can manifest itself in an imbalance that results in the development of coping or discarding actions in the face of academic demands (Vorontsova et al., 2022), regardless of whether the learning takes place in the classroom or during internships (Gijón, 2020). If stress is severe and prolonged in students, it can reduce academic performance, decrease academic and social integration and may even increase the likelihood of dropping out of school (Amanvermez et al., 2022; Ang et al., 2022; Lizarte and Gijón, 2022a).

In certain situations, cyberloafing can be an isolating behavior (Askew et al., 2019; Tandon et al., 2021), a form of procrastination during lectures (Lowe-Calverley and Grieve, 2017), which can lead to decreased task completion satisfaction (Khansa et al., 2017) and increased emotional exhaustion (Koay and Soh, 2018). Thus, cyberloafing can be used by students as an avoidance strategy to temporarily avoid the stressful demands of certain actions in theory and practice classes. Therefore, we propose the following hypothesis:

H1: Obstacle stressors are positively related to cyberloafing.

Considering that challenge stressors motivate workers to exert effort and deploy their capacity to cope with work demands and goal fulfillment (Umer et al., 2022) and that obstacle stressors generate stress levels that hinder good performance and task achievement (Kassam et al., 2009; Kim and Beehr, 2020), it can be inferred that the teacher's assessment of these situations will lead him/her to select appropriate methodologies to avoid cyberloafing as far as possible. One of the most commonly used methodologies in vocational training is project work, which could be appropriate to combat the negative effects of cyberloafing. This pedagogical approach refers to a didactic methodology by which one tries to acquire and develop competences through the realization of a project elaborated individually or in a group (Project-Based Learning). Project work enables cooperative learning and helps students to develop teamwork skills and improve the level of goal achievement (Lau et al., 2013). Thus, students could be more focused on achieving their goals, spending less time on cyberloafing. We therefore propose the hypothesis:

H2: Challenge stressors will be negatively related to cyberloafing.

1.4 Smartphone addiction and Maladaptive as mediating variables

Both challenge and obstacle stress affect academic performance (Kötter et al., 2017) and the acquisition of vocational competences during vocational studies (Lizarte and Gijón, 2022b). The impact of this ongoing stress during vocational training and the well-being of vocational students has not been sufficiently explored so far.

One of the issues of greatest interest in relation to stress, academic performance and cyberloafing is the use of mobile devices, especially

the smartphone. While the use of smartphones for learning activities provides teaching opportunities, it can also have a negative impact on learning environments and processes (Wagner et al., 2012). Thus, cyberloafing behavior can lead to disengagement with class and reduced motivation (Gökçearsan et al., 2018). Cyberloafing behavior and smartphone addiction might be related as both are linked to addictive behavior (Gökçearsan et al., 2016, 2018) and avoidance of situations that can generate stress by fulfilling learning activities developed in class (Yildiz and Saritepeci, 2019). In this context, we propose the following hypothesis:

H3: Challenge stressors will be negatively related to smartphone addiction.

H4: Obstacle stressors will be positively related to smartphone addiction.

Since stress is related to cyberloafing, a reduction of stress would reduce the reasons for cyberloafing (Lim and Teo, 2005). Considering that hindrance stress increases negative emotions and cyberloafing (Koay et al., 2017), smartphone addiction could be related to cyberloafing (Gökçearsan et al., 2016, 2018). Considering that mobile phone addiction takes the form of uncontrollable use of smartphones (Park and Park 2017; Saritepeci, 2020), we could assume that students with high levels of smartphone addiction during lessons would increase levels of cyberloafing. Students with higher levels of smartphone addiction influenced by higher levels of stress are likely to be less responsive to academic demands, thus cyberloafing may increase. Two hypotheses derive from these statements:

H5: Challenge stressors will gain a positive indirect relationship with cyberloafing, mediated by Smartphone addiction.

H6: Obstacle stressors will gain a positive indirect relationship with cyberloafing, mediated by Smartphone addiction.

H7: Smartphone addiction will be positively related to cyberloafing.

Reactions to stressful situations are associated with Maladaptive rather than adaptive cognitive emotional regulation strategies (Kraaij et al., 2003). When people experience more stressors, they tend to develop Maladaptive cognitive emotional regulation strategies to cope with these stressors (Kraaij et al., 2003). Challenging stressors tend to favor positive behaviors and increase effort (Pearsall et al., 2009), so that self-blame, rumination, catastrophizing and blaming others (Maladaptive subscale) will have low values. However, obstacle stressors would increase the values of the Maladaptive subscale (Wallace et al., 2009; Sawhney and Michel, 2022). These facts allow us to indicate that challenge stressors could reduce the values of the Maladaptive scale and obstacle stressors would increase the values obtained when applying the Maladaptive scale (Wallace et al., 2009; Sawhney and Michel, 2022).

The use of Maladaptive emotional regulation strategies is associated with increased stress symptoms (Garnefski et al., 2001) and the use of adaptive emotional regulation strategies is associated with lower stress and higher self-esteem (Cutuli, 2014). Hence, the use of unadaptive emotional regulation strategies may increase smartphone

addiction (Use this paragraph for analysis and conclusions). We propose:

H8: Challenging stressors will be negatively related to the Maladaptive.

H9: Obstacle stressors will be positively related to the Maladaptive.

H10: Smartphone addiction will be positively related to Maladaptive.

Based on the above, we can assume that increasing the means of access to ICT tools in theory and practice classes may promote cyberloafing behaviors (Akbulut et al., 2017; Saritepeci, 2020). Thus, considering that passive coping tends to be more Maladaptive, it could hinder growth opportunities, lower engagement and performance of students in situations that require some degree of stress for goal achievement (Wallace et al., 2009). Thus, the use of Maladaptive emotional regulation strategies in class would promote cyberloafing (Metin and Demirtepe, 2021).

We therefore propose:

H11: Challenge stressors will gain a positive indirect relationship with cyberloafing, mediated by the Maladaptive.

H12: Obstacle stressors will have a positive indirect relationship with cyberloafing, mediated by the Maladaptive.

H13: Maladaptive will relate positively to cyberloafing.

2 Method

2.1 Participants

The instruments were sent to 44 Secondary Schools teaching Higher Level Training Cycles in all the provinces that make up the Andalusian Autonomous Community during the 2021/22 academic year. A response of 427 questionnaires was obtained, of which 24 were excluded because they were incomplete. The final sample that participated in this study is 403 students of Higher Vocational Training, of which 46.15% are women and 53.85% are men, aged between 16 and 59 years old.

Table 1 shows the descriptive analysis of the demographic variables: in relation to the degree, 40.94% of the students are in the first year – corresponding to theory classes – while 59.06% are in the second year – work experience period in a company. As for the influence on the decision to enroll in Vocational Training, 75.68% did so because of the job opportunities, 16.63% because of the academic program, 4.47% because of friends and only 3.22% because of the possibility of participating in Erasmus+ or internationalization projects.

2.2 Instruments

Obstacle and challenge stressors were measured using Cavanaugh et al. (2000) scale, which consists of 11 items. The

TABLE 1 Descriptive of socio-demographic variables.

	<i>n</i>	%
Qualification		
Theory students	165	40,94
Traineeships students	238	59,06
Sex		
Male	217	53,85
Female	186	46,15
What influenced you most in your decision to enroll in Vocational Training?		
Job opportunities	305	75,68
Its academic program	67	16,63
My friends	18	4,47
Possibility to participate in Erasmus+ projects or internationalization projects.	13	3,22

obstacle stressors consist of 5 items (e.g., “Inability to clearly understand what is expected of me at work”) and the challenge stressors consist of 6 items (e.g., “Time pressures I experience”). Respondents were asked to rate the degree of stress at work for each item on a 5-point Likert scale (1 = not at all stressful, 5 = very stressful). The scores of the stress-challenge and stress-obstacle subscales were computed by averaging the items. Cronbach’s alpha value (Cronbach’s α) of the obstacle stressors was 0.83 and that of the challenge stressors was 0.85.

The Smartphone addiction scale-short version (SAS-SV) instrument by Kwon et al. (2013) is used to measure the level of smartphone addiction. The SAS-SV instrument is composed of 10 items (e.g., “Having a hard time concentrating in class while doing assignments, or while working due to smartphone use”). The questions were answered on a 6-point Likert scale, where 1 is “Strongly disagree” and 6 is “Strongly agree,” according to the self-report. Possible scores range from 10 to 60. Higher overall scores show a higher degree of smartphone addiction (Kwon et al., 2013). Cronbach’s alpha value (Cronbach’s α) for this scale was 0.92.

The Cognitive Emotional Regulation Questionnaire (CERQ) instrument was used to measure the Maladaptive subscale. The CERQ is used to measure nine types of specific cognitive skills. Self-blame, rumination, catastrophizing and blaming others are considered Maladaptive subscales. However, acceptance, positive refocusing, refocus on planning, positive reappraisal, and putting into perspective are considered adaptive subscales (Garnefski et al., 2002). The CERQ is composed of 36 items with 5-point Likert-type responses (1 = almost never to 5 = almost always). Higher scores indicate higher frequency of use of the specific cognitive strategy being measured. In our research only the Maladaptive subscale was adopted. Cronbach’s alpha value (Cronbach’s α) for this subscale was 0.93.

Lim’s (2002) 11-item scale (e.g., “visit sports-related websites”) was used to measure cyberloafing. Respondents were asked to rate their cyberloafing during school hours on a 5-point Likert scale (1 = not at all to 5 = extremely). Cronbach’s alpha value (Cronbach’s α) for this scale was 0.93.

2.3 Procedures

The research presented here was structured in four phases. In the first – instrument selection – the questionnaires indicated in the previous section were tailored and adapted for online application. The questionnaire included an explanation of the research, assuring the anonymity of the participants and the confidentiality of the data. In the second phase – sample design – a criterion sampling was carried out for the application of the adapted online instruments. For this purpose, a list was drawn up of educational centers offering full teaching of higher level training cycles belonging to the eight provinces that make up the autonomous community of the Junta de Andalucía during the 2021/22 academic year. Using the criterion of accessibility, several vocational training teachers were contacted to facilitate the distribution of the instruments among their students and the rest of the teachers in the participating centers. In the third phase – implementation – the instruments were applied via the online platform (end of April 2021). The average time to complete the two questionnaires was 40 min. In the fourth phase -analysis- the data were downloaded from the platform, and different descriptive statistics (mean and standard deviation) and multivariate statistics (structural equation model and multigroup analysis) were applied and the results analyzed.

2.4 Statistical analysis

Data were analyzed using IBM SPSS 25.0 and Amos. To understand the relationships between each construct, structural equation modeling (SEM) was used to examine causal relationships between constructs. Bootstrapping was used to demonstrate mediating effects. Finally, when comparing the differences between whether students are in theory or in practice, a multigroup invariance analysis using z-scores was conducted to test the possible moderating effect of this factor. Differences in z-score values between the constrained and unconstrained model were calculated, and structural trajectories in each subgroup were compared on the basis of the z-scores (Gaskin, 2012).

3 Results

Table 2 outlines the means and standard deviations, the Cronbach’s alpha reliability indices, and the correlations between the scales. In terms of internal consistency, the alpha index values were above 0.80, indicating high reliability.

As can be seen in Table 2, all scales correlate significantly with each other. Specifically, a negative and moderate correlation was observed between Challenge stressors and Smartphone Addiction ($r = -0.33$; $p < 0.01$); a significant negative correlation between Challenge stressors and Maladaptive, although weak ($r = -0.23$; $p < 0.05$); and a significant positive and strong correlation between Challenge stressors and Cyberloafing ($r = -0.55$; $p < 0.001$).

On the other hand, significant positive correlations were found between Challenge stressors and Hindrance stressors of strong size ($r = 0.58$; $p < 0.001$). There was a weak positive significant correlation between Hindrance stressors and Smartphone Addiction ($r = 0.28$;

TABLE 2 Means, standard deviations, reliabilities, and zero-order correlations among variables.

	Mean (SD)	Cronbach's α	1	2	3	4	5
1. Challenge stressors	3.21 (0.77)	0.85	1				
2. Hindrance stressors	2.87 (0.62)	0.83	0.58***	1			
3. Smartphone Addiction	28.04 (9.25)	0.92	−0.33**	0.28**	1		
4. Maladaptive	2.32 (0.70)	0.94	−0.23*	0.31**	0.41***	1	
5. Cyberloafing	3.28 (0.73)	0.93	−0.55***	0.49***	0.37***	0.47**	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

$p < 0.01$); a moderate positive significant correlation between Hindrance stressors and Maladaptive ($r = 0.31$; $p < 0.01$); and a moderate positive significant correlation between Hindrance stressors and Cyberloafing ($r = 0.49$; $p < 0.001$). A significant positive correlation was observed between Smartphone Addiction and moderate-sized Maladaptive ($r = 0.41$; $p < 0.001$); a significant positive correlation between Smartphone Addiction and Cyberloafing ($r = 0.37$; $p < 0.001$) was also moderate. Finally, a significant positive and moderate correlation is found between Maladaptive and Cyberloafing ($r = 0.47$; $p < 0.01$). In general terms, we can observe that the means and standard deviations in the variables are obtained around mean values in all of them: Challenge stressors ($M = 3.21$; $SD = 0.77$); Hindrance stressors ($M = 2.87$; $SD = 0.62$); Smartphone Addiction ($M = 28.04$; $SD = 9.25$); Maladaptive ($M = 2.36$; $SD = 0.945$) y Cyberloafing ($M = 3.28$; $SD = 0.73$).

3.1 Structural model testing

The model was tested through structural equation modeling (SEM), using AMOS 26.0. The relationships between variables and the unstandardized path coefficients of the model are shown in Figure 2. Challenge stressors were negatively related to mobile addiction ($\beta = -0.35$, $SE = 0.16$, $t = -2.15$; $p = 0.032$), as were Maladaptive ($\beta = -0.19$, $SE = 0.09$, $t = -2.01$; $p = 0.045$), supporting Hypotheses 3 and 8, respectively. However, obstacle stressors were positively related to mobile phone addiction ($\beta = 0.55$, $SE = 0.24$, $t = 2.30$; $p = 0.022$) and to Maladaptive ($\beta = 0.23$, $SE = 0.07$, $t = 3.18$; $p = 0.002$) supporting hypotheses 4 and 9.

Smartphone addiction was positively related to Maladaptive ($\beta = 0.39$, $SE = 0.14$, $t = 2.69$; $p = 0.007$), thus supporting hypothesis 10.

On the other hand, both Smartphone addiction ($\beta = 0.53$, $SE = 0.15$, $t = 3.54$; $p < 0.001$) and Maladaptive ($\beta = 0.54$, $SE = 0.24$, $t = 2.28$; $p = 0.023$) were significant predictors of cyberloafing, supporting hypotheses 7 and 13. The direct pathway from challenge stressors to cyberloafing ($\beta = -0.43$, $SE = 0.16$, $t = -6.90$; $p < 0.001$) was significantly negative, thus not supporting hypothesis 2, while the direct route from obstacle stressors to cyberloafing ($\beta = 0.41$, $SE = 0.07$, $t = 6.17$; $p < 0.001$) was significantly positive, thus supporting hypothesis 1.

The fact that the direct effects of challenge and hindrance stressors on mobile addiction and cyberloafing were significant supported the possibility of an indirect effect. Bias-corrected bootstrapping was used to test for an indirect effect (Preacher and Hayes, 2008; Hayes et al., 2011). The results indicate that mobile phone addiction mediates the relationship between stressors and

cyberloafing. The indirect effects of challenge and hindrance stressors on cyberloafing through mobile addiction were significant (for challenge stressors, indirect effect = -0.19 , 95% CI [0.03, 0.32] supporting hypothesis 5; for hindrance stressors, indirect effect = 0.29 , 95% CI [0.07, 0.43] supporting hypothesis 6). Similarly, the fact that the significant direct effects of challenge and hindrance stressors on Maladaptive and Maladaptive on cyberloafing also suggested the possibility of an indirect effect. Thus, the Maladaptive measured the relationship between stressors and cyberloafing. The indirect effects of challenge and hindrance stressors on cyberloafing via the Maladaptive were significant (for challenge stressors, indirect effect = -0.10 , 95% CI [0.01, 0.21] supporting hypothesis 11; for hindrance stressors, indirect effect = 0.12 , 95% CI [0.01, 0.23] supporting hypothesis 12).

3.2 Multigroup analysis

Multigroup analysis was performed to compare the differences in the coefficients of the trajectories according to whether the students are in theory or in practice, using the z-score comparison (Table 3). The result reveals that the relationship between Maladaptive and cyberloafing is positive in students who are in the theoretical part while in students who are in the practical part this relationship is negative, this difference being statistically significant.

4 Discussion and conclusions

With this research we have been able to analyze the relationship between challenge stressors, obstacle stressors, smartphone addiction, maladaptation and Cyberloafing in higher vocational training students. The mediating role of smartphone addiction and maladaptation in the relationship between challenge-obstacle stressors and Cyberloafing was analyzed. Findings suggest that students' challenge stressors do not increase Cyberloafing, allowing them to cope with academic demands and work challenges during the theory and internship period. In contrast, obstacle stressors generate stressful situations that impair the acquisition of academic goals and the development of academic skills. In our research we observed that in vocational students, challenge-obstacle stressors have a disparate influence on cyberloafing.

Evidence was obtained that challenge stressors are negatively related to Smartphone addiction, taking into consideration that there are several methodologies that involve the use of Smartphones as a didactic resource, and smartphone addiction is related to cyberloafing.

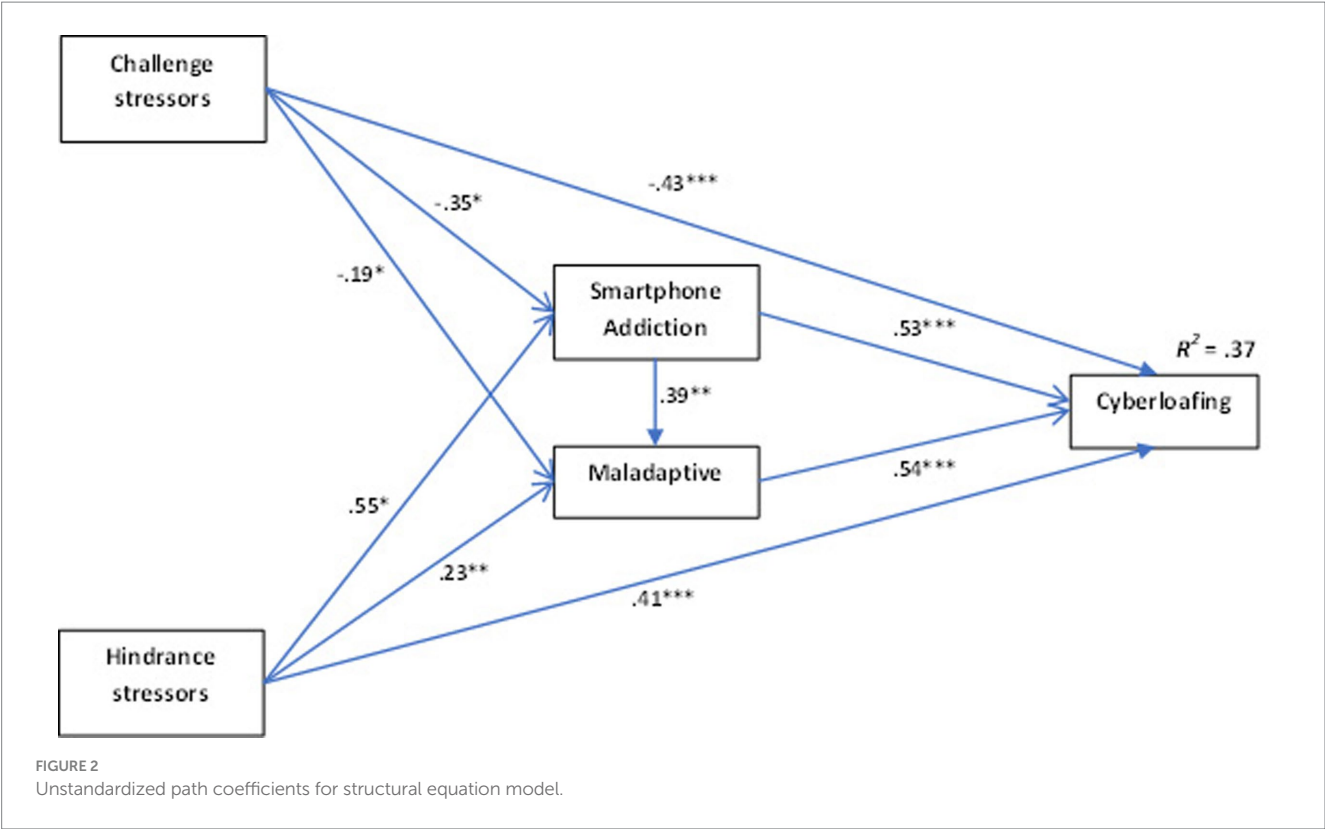


TABLE 3 Multi-group analysis regression coefficients.

	Grupo		
	Teoría	Práctica	z-score
CS--->CYB	−0.37***	−0.40***	−0.57
CS--->SMA	−0.31*	−0.29*	0.32
CS--->MAL	−0.17*	−0.15*	0.29
HS--->CYB	0.35**	0.37***	0.26
HS--->SMA	0.51*	0.49*	−0.27
HS--->MAL	0.15***	0.18**	0.43
SMA--->CYB	0.47***	0.44***	−0.44
MAL--->CYB	0.41**	−0.18**	−5.36***
SMA--->MAL	0.32**	0.34**	−0.29

CS, Challenge stressors; HS, Hindrance stressors; SMA, Smartphone Addiction; MAL, Maladaptive; CYB, Cyberloafing. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-score: critical ratio of differences between parameters.

Challenge stressors were negatively related to Smartphone addiction (3). Furthermore, challenge stressors obtain a positive indirect relationship with cyberloafing, mediated by Smartphone addiction (5). The findings allow us to suggest that challenge stressors have a similar behavior to that found in studies conducted with employees (Agarwal and Avey, 2020; Zhou et al., 2021); thus, we could suppose that this type of stress would allow reinforcing the commitment and effort to meet academic and work challenges, reducing the levels of Cyberloafing (Kunzelmann and Rigotti, 2021).

The results of our research showed that, unlike challenge stressors, obstacle stressors are positively related to cyberloafing (h4), indicating that this type of stress causes an increase in cyberloafing as a

temporary avoidance response to certain academic demands. Obstacle stressors have an indirect positive relationship with cyberloafing, mediated by smartphone addiction (6). This could lead to reduced academic performance and low academic and social integration. It is well known that smartphones are changing the system of relationships, lifestyles and ways of learning (Karaoglan Yilmaz et al., 2023). However, incorrect use can have a negative impact on the learning process (Wagner et al., 2012). Thus, addictive behavior and avoidance of stressful obstacle situations lead to increased levels of Smartphone addiction and Cyberloafing. As a result of this finding, it is necessary to regulate didactic planning so that the use of smartphones does not promote addiction. It must be considered that the dependence on smartphones has increased in the school environment. The smartphone has become a working tool for both teacher and pupil, and its use has become a commonly accepted habit for the teacher's decision making in monitoring academic tasks. All this leads us to believe that the learner will become increasingly dependent on smartphones for school work. It is clear that educational programs on the responsible use of mobile phones need to continue to be implemented for students in stages prior to higher vocational education.

Our findings indicate that challenging stressors are negatively related to Maladaptive, as they tend to favor positive behaviors and effort. This relationship indicates a reduction in the values of self-blame, rumination, catastrophizing and blaming others. The same is not true for hindering stressors, increasing the use of Maladaptive emotional regulation strategies. Maladaptation leads to a further strengthening of smartphone addiction. As a result of our research, we observed that Maladaptive is higher when high values of Cyberloafing are registered; therefore, the use of Maladaptive

emotional regulation strategies corresponds to avoidance behaviors that could favor Cyberloafing. Another important finding in this study, through multigrade analysis, is that we obtained a significant negative relationship between Maladaptive and Cyberloafing in internship students: higher levels of Maladaptive lower levels of Cyberloafing. During the internship period, students perceive the challenges differently than in theory class.

Thus, the perception that internship students have of the academic and work challenges or demands favor positive behavior and effort. We should bear in mind, firstly, that motivation at this final point of training may be of great importance, as the student will present a solid state of mind, which would allow a higher level of adaptation to achieve the academic objectives in a more effective and efficient way. Secondly, related to the previous statement, we can indicate that the vocational training student has chosen this training pathway oriented toward a more immediate employment opportunity, which maximizes the relationship between academic effort and the result of professional insertion. This choice is the majority of all students enrolled in higher vocational training. Despite what has been indicated in relation to trainees, some studies carried out with these students during the pandemic have shown the existence of maladjustment in an irregular way, with new forms of maladjustment appearing that had not previously been clearly established, such as the practice of being on the phone and ignoring those around them (phubbing).

Finally, we can indicate that our research also allows for a number of practical reflections on challenging stressors, which could mobilize the development of academic and professional competences. It should be noted that, while this type of stress provokes positive behaviors, it would also encourage innovative behavior in challenging situations (Ren and Zhang, 2015). These innovative behaviors can relate to the learner, the teacher, the educational program and the organization of the school institution. As for the student, learning to manage stressful situations – i.e. developing soft skills related to stress management – can be a key element for academic success and be projected as an added value for their insertion into the world of work. Related to this learning, the development in the classroom of activities that put students in the situation of having to face challenging stressors and regulate them, will allow the teacher to promote competences related to stress management. This should take the form of the design and application of activities whose learning outcomes include the aforementioned soft competences. Finally, the very organization of spaces and times in the educational institution can also be oriented toward the development of stress management skills, with spaces for relaxation or the development of actions that include activities such as Dog Assisted Therapy in pet-friendly institutions, to reduce stress at specific times, such as exam periods.

4.1 Limitations and options for future research

The study has several limitations. Firstly, the sample consists of higher vocational students from a limited number of vocational families. It will be necessary in the future to expand the number of vocational families in order to carry out comparative analyses. In addition, research at other levels of vocational education, such as intermediate vocational training, may give a more complete picture of the role of stressors in academic success. An additional limitation may

derive from the timing of the study – during the pandemic – when personal and academic relationships were disrupted by the excessive use of smartphones, due to confinement and virtual learning. At the present time, the relationship system may have changed with the return to normality, with personal relationships recovering in the face of mobile device use. However, this needs to be investigated, to find out whether the pandemic has been a break in the normal development of educational programs or, on the contrary, has substantially changed the system of relationships, teaching methodologies and modes of learning, as well as its influence on levels of academic stress and its management.

Data availability statement

The datasets presented in this article are not readily available because the full dataset is not available for this study. Contact the first (EL) or last author (JG) for access to specific variables. Requests to access the datasets should be directed to josegp@ugr.es.

Ethics statement

The studies involving humans were approved by the Ethics Committee in Human Research of the University of Granada were followed 2758CEIH2022. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

EL: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft. MKG: Conceptualization, Investigation, Supervision, Validation, Visualization, Writing – review & editing. MCG: Conceptualization, Data curation, Project administration, Visualization, Writing – review & editing. JG: Conceptualization, Investigation, Software, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This article is part of the research project Design of an instrument for the detection of the risk of early dropout in students of Higher Vocational Training, PPJIA2020.14, funded by the University of Granada.

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

References

- Afshari, L., Hayat, A., Ramachandran, K. K., Bartram, T., and Balakrishnan, B. K. P. D. (2022). Threat or opportunity: accelerated job demands during COVID-19 pandemic. *Pers. Rev.* 51, 2482–2501. doi: 10.1108/pr-02-2021-0098
- Agarwal, U. A., and Avey, J. B. (2020). Abusive supervisors and employees who cyberloaf: examining the roles of psychological capital and contract breach. *Internet Res.* 30, 789–809. doi: 10.1108/INTR-05-2019-0208
- Akbulut, Y., Dönmez, O., and Dursun, Ö. Ö. (2017). Cyberloafing and social desirability bias among students and employees. *Comput. Hum. Behav.* 72, 87–95. doi: 10.1016/j.chb.2017.02.043
- Akgün, F. (2020). Investigation of high school students' cyberloafing behaviors in classes. *Educ. Sci.* 45, 79–108. doi: 10.15390/EB.2019.8419
- Amanvermez, Y., Zhao, R., Cuijpers, P., de Wit, L. M., Ebert, D. D., Kessler, R. C., et al. (2022). Effects of self-guided stress management interventions in college students: a systematic review and meta-analysis. *Internet Interv.* 28:100503. doi: 10.1016/j.invent.2022.100503
- Ang, W. H. D., Lau, S. T., Cheng, L. J., Chew, H. S. J., Tan, J. H., Shorey, S., et al. (2022). Effectiveness of resilience interventions for higher education students: a meta-analysis and metaregression. *J. Educ. Psychol.* 114, 1670–1694. doi: 10.1037/edu0000719
- Askew, K. L., Ilie, A., Bauer, J. A., Simonet, D. V., Buckner, J. E., and Robertson, T. A. (2019). Disentangling how coworkers and supervisors influence employee cyberloafing: what normative information are employees attending to? *J. Leadersh. Org. Stud.* 26, 526–544. doi: 10.1177/1548051818813091
- Batabyal, S. K., and Bhal, K. T. (2020). Traditional cyberloafing, mobile cyberloafing and personal mobile-internet loafing in business organizations: exploring cognitive ethical logics. *J. Inf. Commun. Ethics Soc.* 18, 631–647. doi: 10.1108/JICES-07-2019-0081
- Baturay, M. H., and Toker, S. (2015). An investigation of the impact of demographics on cyberloafing from an educational setting. *Comput. Hum. Behav.* 50, 358–366. doi: 10.1016/j.chb.2015.03.081
- Berweger, S., Krattenmacher, S., Salzmann, P., and Schönenberger, S. (2013). *LiSA. Lernende im Spannungsfeld von Ausbildungserwartungen, Ausbildungsrealität und erfolgreicher Erstausbildung*. St. Gallen: Pädagogische Hochschule.
- Billett, S. (2014). The standing of vocational education: sources of its societal esteem and implications for its enactment. *J. Vocat. Educ. Train.* 66, 1–21. doi: 10.1080/13636820.2013.867525
- Böhn, S., and Deutscher, V. (2022). Dropout from initial vocational training – a meta-synthesis of reasons from the apprentice's point of view. *Educ. Res. Rev.* 35:100414. doi: 10.1016/j.edurev.2021.100414
- Bouw, E., Zitter, I., and De Bruijn, E. (2021). Multilevel design considerations for vocational curricula at the boundary of school and work. *J. Curric. Stud.* 53, 765–783. doi: 10.1080/00220272.2021.1899290
- Bürgi, R., and Gonon, P. (2021). Varieties within a collective skill formation system: how VET governance in Switzerland is shaped by associations. *Int. J. Res. Vocat. Educ. Train.* 8, 46–64. doi: 10.13152/IJRJET.8.1.3
- Cavanaugh, M. A., Boswell, W. R., Roehling, M. V., and Boudreau, J. W. (2000). An empirical examination of self-reported work stress among US managers. *J. Appl. Psychol.* 85, 65–74. doi: 10.1037//0021-9010.85.1.65
- Cutuli, D. (2014). Cognitive reappraisal and expressive suppression strategies role in the emotion regulation: an overview on their modulatory effects and neural correlates. *Front. Syst. Neurosci.* 8, 105–106. doi: 10.3389/fnsys.2014.00175
- Ersoy, H., Duman, E., and Öncü, S. (2016). Motivation and success with augmented reality: an experimental study. *J. Instruct. Technol. Teacher Educ.* 5, 39–44.
- Farivar, F., and Richardson, J. (2018). Cross-domain online social networking and job/life satisfaction. *Academy of Management Global Proceedings*, 160. Available at: <https://journals.aom.org/doi/abs/10.5465/amgbproc.surrey.2018.0160.abs>
- Ferm, L. (2021). Vocational students' ways of handling the academic/vocational divide. *Int. J. Res. Vocat. Educ. Train.* 8, 1–20. doi: 10.13152/IJRJET.8.1.1
- Flanigan, A. E., and Babchuk, W. A. (2022). Digital distraction in the classroom: exploring instructor perceptions and reactions. *Teach. High. Educ.* 27, 352–370. doi: 10.1080/13562517.2020.1724937
- Garnefski, N., Kraaij, V., and Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personal. Individ. Differ.* 30, 1311–1327. doi: 10.1016/S0191-8869(00)00113-6
- Garnefski, N., Kraaij, V., and Spinhoven, P. (2002). *Manual for the use of the cognitive emotion regulation questionnaire*. Leiderdorp: DATEC.
- Gaskin, J. (2012). Group Comparison, Stats Tools Package. Available at: <http://statwiki.kolobkreaties.com>
- Gijón, P. (2020). *Aprendizaje y educación de adultos: evolución, tendencias y organización* Técnica Avicam.
- Gökçeşlan, Ş., Mumcu, F. K., Haşlamam, T., and Çevik, Y. D. (2016). Modeling smartphone addiction: the role of smartphone usage, self-regulation, general self-efficacy and cyberloafing in university students. *Comput. Hum. Behav.* 63, 639–649. doi: 10.1016/j.chb.2016.05.091
- Gökçeşlan, Ş., Uluyol, Ç., and Şahin, S. (2018). Smartphone addiction, cyberloafing, stress and social support among university students: a path analysis. *Child Youth Serv. Rev.* 91, 47–54. doi: 10.1016/j.childyouth.2018.05.036
- Hayes, A. F., Preacher, K. J., and Myers, T. A. (2011). "Mediation and the estimation of indirect effects in political communication research" in *The sourcebook for political communication research: Methods, measures, and analytical techniques*. eds. E. P. Bucy and R. L. Holbert (New York, NY: Routledge), 434–465.
- Hernández, C., Ferrer, C., Montagud, S., and Blanco, M. C. (2023). Estrés en tiempos de confinamiento: estrategias de afrontamiento y crecimiento postraumático en población universitaria. *Revista de Psicología y Educación* 18, 1–10. doi: 10.23923/rpye2023.01.229
- Hösl, S., Wade, L., and Von Wyl, A. (2018). "Stress und soziale Unterstützung im ersten Jahr einer Berufsausbildung" in *Jugendliche im Übergang zwischen Schule und Beruf*. eds. F. Sabatella and A. von Wyl (Berlin: Springer), 23–39.
- Johnson, C. C., Walton, J. B., Strickler, L., and Elliott, J. B. (2023). Online teaching in K-12 education in the United States: a systematic review. *Rev. Educ. Res.* 93, 353–411. doi: 10.3102/00346543221105550
- Junco, R., and Cotten, S. R. (2012). No a 4 U: the relationship between multitasking and academic performance. *Comput. Educ.* 59, 505–514. doi: 10.1016/j.compedu.2011.12.023
- Karaoglan Yilmaz, F. G., Avci, U., and Yilmaz, R. (2023). The role of loneliness and aggression on smartphone addiction among university students. *Curr. Psychol.* 42, 17909–17917. doi: 10.1007/s12144-022-03018-w
- Karatepe, O. M., Yavas, U., Babakus, E., and Deitz, G. D. (2018). The effects of organizational and personal resources on stress, engagement, and job outcomes. *Int. J. Hosp. Manag.* 74, 147–161. doi: 10.1016/j.ijhm.2018.04.005
- Kassam, K. S., Koslov, K., and Mendes, W. B. (2009). Decisions under distress. *Psychol. Sci.* 20, 1394–1399. doi: 10.1111/j.1467-9280.2009.02455.x
- Khansa, L., Kuem, J., Siponen, M., and Kim, S. S. (2017). To cyberloaf or not to cyberloaf: the impact of the announcement of formal organizational controls. *J. Manag. Inf. Syst.* 34, 141–176. doi: 10.1080/07421222.2017.1297173
- Kim, M., and Beehr, T. A. (2020). Thriving on demand: challenging work results in employee flourishing through appraisals and resources. *Int. J. Stress. Manag.* 27, 111–125. doi: 10.1037/str0000135
- Koay, K. Y., Soh, P. C. H., and Chew, K. W. (2017). Do employees' private demands lead to cyberloafing? The mediating role of job stress. *Management Research Review* 40, 1025–1038. doi: 10.1108/MRR-11-2016-0252
- Koay, K. Y., and Soh, P. C. H. (2018). Should cyberloafing be allowed in the workplace? *Hum. Resour. Manag. Int. Dig.* 26, 4–6. doi: 10.1108/HRMID-05-2018-0107
- Kötter, T., Wagner, J., Brüheim, L., and Voltmer, E. (2017). Perceived medical school stress of undergraduate medical students predicts academic performance: an observational study. *BMC Med. Educ.* 17, 1–6. doi: 10.1186/s12909-017-1091-0
- Kraaij, V., Garnefski, N., de Wilde, E. J., Dijkstra, A., Gebhardt, W., Maes, S., et al. (2003). Negative life events and depressive symptoms in late adolescence: bonding and cognitive coping as vulnerability factors? *J. Youth Adolesc.* 32, 185–193. doi: 10.1023/A:1022543419747
- Kunzelmann, A., and Rigotti, T. (2021). Challenge demands and resilience: strain and learning as different pathways. *Zeitschrift für Arbeits- und Organisationspsychologie* 66, 59–71. doi: 10.1026/0932-4089/a000363
- Kwon, M., Kim, D.-J., Cho, H., and Yang, S. (2013). The smartphone addiction scale: development and validation of a short version for adolescents. *PLoS One* 8:83558. doi: 10.1371/journal.pone.0083558
- Lau, P., Kwong, T., Chong, K., and Wong, E. (2013). Developing students' teamwork skills in a cooperative learning project. *Int. J. Lesson Learn. Stud.* 3, 80–99. doi: 10.1108/IJLLS-03-2013-0018
- LePine, J. A., Podsakoff, N. P., and LePine, M. A. (2005). A meta-analytic test of the challenge stressor–hindrance stressor framework: an explanation for inconsistent

- relationships among stressors and performance. *Acad. Manag. J.* 48, 764–775. doi: 10.5465/AMJ.2005.18803921
- Leu, S. (2014). Stress und soziale Unterstützung bei Jugendlichen im ersten Jahr einer Berufsausbildung (Trabajo Fin de Grado, Zurich University of Applied Sciences). Repository Institutional ZHAW. *Appl. Psychol.* doi: 10.21256/zhaw-770
- Ley Orgánica. (2022). de 31 de marzo, de ordenación e integración de la Formación Profesional.
- Lim, V. K. G. (2002). The IT way of loafing on the job: Cyberloafing, neutralizing and organizational justice. *J. Organ. Behav.* 23, 675–694. doi: 10.1002/job.161
- Lim, V. K., and Teo, T. S. (2005). Prevalence, perceived seriousness, justification and regulation of cyberloafing in Singapore: An exploratory study. *Information & Management* 42, 1081–1093. doi: 10.1016/j.im.2004.12.002
- Lizarte, E. J. (2021). “Quality in higher education: drop-out profiles of students from the University of Granada” in *Educación y formar a la generación Z: experiencias derivadas de la investigación*. eds. A. J. Moreno, J. A. Marín, J. M. Trujillo and M. P. Cáceres (Barcelona: Octaedro), 364–372.
- Lizarte, E. J., and Gijón, J. (2022a). “El abandono en la formación profesional superior: Un reto para investigadores y actores políticos” in *Experiencias e investigaciones en contextos educativos*. eds. J. M. Trujillo, D. Capperucci, C. Rodríguez and N. Campos (Madrid: Dykinson), 364–372.
- Lizarte, E. J., and Gijón, J. (2022b). Prediction of early dropout in higher education using the SCPQ. *Cogent Psychol.* 9:2123588. doi: 10.1080/23311908.2022.2123588
- Lowe-Calverley, E., and Grieve, R. (2017). Web of deceit: relationships between the dark triad, perceived ability to deceive and cyberloafing. *Cyberpsychology* 11:5. doi: 10.5817/CP2017-2-5
- Mazzola, J. J., and Disselhorst, R. (2019). Should we be “challenging” employees?: a critical review and meta-analysis of the challenge-hindrance model of stress. *J. Organ. Behav.* 40, 949–961. doi: 10.1002/job.2412
- Metin, I., and Demirtepe, D. (2021). Cyberloafing behaviors among university students: their relationships with positive and negative affect. *Curr. Psychol.* 42, 11101–11114. doi: 10.1007/s12144-021-02374-3
- Ministerio de Educación y Formación Profesional (MEFP) (2021). *Informe 2021 sobre el estado del sistema educativo. Curso 2019–2020*. Madrid: Secretaría General Técnica.
- OECD. (2018). *Getting skills right: Spain, getting skills right*, OECD Publishing, Paris.
- Park, J. H., and Park, J. H. (2017). The relationships among interpersonal relationship anxiety, college adjustment, self-control, and smartphone addiction in nursing students. *Journal of the Korean Data and Information Science Society* 28, 185–194. doi: 10.7465/jkdi.2017.28.1.185
- Pearsall, M. J., Ellis, A. P. J., and Stein, J. H. (2009). Coping with challenge and hindrance stressors in teams: behavioral, cognitive, and affective outcomes. *Organ. Behav. Hum. Decis. Process.* 109, 18–28. doi: 10.1016/j.obhdp.2009.02.002
- Podsakoff, N. P., LePine, J. A., and LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: a meta analysis. *J. Appl. Psychol.* 92, 438–454. doi: 10.1037/00219010.92.2.438
- Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/brm.40.3.879
- Pusceddu, T. (2016). Stresserleben und Stressbewältigung von kaufmännischen Lernenden (Trabajo Fin de Grado, Fachhochschule Nordwestschweiz). Repository Institutional FHNW. Available at: <https://doi.org/10.26041/fhnw-811>
- Reizer, A., Galperin, B. L., Chavan, M., Behl, A., and Pereira, V. (2022). Examining the relationship between fear of COVID-19, intolerance for uncertainty, and cyberloafing: a mediational model. *J. Bus. Res.* 145, 660–670. doi: 10.1016/j.jbusres.2022.03.037
- Ren, F., and Zhang, J. (2015). Job stressors, organizational innovation climate, and employees’ innovative behavior. *Creat. Res. J.* 27, 16–23. doi: 10.1080/10400419.2015.992659
- Saritepeci, M. (2020). Predictors of cyberloafing among high school students: unauthorized access to school network, metacognitive awareness and smartphone addiction. *Educ. Inf. Technol.* 25, 2201–2219. doi: 10.1007/s10639-019-10042-0
- Sawhney, G., and Michel, J. S. (2022). Challenge and hindrance stressors and work outcomes: the moderating role of day-level affect. *J. Bus. Psychol.* 37, 389–405. doi: 10.1007/s10869-021-09752-5
- Schaap, H., Baartman, L. K. J., and De Bruijn, E. (2012). Students’ learning processes during school-based learning and workplace learning in vocational education: a review. *Vocat. Learn.* 5, 99–117. doi: 10.1007/s12186-011-9069-2
- Schulten, D., and Wussler, K. (2013). Auszubildende im Stress? Was Unternehmen tun können. *Wirtschaftspsychologie aktuell* 2, 40–44.
- Steinmann, R. M. (2005). “Psychische Gesundheit – Stress” in *Wissenschaftliche Grundlagen für eine nationale Strategie zur Stressprävention und Förderung psychischer Gesundheit in der Schweiz* (Bern: Gesundheitsförderung Schweiz)
- Tandon, A., Kaur, P., Ruparel, N., Islam, J. U., and Dhir, A. (2021). Cyberloafing and cyberslacking in the workplace: systematic literature review of past achievements and future promises. *Internet Res.* 32, 55–89. doi: 10.1108/INTR-06-2020-0332
- Toepper, M., Zlatkin-Troitschanskaia, O., and Kühling-Thees, C. (2022). Literature review of international empirical research on transfer of vocational education and training. *Int. J. Train. Dev.* 26, 686–708. doi: 10.1111/ijtd.12276
- Umer, M., Ul, I., Murtaza, G., and Jaffery, H. (2022). Challenge–hindrance stressors, helping behavior and job performance: double-edged sword of religiousness. *J. Bus. Ethics* 179, 1–13. doi: 10.1007/s10551-022-05129-7
- Virgós, M., Burguera, J., and Pérez, M. H. (2022). La Formación Profesional Dual en la empresa desde la perspectiva de sus protagonistas. *Revista Complutense de Educación* 33, 27–39. doi: 10.5209/rced.70992
- Vorontsova, O., Ghisletta, P., Ababkov, V., Bondolfi, G., and Barisnikov, K. (2022). Short mindfulness-based intervention for psychological and academic outcomes among university students. *Anxiety Stress Coping* 35, 141–157. doi: 10.1080/10615806.2021.1931143
- Wagner, D. T., Barnes, C. M., Lim, V. K., and Ferris, D. L. (2012). Lost sleep and cyberloafing: evidence from the laboratory and a daylight saving time quasi-experiment. *J. Appl. Psychol.* 97, 1068–1076. doi: 10.1037/a0027557
- Wallace, J. C., Edwards, B. D., Arnold, T., Frazier, M. L., and Finch, D. M. (2009). Work stressors, role-based performance, and the moderating influence of organizational support. *J. Appl. Psychol.* 94, 254–262. doi: 10.1037/a0013090
- Wu, J. Y., and Xie, C. (2018). Using time pressure and note-taking to prevent digital distraction behavior and enhance online search performance: Perspectives from the load theory of attention and cognitive control. *Computers in Human Behavior* 88, 244–254. doi: 10.1016/j.chb.2018.07.008
- Yildiz, H., and Saritepeci, M. (2019). Occupational burnout and cyberloafing among teachers: analysis of personality traits, individual and occupational status variables as predictors. *Soc. Sci. J.* 56, 69–87. doi: 10.1016/j.soscj.2018.10.011
- Zhou, B., Li, Y., Hai, M., Wang, W., and Niu, B. (2021). Challenge-hindrance stressors and cyberloafing: a perspective of resource conservation versus resource acquisition. *Curr. Psychol.* 42, 1172–1181. doi: 10.1007/s12144-021-01505-0



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Ann Dowker,
University of Oxford, United Kingdom
Alper Çiltaş,
Ataturk University, Türkiye

*CORRESPONDENCE

Nurullah Eryilmaz
✉ ne331@bath.ac.uk

RECEIVED 06 February 2024

ACCEPTED 13 May 2024

PUBLISHED 13 June 2024

CITATION

Kandemir MA, Eryilmaz N and Shelley M (2024)
A study of factors and their influence on
mathematics-related beliefs and performance
in Turkey. *Front. Educ.* 9:1376521.
doi: 10.3389/feduc.2024.1376521

COPYRIGHT

© 2024 Kandemir, Eryilmaz and Shelley. 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 study of factors and their influence on mathematics-related beliefs and performance in Turkey

Mehmet Ali Kandemir¹, Nurullah Eryilmaz^{2*} and Mack Shelley³

¹Faculty of Education, Balıkesir University, Balıkesir, Türkiye, ²Department of Education, University of Bath, Bath, United Kingdom, ³Department of Political Science, Iowa State University, Ames, IA, United States

The present study endeavors to uncover the intricate nexus among pertinent variables, including self-efficacy, perceived usefulness, parental education levels, and gender, and their relationship with mathematics performance. To achieve this objective, several hypotheses were tested, and a path model was estimated. The study cohort comprised 117 8th-grade students (64 females, 53 males) enrolled in a secondary school in Turkey. Results, in general, supported research hypotheses and the path model. Males' mathematics self-efficacy was found to be higher than females. No significant gender difference was found in the mathematics performance score in the initial analysis. Nevertheless, gender made a significant contribution to the prediction of mathematics performance beyond the effects of other variables in the hierarchical regression analysis. Consistent with theoretical predictions, females' mathematics self-efficacy expectations were unrealistically low compared to males. However, the effect of self-efficacy on performance was found to be invariant across gender. Perceived usefulness was found to be moderately correlated with performance but made little contribution to the prediction of performance above and beyond the effects of other variables.

KEYWORDS

gender, mathematics performance, mathematics-related beliefs, secondary school students, Turkey

1 Introduction

Understanding the factors that influence students' mathematics performance is of paramount importance in educational research (Pajares and Kranzler, 1995; Ursini and Sanchez, 2008; Guven and Cabakcor, 2013). Mathematics is not only a fundamental subject but also provides a basis for developing problem-solving skills, logical reasoning, and mental discipline (Park et al., 2021). Moreover, mathematical knowledge plays a crucial role in understanding the contents of other school subjects, making it an essential component of overall curriculum (Cheema and Kitsantas, 2014; Uysal, 2015). By delving into the interplay of various factors such as gender, self-efficacy, perceived usefulness, and parental education, this study aims to unravel the complex relationships that underpin students' mathematics performance. The findings of this research can have far-reaching implications for educational policies and practices in Turkey, ultimately contributing to the enhancement of mathematics education and students' academic achievement.

1.1 Self-efficacy

Self-efficacy, defined as “beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (Bandura, 1997, p. 3), serves as a robust predictor of mathematics performance (Chen, 2003; Pietsch et al., 2003; Stevens et al., 2004, 2007; Ferla et al., 2009; Fast et al., 2010). Pajares and Miller (1994) found that mathematics self-efficacy surpassed math self-concept, perceived usefulness of mathematics, prior experience with mathematics, and gender in predicting problem-solving ability. In a study involving 416 high school students, Pietsch et al. (2003) highlighted a stronger correlation between mathematics self-efficacy and performance compared to mathematics self-concept. Wang et al. (2012) emphasized the centrality of students’ self-concept of ability in mathematics-to-mathematics achievement among 8th graders in various countries. This trend was also observed in Hong Kong and Singapore (Chen, 2014), where mathematics achievement correlated significantly with contextual factors. In Shanghai, a top-performing region in mathematics according to the Program for International Student Assessment, Yao et al. (2018) found a moderate correlation between students’ sense of wellbeing and mathematics performance.

Several path-analytic studies investigated the nuanced relationship between mathematics self-efficacy and performance alongside other variables. In one study, mathematics self-efficacy mediated between mathematics attitude and performance (Randhawa et al., 1993). Pajares and Kranzler (1995) identified strong direct effects of ability and self-efficacy on mathematics performance, with ability also exerting a significant direct effect on self-efficacy, which, in turn, mediated the indirect influence of ability on performance.

Pajares (1996) explored the predictive and mediational role of self-efficacy beliefs in mathematical problem-solving among middle school gifted students, highlighting self-efficacy’s unique contribution even after accounting for various factors. Pajares and Graham (1999) identified self-efficacy as the sole motivational variable predicting performance.

Furthermore, Schukajlow et al. (2018) study observed that prompting students to generate multiple mathematical solutions for real-world problems did not directly enhance self-efficacy. However, they noted an indirect positive impact on perceived competence and self-efficacy, particularly among students with initially low self-efficacy.

1.2 Perceived usefulness

Similar to self-efficacy, attitude, often referred to as “perceived usefulness,” exhibits a significant association with mathematics performance (Hackett and Betz, 1989; Pajares and Miller, 1994; Peters, 2013). However, its predictive efficacy and impact on mathematics performance appear to be somewhat overshadowed by self-efficacy. For instance, in a study by Hackett and Betz (1989), attitude demonstrated a moderate correlation with both self-efficacy ($r = 0.47$) and mathematics performance ($r = 0.40$). Crucially, in stepwise regression analysis, perceived usefulness failed to make a significant contribution to predicting mathematics

performance, while self-efficacy emerged as a significant predictor. Further corroborating this, a path analytic study by Pajares and Miller (1994) revealed that self-efficacy exerted a direct influence on mathematics performance, whereas perceived usefulness did not.

Hemmings et al. (2011) findings based on Australian secondary school students indicated that female students tended to harbor more positive attitudes toward mathematics. Notably, the primary predictors of scores on a nationally recognized mathematics examination were prior achievement in mathematics and attitudes toward mathematics. Even after controlling for these predictors, gender differences did not prove to be statistically significant.

1.3 Gender dynamics in mathematics: self-efficacy, attitude, and performance

The role of gender in the realm of mathematics encompasses intricate dynamics, as underscored by Gallagher and Kaufman (2005), who asserted that the relationship between gender and mathematics is more nuanced than conventional beliefs suggest. In the United States, female students tend to score lower than male counterparts on standardized tests (e.g., Willingham and Cole, 1997) but demonstrate equal or superior performance in mathematics classes (e.g., Bridgeman and Wendler, 1991; Kessel and Linn, 1996). However, this pattern does not universally apply, as studies in South Africa (Cherian and Siweya, 1996) and the Netherlands (Taal, 1994) reveal no consistent male advantage in mathematics performance or attitudes.

Numerous investigations have delved into gender differences in mathematics self-efficacy, with a recurring theme indicating that boys often exhibit higher self-efficacy than girls (Lent et al., 1993; Randhawa, 1994; Junge and Dretzke, 1995; Lussier, 1996; Gwilliam and Betz, 2001; Di Martino and Zan, 2011). Nonetheless, some studies report no significant disparity between boys and girls in mathematics self-efficacy (Fouad and Smith, 1996; Middleton and Midgley, 1997). Pajares (2005) challenged the notion that gender differences in mathematics self-efficacy favor girls, emphasizing the absence of such bias across various educational domains. Hackett and Betz’s (1981) hypothesis that girls tend to underestimate their mathematics self-efficacy relative to boys was not substantiated in a subsequent study (Hackett and Betz, 1989).

Previous research has indeed highlighted these conflicting trends. For instance, Hembree (1990) identified a stronger correlation between mathematics anxiety and mathematical performance in males compared to females. This finding suggests that anxiety might more significantly impede males’ mathematical performance, a notable contrast to other studies. Conversely, Devine et al. (2012) reported an opposite trend, where females exhibited a stronger relationship between mathematics anxiety and performance. Such discrepancies underscore the complex interplay between gender and attitudes toward mathematics.

Moreover, the findings by Miller and Bichsel (2004) contribute additional layers to this complexity. They discovered that mathematics anxiety correlates differently with various types of mathematical skills across genders: it is more related to basic mathematics scores in males, whereas in females, it is more pertinent to applied mathematics scores. This differentiation

suggests that the impact of mathematics anxiety may not only vary between genders but also across different mathematical domains.

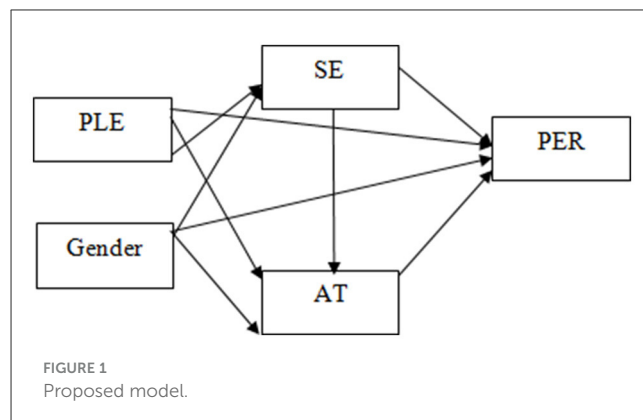
These studies collectively highlight that gender can influence the relationship between mathematics-related attitudes, such as anxiety, and performance in nuanced ways. Given these precedents, our current focus on self-efficacy must be contextualized within this broader framework. It remains essential to consider these varied influences as we further explore how self-efficacy specifically interacts with mathematical performance across genders. Understanding these dynamics can help tailor educational strategies that address the specific needs and challenges of both male and female students in mathematics education.

Attitudinal variations also emerge, with Fennema and Sherman (1977) finding that boys generally hold more positive attitudes toward mathematics. Beal (1999) observed that while girls tend to excel in receiving high grades in mathematics classes, boys may outperform them in formal testing situations. Halpern et al. (2005) proposed a psychobiological model, challenging nature vs. nurture dichotomies and emphasizing the interplay between biological and psychosocial factors. An alternative model by Byrnes (2005) posited that optimal academic achievement is facilitated when students encounter learning opportunities aligned with their motivation and abilities. Australian results (Forgasz and Hill, 2013) from the Program for International Student Assessment indicated a widening gender gap in mathematics achievement, favoring males, students from higher socioeconomic backgrounds, and those attending metropolitan schools, particularly in more challenging mathematical content.

In exploring the impact of task context, Zohar and Gershikov (2008) discerned gender-specific responses across three contexts of mathematical tasks (stereotypically boys', girls', and gender-neutral). Girls' performance was notably affected by task context, with boys demonstrating higher scores in stereotypically boys' contexts, while a significant interaction between age and gender suggested age-dependent effects on girls in stereotypically girls' contexts. In neutral contexts, gender differences were not discernible. These findings emphasize the nuanced interplay between gender, task context, and mathematical performance.

1.4 Parental and cultural influences on mathematics achievement

In addition to gender dynamics, self-efficacy, and attitudes, parental involvement and education levels emerge as crucial factors intertwined with mathematics performance. Illustratively, a study conducted in Slovenia by Levpušček et al. (2013) revealed the positive and direct impact of student intelligence, self-rated openness, and parental education on mathematics performance. Similarly, research in Turkey by Gun and Erdem (2014) identified the influence of fathers' academic backgrounds, alongside attitudes toward the course and teachers, on students' mathematics performance. Ismail and Awang (2008) study in Malaysia underscored the significance of various factors, including gender, language spoken at home, expected educational level, family background, and home educational resources, in shaping eighth-grade students' mathematics achievement.



Moreover, McConney and Perry (2010) found that student socioeconomic status, emanating from parental and family factors, accounted for a substantial portion of explained variance in students' mathematics literacy performance in Australia. Ing (2014), utilizing latent growth curve analysis, explored the relationship between perceived early parental support, mathematics and science achievement trajectories, and persistence in STEM careers by gender. The results indicated a positive link between mathematics and science trajectories and STEM career persistence for both genders yet perceived early parental support specifically enhanced mathematics achievement for males.

Investigating cross-national cultural differences, Hu et al. (2018) revealed that national culture played a pivotal role, accounting for a significant portion of variations in mathematics performance across 51 countries. Particularly, the cultural trait of "long-term orientation" demonstrated a robust positive association with mathematics achievement. These findings suggest a substantial role for parents in shaping students' mathematics achievement, given that cultural norms are often established and reinforced within the family structure.

A study conducted in India by Soni and Kumari (2017) employed path analysis to elucidate that parental anxiety and attitude toward mathematics serve as precursors to their children's mathematics anxiety, attitude, and ultimately, achievement. This comprehensive exploration underscores the intricate interplay of various factors, extending beyond individual attributes and encompassing the pivotal role of parental involvement and cultural context in shaping mathematics performance.

1.5 Mathematics performance model

These studies collectively underscore the unmistakable significance of self-efficacy, attitude, gender, and parental education in shaping the landscape of mathematics performance. Notably, the intricacies surrounding gender roles in mathematics performance and mathematics self-efficacy add a layer of complexity to our understanding. The present research aims to elucidate the nuanced relationship between gender and other mathematics-related constructs. In pursuit of this goal, we constructed and tested the model depicted in Figure 1, drawing insights from the literature to inform its foundation.

The model postulates a multifaceted and intricate relationship among key factors influencing mathematics performance (PER). Attitude (AT) is conceptualized as a dynamic outcome influenced by self-efficacy (SE), parents' level of education (PLE), and gender. Drawing from the literature, we hypothesize that self-efficacy exerts both direct and indirect influences on performance—its impact on attitude serving as a crucial mediating factor. This research framework is constructed upon a synthesis of findings from the provided literature, aiming to unravel the interplay of these factors and their collective influence on mathematics performance.

1.6 Purpose and research questions

The main thrust of the current study lies in evaluating the effectiveness of the proposed model and exploring potential gender-based variations in the impact of self-efficacy on performance. To achieve this overarching goal, the study systematically investigates nine key aspects. Firstly, it conducts a detailed exploration of the student profile concerning math-related constructs, encompassing factors such as self-efficacy, attitude, gender, and parents' level of education. Subsequently, the study delves into potential gender differences within these math-related constructs. The third aspect focuses on unraveling the intricate relationships among the factors influencing mathematics performance.

Moving forward, the study probes the specific relationship between self-efficacy and attitude, seeking to elucidate the dynamics between these pivotal constructs. The correspondence between self-efficacy and performance is then examined for potential gender-based variations. Following this, the study investigates which variable emerges as the most robust predictor of mathematics performance, contributing to a deeper understanding of the determinants of academic success in this context.

The role of attitude is explored by examining whether it accounts for a statistically significant amount of variance beyond the effects of self-efficacy. Additionally, the study assesses the predictive utility of parents' level of education on mathematics performance. Lastly, the study evaluates the overall utility of the proposed model, aiming to determine whether it provides a meaningful and insightful framework for comprehending the complexities of mathematics performance. Through the systematic investigation of these aspects, the study aims to contribute valuable insights to the existing body of literature in the field.

2 Method

2.1 Participants

The participants consisted of 117 8th graders (64 females, 53 males) in a secondary school in the city of Balıkesir, Turkey. The frequency and percentage of the participants' parents' level of education are presented in [Table 1](#). Mean scores were 3.37 ($SD = 0.86$) for the father's level of education and 2.96 ($SD = 1.01$) for the mother's level of education. The parents' combined level of education scores was calculated by averaging these two values, resulting in a score of 3.16.

2.2 Research instruments

The first part of the survey included demographic information of the participants, including gender and the highest education that their mother and father earned. Corresponding to the Turkish educational system, they indicated their response on a 5-point scale: (1) primary school (5 years), (2) secondary school (8 years), (3) high school (12 years), (4) university, and (5) graduate degree. Their responses to these two items were averaged to find parents' level of education score (PLE). Thus PLE scores ranged from 1 to 5, with higher scores indicating a higher level of education attained.

The second part of the survey consisted of two scales: Mathematics Self-efficacy and Attitude toward Mathematics. The participants indicated their response on a 5-point range from 1 (Strongly Disagree) to 5 (Strongly Agree). Negatively-worded items were reverse-scored before calculating the total score. The participants' scores were calculated by averaging all of their responses to the self-efficacy and attitude scale items separately. Thus, their overall self-efficacy and attitude scores ranged from 1 to 5, with higher scores indicating stronger self-efficacy and more positive attitude toward mathematics.

The 14-item Self-efficacy scale employed in the current study was developed by [Umay \(2001\)](#). Some of the items included: "I feel confident that I can solve mathematical problems," "If I spend enough time and put necessary efforts, I can solve different types of mathematical problems," and "I can help my friends if they have difficulty solving problems."

Attitude scale items used in the current study were adapted from [Fennema and Sherman \(1976\)](#) mathematics attitude scale's 10-item perceived usefulness sub-scale. These 10 items were translated into Turkish with the researcher and were shown to three experts in mathematics education to establish face validity. All of the experts agreed that the scale measures attitude toward mathematics. The items included "Math will not be important to me in my life's work" and "I will use mathematics in many ways in the future." Like the self-efficacy scale, the participants indicated their response on a 5-point range from 1 (Strongly Disagree) to 5 (Strongly Agree). Negatively-worded items were reverse-scored before calculating the total score by averaging all of their responses to the scale items. Thus, their overall attitude scores ranged from 1 to 5, with higher scores indicating more positive attitudes.

To establish the validity of the attitude and self-efficacy scales, two-factor confirmatory factor analysis was conducted with another group of 300 students. All of the index values were met or exceed established threshold values: RMSEA (0.045) was <0.05 , and CFI (0.92) and NFI (0.93) were higher than the usual 0.90 cutoff value ([Cheung and Rensvold, 2002](#)). Item loadings ranged from 0.50 to 0.75 in the attitude scale and 0.45 to 0.80 for the self-efficacy scale, which verified the construct validity of these components of the model. Average Variance Extracted (AVE) values were 0.51 for the attitude scale and 0.54 for the self-efficacy scale. Since these values exceeded the threshold value of 0.50, convergent validity was established. To assess discriminant validity, the square root of AVE was calculated and compared with the correlation between these two constructs. The square root of AVE was 0.71 for attitude and 0.73 for self-efficacy. Discriminant validity was established since both values were greater than the correlation between these two constructs ($r = 0.68$).

TABLE 1 Parents' level of education.

Highest degree	Father level of education		Mother level of education	
	Frequency	Percentage	Frequency	Percentage
Primary school	7	6	14	12
Secondary school	7	6	18	15.4
High school	40	34.2	44	37.6
University	63	53.8	40	35.1

The participants' mathematics performance scores were assessed using the TEOG nationwide standardized test selection exam for entering high school, which had been used since 2013 in Turkey. It consists of 100 multiple-choice questions, 20 of which were related to mathematics. Students receive this exam twice a year. The mathematics section of the exam measures the basic skills for success in high school. Each question is worth 5 points. Thus, students' scores range from 0 to 100. Their scores on two exams were averaged to calculate students' mean mathematics performance.

2.3 Data analysis

Data analysis started with examining Cronbach's alpha values to establish the reliability of the scales. Then, descriptive statistics were used to explore the participants' scores on math-related constructs. *T*-tests were used to examine whether there is a difference between male and female students in self-efficacy, attitude, and performance mean scores. Cohen's *d* was employed to measure effect size. According to Cohen (1988), effect sizes <0.2 are considered small, values <0.3 are medium, and values 0.5 or higher are considered large. To explore the correspondence between self-efficacy and performance by gender at the individual level, *D* scores proposed by Dowling (1978) were calculated for each participant.

Pearson correlations were estimated to explore interrelationships among variables of interest. Then, the same analysis was done for males and females separately. Fisher's *z*-test was conducted to determine if correlations between self-efficacy and performance differ significantly by gender. Hierarchical regression analysis was employed to explore relative contributions of variables in predicting performance. Based on previous research studies and theoretical framework, variables were entered into the prediction equation in the following order: gender, PLE, SE, and AT. To explore the unique contributions of gender and PLE to performance, another hierarchical regression analysis was conducted, with PLE entered into the equation in the third step and gender entered in the last step.

Finally, a path analysis model was estimated to test the effects of gender, PLE, SE, and AT on mathematics performance. Multi-group invariance analysis was conducted to examine the role of self-efficacy and other variables on performance by gender. In this analysis, the constrained model was compared with the unconstrained model utilizing a χ^2 difference test.

TABLE 2 Descriptive statistics for math-related constructs.

Construct	Mean	Standard deviation	Skewness	Kurtosis
Self-efficacy	3.51	1.04	−0.40	−0.47
Attitude	2.71	1.10	0.08	−0.41
Performance	72.50	20.63	−0.45	−0.92

3 Results

Cronbach's alpha values were 0.85 for attitude and 0.88 for self-efficacy scales. These results indicate high reliability. Table 2 presents descriptive statistics for each variable. The self-efficacy scores ranged from 1 to 5 and the mean score was found to be more than the mid-point of 3, which indicated overall positive responses to the items by the participants. On the other hand, the attitude scores ranged from 1 to 5 with mean of 2.71, which suggested that the participants, in general, did not possess positive attitudes toward mathematics. The standard deviations of self-efficacy and attitude scores were relatively low, meaning that data are clustered closely around the mean in these two variables. The participants' performance scores ranged from 26.5 to 100 with mean of 72.50. This indicated that the participants' performance was high; however, the standard deviation of performance was relatively high, indicating that the data are widely spread. All of the skewness and kurtosis values fell in the desirable range of between −2 and +2 recommended by George and Mallery (2010), which indicated multivariate normality.

Table 3 shows the difference between male and female students in their self-efficacy, attitude, and performance scores. Male students' mean self-efficacy score was statistically higher than the mean for female students ($t = -2.33$, $p < 0.01$). Cohen's *d* was 0.44, which indicated a moderate effect size. No significant difference between male and female students existed in mean performance scores. In addition, no significant difference between male and female students was found in the attitude scores.

To explore self-efficacy/performance correspondence by gender, self-efficacy and performance scores were transformed into *z* scores separately and the standardized performance scores were subtracted from the standardized self-efficacy to calculate *D* scores, which measured congruence between self-efficacy and performance scores. *D* values of 0 indicate perfect congruence. Dowling (1978) noted that *D* scores fell into five categories: overconfident ($D > 0.8$), somewhat overconfident ($0.4 < D < 0.8$),

TABLE 3 Gender differences in math-related constructs.

Construct	Mean (standard deviation)		<i>t</i>	Effect size
	Female	Male		
Self-efficacy	3.32 (1.14)	3.76 (0.86)	−2.33	0.44
Attitude	2.67 (1.04)	2.77 (1.05)	−0.55	0.09
Performance	72.49 (20.87)	72.51 (20.53)	0.01	0.00

TABLE 4 Percentage of overconfidence and under confidence by gender.

D score category	Female		Male	
	<i>N</i>	%	<i>N</i>	%
Overconfident	7	10.9	11	20.8
Somewhat confident	7	10.9	10	18.9
Congruent	27	42.2	18	3
Somewhat under confident	10	15.6	9	17
Under confident	13	20.3	5	9.4

TABLE 5 Correlations among the variables.

	1	2	3	4
1. Parents education	–			
2. Self-efficacy	0.33	–		
3. Attitude	0.28	0.68	–	
4. Performance	0.37	0.67	0.60	–

All of the correlations were significant ($p < 0.01$).

congruent ($0.42 D > -0.4$), somewhat under confident ($-0.4 > D > -0.8$), and under confident ($D < -0.8$).

The percentages and number of male and female students in each category are shown in Table 4, with 10.9% of female and 20.8% of male students falling into the overconfident category. Similarly, 20.3% of female students and 9.4% of male students fell into the under confident category. These results suggest that while female students tend to underestimate their mathematics skills, male students tend to overestimate and exaggerate their skills. A chi-square test was conducted to examine if there is a statistically significant difference between male and female students in D scores. Results revealed a significant difference between the two groups [$\chi^2(4, N = 117) = 5.84, p < 0.05$].

Table 5 shows the results of Pearson correlation analysis. All of the variables were found to be significantly correlated with each other. Self-efficacy yielded the highest correlation with performance ($r = 0.67, p < 0.01$), followed by attitude ($r = 0.60, p < 0.01$). The correlation between self-efficacy and attitude ($r = 0.68, p < 0.01$) was higher than expected.

Correlations among key variables are presented by gender (females above the diagonal, males below) in Table 6. Self-efficacy seemed to be more strongly associated with performance in females (0.73) than males (0.52). Nevertheless, Fisher's *z*-test revealed that none of the differences between males and females was significant,

TABLE 6 Correlations among variables by gender.

	1	2	3	4
1. Parents education	–	0.33*	0.27*	0.37**
2. Self-efficacy	0.33*	–	0.66**	0.73**
3. Attitude	0.30	0.67**	–	0.62**
4. Performance	0.35**	0.62**	0.58**	–

* $p < 0.05$, ** $p < 0.01$.

TABLE 7 Results of hierarchical regression analysis.

Predictor	Beta	<i>F</i>	<i>R</i>	<i>R</i> ² change
Regression 1				
Gender	0.00	0	0.00	0.00
Regression 2				
		8.68	0.37	0.13**
Gender	−0.01	0		
Parents education	0.36	17.37		
Regression 3				
		35.78	0.69	0.35**
Gender	−0.14	0		
Parents education	0.16	17.37		
Self-efficacy	0.64	78.22		
Regression 4				
		28.72	0.71	0.02**
Gender	−0.12	0		
Parents education	0.14	17.37		
Self-efficacy	0.49	78.22		
Attitude	0.20	4.35		

** $p < 0.01$.

which means that the intercorrelations among key variables were invariant across gender in the current study.

Table 7 shows the relative contribution of predictor variables on performance. The regression results reveal that attitude accounted for 2% of variation in performance above and beyond the effects of other variables. The effect sizes for attitude and parents' level of education were 0.02 and 0.17, respectively. While attitude had a very small effect size, parents' level of education had a more substantial effect size on performance. Even after gender was entered into the equation predicting performance in the last step after the effects of self-efficacy, attitude, and parents' education were partial led out, attitude still accounted for a significant amount of variance in performance.

Figure 2 shows the resulting path coefficients of the proposed research model. The results showed that self-efficacy had the strongest effect on performance ($\beta = 0.54, p < 0.01$).

Parents' education influenced performance directly ($\beta = 0.17, p < 0.01$) and indirectly through its influence on self-efficacy ($\beta = 0.34, p < 0.01$). Attitude had a significant direct influence on performance ($\beta = 0.21, p < 0.01$). While gender had a direct significant influence on performance ($\beta = 0.24, p < 0.01$) and attitude ($\beta = 0.22, p < 0.01$), in both instances favoring females, it had a direct significant influence on self-efficacy ($\beta = -0.36, p <$

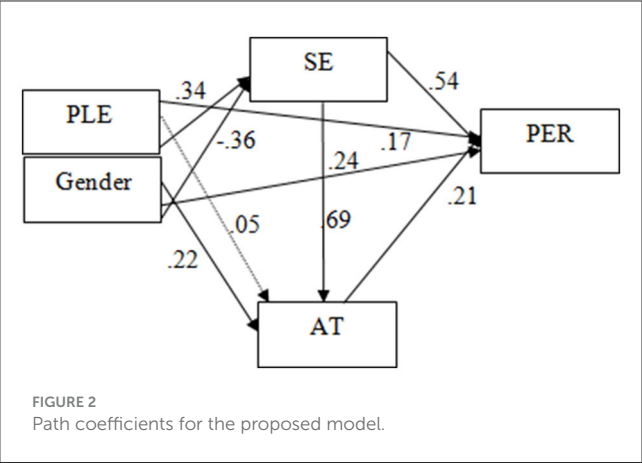


TABLE 8 Direct, indirect, and total effects of the proposed model.

Outcome		Determinant	Standardized estimates	
		Direct	Indirect	Total
Performance $R^2 = 0.53$	Gender	0.24	−0.21	0.03
	Self-efficacy	0.54	0.17	0.71
	Attitude	0.21	–	0.21
	Parents education	0.17	0.25	0.42
Attitude $R^2 = 0.57$	Gender	0.22	−0.29	−0.07
	Self-efficacy	0.69	–	0.80
	Parents education	0.05	0.27	0.33
Self-efficacy $R^2 = 0.32$	Gender	−0.36	–	−0.36
	Parents education	0.34	–	0.34

0.01) in favor of males. The path coefficient from parents' level of education to attitude was not significant ($\beta = 0.05, p > 0.01$).

As Table 8 indicates, the most dominant determinant of performance was self-efficacy, with a total effect of 0.71, which is regarded as a relatively large effect. Like self-efficacy, parents' level of education had a medium total effect (0.42) on performance. Gender, parents' level of education, and self-efficacy accounted for 47% of variation in attitude. The predictor variables collectively accounted for 53% of variation in performance.

To explore whether the effects of self-efficacy, parents' education, and attitude on performance differ by gender, another model was estimated with gender removed first, and then path analysis was conducted for the pooled sample of all groups and for females and males separately. Table 9 shows the path coefficients for this unconstrained base model in which all path coefficients were freely estimated.

TABLE 9 Path coefficients across gender.

Path	Total sample	Female	Male
PLE→ SE	0.38	0.37	0.40
PLE→ AT	0.06	0.02	0.10
SE→ AT	0.72	0.76	0.67
PLE→ PER	0.18	0.17	0.18
SE→ PER	0.45	0.57	0.39
AT→ PER	0.23	0.14	0.25

TABLE 10 Multi-sample analysis of paths for males and females.

	χ^2	$\Delta\chi^2$ from unconstrained base model	P-value
Base model constrained path	7.07		
SE→ PER	7.15	0.08	0.77
PLE→ PER	7.08	0.01	0.98
AT→ PER	7.52	0.50	0.48

The path coefficient between self-efficacy and performance was stronger for females than for males (0.41 vs. 26). To examine if this difference is significant, multi-group invariance analysis was conducted with the path from self-efficacy to performance constrained to be equal across gender. The resulting model fit was then compared to the unconstrained model. A χ^2 difference test then was conducted to see if the difference is statistically significant. If no significant difference was detected between the two models, then they are considered invariant. The χ^2 difference value (χ^2 change = 1.88, $p > 0.05$) was not significant. In addition to this analysis, two more multi group analyses were conducted by constraining two paths separately. Results presented in Table 10 revealed that the effects of self-efficacy, attitude, and parents' education on performance are invariant across gender.

4 Discussion

The primary aim of this paper was to investigate the interrelationships among various variables influencing mathematics performance, specifically exploring whether the impact of self-efficacy on performance differs by gender. A comprehensive model was postulated, outlining a complex relationship involving parents' level of education, gender, self-efficacy, attitude, and mathematics performance, and its predictive utility was rigorously tested. Notably, the analysis employed national exam scores as an indicator of students' mathematics performance, distinguishing it from studies utilizing local assessment tools like classroom grades. The results robustly supported the utility of the proposed model, affirmatively addressing the research questions. Consequently, this study contributes to the literature on mathematics performance and its predictors.

Unsurprisingly, self-efficacy emerged as the variable with the highest correlation and the strongest direct effect on mathematics performance. In line with prior research, including [Lent et al. \(1993\)](#), [Randhawa et al. \(1993\)](#), [Junge and Dretzke \(1995\)](#), [Lussier \(1996\)](#), and [Gainor and Lent \(1998\)](#), males exhibited statistically higher self-efficacy than females. Additionally, females demonstrated unrealistically low mathematics self-efficacy expectations compared to males. However, contrary to [Hackett and Betz \(1989\)](#) and in contrast to their earlier results ([Hackett and Betz, 1981](#)), this study did not support the notion that the effect of self-efficacy on performance is stronger for males than for females.

Consistent with [Neale \(1969\)](#), [Ma \(1997\)](#), [Ma and Kishor \(1997\)](#), and [Choi and Chang \(2011\)](#), attitude was moderately correlated with mathematics performance but made only a minimal contribution to predicting it. Although path analysis revealed a direct and significant influence of attitude on performance, its effect appeared markedly lower than that of self-efficacy.

While our findings suggest a relationship between self-efficacy and mathematical performance, it is important to consider that this does not definitively establish the direction of causality. Self-efficacy might not only influence mathematical performance, but could also reflect an individual's actual performance in mathematics. Furthermore, the relationship could indeed be bidirectional, suggesting a complex interplay where performance influences self-efficacy just as self-efficacy impacts performance.

Parents' level of education emerged as a noteworthy variable, moderately related to performance, and exerting a direct and significant effect. One plausible explanation is that parents with higher educational attainment may be more actively involved in their children's learning, supporting their intellectual and academic growth, as suggested by [Levpušček et al. \(2013\)](#). In addition to parents' level of education, future researchers might consider measuring parents' involvement and utilizing it to predict mathematics performance.

Crucially, the proposed model accounted for 53% of the variation in performance, affirming its predictive utility and shedding light on the intricate relationships among gender, parents' level of education, self-efficacy, attitude, and mathematics performance. Beyond providing insights into performance disparities, the model offers a valuable perspective for understanding why some students excel in mathematics while others do not. As such, it can be instrumental for future researchers and mathematics teachers in explaining and predicting students' mathematics performance.

It is imperative to acknowledge the study's limitations, notably the collection of data from a single school with a relatively small sample size. Consequently, the generalizability of findings to other settings and student populations may be challenging. Future researchers are encouraged to test the proposed model in diverse settings, and if consistent with this study's findings, the model may gain broader applicability.

5 Implications

This article makes a distinctive contribution to the realm of mathematics education research by underscoring the paramount importance of self-efficacy and emphasizing that its effect remains

consistent across genders. A key implication of this pivotal finding for the enhancement of STEM expertise and achieving gender parity in students is that policymakers and education planners should prioritize the maintenance of high levels of self-efficacy among all students. This insight serves as a valuable guide for shaping teaching and learning strategies in mathematics education.

Moreover, our findings highlight the noteworthy conclusion that attitudes toward mathematics, while significant, pale in comparison to the predictive power of self-efficacy regarding student mathematics achievement. The considerable variability in attitudes, shaped by diverse experiences with instructors, curricular innovations, and parental background in mathematics, suggests that policymaking efforts and administrative focus should be judiciously directed toward addressing the more foundational factor of self-efficacy. This recognition helps to streamline resources and efforts, steering them away from potentially elusive attempts to address students' fluctuating attitudes toward mathematics.

Understanding the role of parental education attainment emerges as a crucial insight from our research. Given the active involvement of parents in homework assistance, school system interactions, and influencing curriculum sentiments, the consequential impact of parental education on students' mathematics performance is evident. While better-educated parents are more likely to actively engage in their children's education, fostering societal equity requires efforts to stimulate greater interest and involvement among less well-educated parents. This approach aims to increase the likelihood of mathematics success for students from less advantaged backgrounds, contributing to the diversification and expansion of the future STEM workforce with enhanced numeracy skills.

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 humans were approved by Balikesir University Ethic Committee. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. 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

NE: Supervision, Writing – review & editing. MK: Conceptualization, Data curation, Formal analysis, Funding

acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.
- Beal, C. R. (1999). Introduction: special issue on the math-fact retrieval hypothesis. *Contemp. Educ. Psychol.* 24, 171–180. doi: 10.1006/ceps.1999.1006
- Bridgeman, B., and Wendler, C. (1991). Gender differences in predictors of college mathematics performance and in college mathematics course grades. *J. Educ. Psychol.* 83, 275–284. doi: 10.1037/0022-0663.83.2.275
- Byrnes, J. P. (2005). “Gender differences in math cognitive processes in an expanded framework,” in *Gender Differences in Mathematics*, eds. A. M. Gallagher, and J. C. Kaufman (New York, NY: Cambridge University Press), 73–98. doi: 10.1017/CBO9780511614446.005
- Cheema, J. R., and Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: a look at gender and racial-ethnic differences. *Int. J. Sci. Math. Educ.* 12, 1261–1279. doi: 10.1007/s10763-013-9454-4
- Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learn. Individ. Differ.* 14, 79–92. doi: 10.1016/j.lindif.2003.08.003
- Chen, Q. (2014). Using TIMSS 2007 data to build mathematics achievement model of fourth graders in Hong Kong and Singapore. *Int. J. Sci. Math. Educ.* 12, 1519–1545. doi: 10.1007/s10763-013-9505-x
- Cherian, V. I., and Siweya, J. (1996). Gender and achievement in mathematics by indigenous African students majoring in mathematics. *Psychol. Rep.* 78, 27–34. doi: 10.2466/pr0.1996.78.1.27
- 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
- Choi, N., and Chang, M. (2011). Interplay among school climate, gender, attitude toward mathematics, and mathematics performance of middle school students. *Middle Grades Res. J.* 6:14.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Erlbaum.
- Devine, A., Fawcett, K., Szucs, D., and Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behav. Brain Funct.* 8, 1–9. doi: 10.1186/1744-9081-8-33
- Di Martino, P., and Zan, R. (2011). Attitude towards mathematics: a bridge between beliefs and emotions. *ZDM Math. Educ.* 43, 471–482. doi: 10.1007/s11858-011-0309-6
- Dowling, D. M. (1978). *The development of a mathematics confidence scale and its application in the study of confidence in women college students* [Doctoral dissertation]. The Ohio State University, Columbus, OH.
- Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., et al. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *J. Educ. Psychol.* 102, 729–740. doi: 10.1037/a0018863
- Fennema, E., and Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *Am. Educ. Res. J.* 14, 51–71. doi: 10.3102/00028312014001051
- Fennema, E., and Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales: instruments designed to measure attitudes toward the learning of mathematics by females and males. *J. Res. Math. Educ.* 7, 324–326. doi: 10.5951/jresmetheduc.7.5.0324
- Ferla, J., Valcke, M., and Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learn. Individ. Differ.* 19, 499–505. doi: 10.1016/j.lindif.2009.05.004
- Forgasz, H. J., and Hill, J. C. (2013). Factors implicated in high mathematics achievement. *Int. J. Sci. Math. Educ.* 11, 481–499. doi: 10.1007/s10763-012-9348-x
- Fouad, N. A., and Smith, P. L. (1996). A test of a social cognitive model for middle school students: math and science. *J. Couns. Psychol.* 43, 338–346. doi: 10.1037/0022-0167.43.3.338
- Gainor, K. A., and Lent, R. W. (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of Black college students. *J. Couns. Psychol.* 45, 403–413. doi: 10.1037/0022-0167.45.4.403
- Gallagher, A. M., and Kaufman, J. C. (2005). *Gender Differences in Mathematics*. New York, NY: Cambridge University Press.
- George, D., and Mallery, M. (2010). *Using SPSS for Windows Step by Step: A Simple Guide and Reference*. Boston, MA: Allyn and Bacon.
- Gun, Z., and Erdem, Z. C. (2014). Uyum analizi yöntemiyle matematik başarısını etkileyen faktörlerin incelenmesi (Assessing the factors affecting the mathematics success via correspondence analysis method). *Adiyaman Üni. Egit. Bilim. Derg. (Adiyaman Univ. J. Educ. Sci.)* 4, 98–118. doi: 10.17984/adyuebd.43951
- Güven, B., and Cabakcor, B. O. (2013). Factors influencing mathematical problem-solving achievement of seventh grade Turkish students. *Learn. Individ. Differ.* 23, 131–137. doi: 10.1016/j.lindif.2012.10.003
- Gwilliam, L. R., and Betz, N. E. (2001). Validity of measures of math- and science-related self-efficacy for African Americans and European Americans. *J. Career Assess.* 9, 261–281. doi: 10.1177/106907270100900304
- Hackett, G., and Betz, N. E. (1981). A self-efficacy approach to the career development of women. *J. Vocat. Behav.* 18, 326–339. doi: 10.1016/0001-8791(81)90019-1
- Hackett, G., and Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *J. Res. Math. Educ.* 20, 261–273. doi: 10.2307/749515
- Halpern, D. F., Wai, J., and Saw, A. (2005). “A psychobiosocial model why females are sometimes greater than and sometimes less than males in math achievement,” in *Gender Differences in Mathematics*, eds. A. M. Gallagher, and J. C. Kaufman (New York, NY: Cambridge University Press), 48–72. doi: 10.1017/CBO9780511614446.004
- Hembree, R. (1990). The nature, effects and relief of mathematics anxiety. *J. Res. Math. Educ.* 21, 33–46. doi: 10.2307/749455
- Hemmings, B., Grootenboer, P., and Kay, R. (2011). Predicting mathematics achievement: the influence of prior achievement and attitudes. *Int. J. Sci. Math. Educ.* 9, 691–705. doi: 10.1007/s10763-010-9224-5
- Hu, X., Leung, F. K. S., and Teng, Y. (2018). The influence of culture on students’ mathematics achievement across 51 countries. *Int. J. Sci. Math. Educ.* 16, S7–S24. doi: 10.1007/s10763-018-9899-6
- Ing, M. (2014). Gender differences in the influence of early perceived parental support on student mathematics and science achievement and STEM career attainment. *Int. J. Sci. Math. Educ.* 12, 1221–1239. doi: 10.1007/s10763-013-9447-3

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.

- Ismail, N. A., and Awang, H. (2008). Differentials in mathematics achievement among eighth- grade students in Malaysia. *Int. J. Sci. Math. Educ.* 6, 559–571. doi: 10.1007/s10763-007-9109-4
- Junge, M. E., and Dretzke, B. J. (1995). Mathematical self-efficacy gender differences in gifted/talented adolescents. *Gifted Child Q.* 39, 22–26. doi: 10.1177/001698629503900104
- Kessel, C., and Linn, M. C. (1996). Grades of scores: predicting future college mathematics performance. *Educ. Meas. Issues Pract.* 15, 10–14. doi: 10.1111/j.1745-3992.1996.tb00573.x
- Lent, R. W., Lopez, F. G., and Bieschke, K. J. (1993). Predicting mathematics-related choice and success behaviors: test of an expanded social cognitive model. *J. Vocat. Behav.* 42, 223–236. doi: 10.1006/jvbe.1993.1016
- Levpušček, M. P., Zupančič, M., and Sočan, G. (2013). Predicting achievement in mathematics in adolescent students: the role of individual and social factors. *J. Early Adolesc.* 33, 523–551. doi: 10.1177/0272431612450949
- Lussier, G. (1996). Sex and mathematical background as predictors of anxiety and self-efficacy in mathematics. *Psychol. Rep.* 79, 827–833. doi: 10.2466/pr0.1996.79.3.827
- Ma, X. (1997). Reciprocal relationships between attitude toward mathematics and achievement in mathematics. *J. Educ. Res.* 90, 221–229. doi: 10.1080/00220671.1997.10544576
- Ma, X., and Kishor, N. (1997). Assessing the relationship between attitude towards mathematics and achievement in mathematics: a meta-analysis. *J. Res. Math. Educ.* 28, 26–47. doi: 10.2307/749662
- McConney, A., and Perry, L. B. (2010). Science and mathematics achievement in Australia: the role of school socioeconomic composition in educational equity and effectiveness. *Int. J. Sci. Math. Educ.* 8, 429–452. doi: 10.1007/s10763-010-9197-4
- Middleton, M. J., and Midgley, C. (1997). Avoiding the demonstration of lack of ability: an underexplored aspect of goal theory. *J. Educ. Psychol.* 89, 710–718. doi: 10.1037/0022-0663.89.4.710
- Miller, H., and Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Pers. Individ. Dif.* 37, 591–606. doi: 10.1016/j.paid.2003.09.029
- Neale, D. C. (1969). The role of attitudes in learning mathematics. *Arithmetic Teach.* 16, 631–640. doi: 10.5951/AT.16.8.0631
- Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemp. Educ. Psychol.* 21, 325–344. doi: 10.1006/ceps.1996.0025
- Pajares, F. (2005). “Gender differences in mathematics self-efficacy beliefs,” in *Gender Differences in Mathematics: An Integrative Psychological Approach*, eds. A. Gallagher, and J. Kaufman (New York, NY: Cambridge University Press), 294–315. doi: 10.1017/CBO9780511614446.015
- Pajares, F., and Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemp. Educ. Psychol.* 24, 124–139. doi: 10.1006/ceps.1998.0991
- Pajares, F., and Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemp. Educ. Psychol.* 20, 426–443. doi: 10.1006/ceps.1995.1029
- Pajares, F., and Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: a path analysis. *J. Educ. Psychol.* 86, 193–203. doi: 10.1037/0022-0663.86.2.193
- Park, K., Brombacher, A., and Brocardo, J. (2021). “The role of mathematics in the overall curriculum,” in *Publications of ICME proceedings materials from ICME-11 Mexico/ICME-11-Topic Study Group Reports/TSG25*.
- Peters, M. L. (2013). Examining the relationships among classroom climate, self-efficacy, and achievement in undergraduate mathematics: a multi-level analysis. *Int. J. Sci. Math. Educ.* 11, 459–480. doi: 10.1007/s10763-012-9347-y
- Pietsch, J., Walker, R., and Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *J. Educ. Psychol.* 95:589. doi: 10.1037/0022-0663.95.3.589
- Randhawa, B. S. (1994). Self-efficacy in mathematics, attitudes, and achievement of boys and girls from restricted samples in two countries. *Percept. Mot. Skills* 79, 1011–1018. doi: 10.2466/pms.1994.79.2.1011
- Randhawa, B. S., Beamer, J. E., and Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *J. Educ. Psychol.* 85, 41–48. doi: 10.1037/0022-0663.85.1.41
- Schukajlow, S., Achmetli, K., and Rakoczy, K. (2018). Does constructing multiple solutions for real-world problems affect self-efficacy? *Educ. Stud. Math.* 100, 43–60. doi: 10.1007/s10649-018-9847-y
- Soni, A., and Kumari, S. (2017). The role of parental math anxiety and math attitude in their children's math achievement. *Int. J. Sci. Math. Educ.* 15, 331–347. doi: 10.1007/s10763-015-9687-5
- Stevens, T., Olivarez Jr, A., Lan, W. Y., and Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *J. Educ. Res.* 97, 208–221. doi: 10.3200/JOER.97.4.208-222
- Stevens, T., Wang, K., Olivarez, Jr., A., and Hamman, D. (2007). Use of self perspectives and their sources to predict the mathematics enrollment intentions of girls and boys. *Sex Roles* 56, 351–363. doi: 10.1007/s11199-006-9180-2
- Taal, M. (1994). How do mathematical experiences contribute to the choice of mathematics. *Sex Roles* 31, 752–769. doi: 10.1007/BF01544291
- Umay, A. (2001). The effects of primary mathematics education program on mathematics' self-efficacy. *J. of Qafqaz Univ.* 8, 1–8.
- Ursini, S., and Sanchez, G. (2008). Gender, technology and attitude towards mathematics: a comparative longitudinal study with Mexican students. *ZDM Math. Educ.* 40, 559–577. doi: 10.1007/s11858-008-0120-1
- Uysal, S. (2015). Factors affecting the Mathematics achievement of Turkish students in PISA 2012. *Educ. Res. Rev.* 10, 1670–1678. doi: 10.5897/ERR2014.2067
- Wang, Z., Osterlind, S. J., and Bergin, D. A. (2012). Building mathematics achievement models in four countries using TIMSS 2003. *Int. J. Sci. Math. Educ.* 10, 1215–1242. doi: 10.1007/s10763-011-9328-6
- Willingham, W. W., and Cole, N. S. (1997). *Gender and Fair Assessment*. Mahwah, NJ: Erlbaum.
- Yao, Y., Kong, Q., and Cai, J. (2018). Investigating elementary and middle school students' subjective well-being and mathematical performance in Shanghai. *Int. J. Sci. Math. Educ.* 16(Supplement 1), S107–S127. doi: 10.1007/s10763-017-9827-1
- Zohar, A., and Gershikov, A. (2008). Gender and performance in mathematical tasks: does the context make a difference? *Int. J. Sci. Math. Educ.* 6, 677–693. doi: 10.1007/s10763-007-9086-7



OPEN ACCESS

EDITED BY
Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY
Nataly Martini,
The University of Auckland, New Zealand
Rita Payan Carreira,
University of Evora, Portugal

*CORRESPONDENCE
Laura Ribeiro
✉ lribeiro@med.up.pt

RECEIVED 26 January 2024
ACCEPTED 20 May 2024
PUBLISHED 14 June 2024

CITATION
Araújo B, Gomes SF and Ribeiro L (2024)
Critical thinking pedagogical practices in
medical education: a systematic review.
Front. Med. 11:1358444.
doi: 10.3389/fmed.2024.1358444

COPYRIGHT
© 2024 Araújo, Gomes and Ribeiro. 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.

Critical thinking pedagogical practices in medical education: a systematic review

Beatriz Araújo¹, Sandra F. Gomes¹ and Laura Ribeiro^{1,2*}

¹Department of Public Health and Forensic Sciences, and Medical Education, Medical Education Unit, Faculty of Medicine of the University of Porto, Porto, Portugal, ²I3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

Introduction: The development of critical thinking (CT) has been a universal goal in higher education. A systematic review of the literature was conducted to evaluate the effectiveness of currently used pedagogical practices to foster CT/ clinical reasoning (CR)/ clinical judgment (CJ) skills and/or dispositions in undergraduate medical students.

Methods: PubMed, Web of Science and Scopus databases were searched from January 2010 to April 2021 with a predefined Boolean expression.

Results: Of the 3221 articles originally identified, 33 articles were included by using PICOS methodology. From these, 21 (64%) reported CR pedagogical practices and 12 (36%) CT pedagogical practices.

Discussion: Overall, pedagogical practices such as cognitive/visual representation, simulation, literature exposure, test-enhancing and team-based learning, clinical case discussion, error-based learning, game-based learning seem to enhance CT/CR skills and/or dispositions. Further research is required to identify the optimal timing, duration and modality of pedagogical interventions for effectively foster CT/CR in medical education.

KEYWORDS

critical thinking, clinical reasoning, skills, dispositions, medical education, systematic review

1 Introduction

Due to demographic and disease pattern changes along with patient's needs and aspirations, healthcare professionals are required to develop new skills such as creativity (1), leadership, teamwork, empathy, and communication skills (2), in order to provide high-quality, safe, and effective patient care (3). Physicians must be prepared to deal with all types of environments and make decisions in situations of crisis and epidemics (4). Therefore, skills for managing people-centered care, managing complex tasks, and creating a positive work culture are needed (5). Being aware of the importance of the development of a set of skills beyond knowledge acquisition in 2017, the United Nations Education, Scientific and Cultural Organization (6) suggested the development of critical thinking (CT) as a learning outcome, which is defined as "the ability to question norms, practices, and opinions; to reflect on own[sic] one's values, perceptions, and actions and to take a position in the sustainability discourse" (7). In fact, CT seems to be a key ingredient for commitment to lifelong learning (8) and a deep learning experience, allowing a better understanding and ability to deal with complex concepts and problems (9). It has been positively related to academic achievements (10) and better patients' assessments, diagnoses, and care in the future (11). Critical thinkers seem to develop a more questioning mind, better critical appraisal abilities, and a positive attitude concerning evidence-based medicine (12). Thus, fostering students to develop CT has been a universal goal "to create better doctors" (13), a

desirable outcome that should be developed in an early stage of their training, as skills develop through experience and practice (11, 14).

According to the American Philosophical Association, critical thinking encompasses a broad set of cognitive skills such as interpretation, analysis, evaluation, inference, explanation, self-regulation, and dispositions, including truth-seeking, open-mindedness, analyticity, systematicity, self-confidence, inquisitiveness, and maturity (15). Definitions of CT in medical education tend to emphasize logical or rational thinking—the ability to reason, analyze information, evaluate alternatives, assess arguments and evidence, and reach relevant and appropriate solutions to a problem (16). Moreover, in medicine, CT has also been described and identified nearly as synonymous with “clinical judgment” (CJ), “clinical reasoning” (CR), “diagnostic thinking,” “problem-solving,” or “type 2 thinking” terms involving a mental process used to think through problems and achieve a final decision (11).

Although CT can be taught, both pedagogical and assessment practices are challenging, and there is no consensus on the most effective teaching approach. In part, the instructional methods are challenging due to the different understanding of CT. For instance, Krupat et al. (17) found that 43% of doctors describe CT as a process, 41% as a skill or ability, and 16% as disposition. Therefore, teaching CT remains both a challenge and a necessity in medical education (9).

Our systematic review aims to investigate the effectiveness of the pedagogical practices that are currently used to foster the development of CT/CR/CJ skills and/or dispositions in undergraduate medical students. The specific objectives of this review are as follows: (1) to identify the pedagogical practices currently in use to promote the development of CT/CR/CJ skills and/or dispositions in undergraduate medical students; (2) to identify the tools that are being used to assess CT/CR/CJ skills and/or dispositions in the above conditions; and (3) to investigate the effectiveness of those pedagogical practices, considering the CT/CR/CJ skills and/or dispositions gains, the assessment tools, and the intervention context.

2 Methodology

This study followed the Cochrane recommendations (18–22) and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Guidelines (23) ([Supplementary material—PRISMA Checklist](#)). The protocol for this systematic review is registered on Open Science Framework (OSF): doi: [10.17605/OSF.IO/8PJ26](https://doi.org/10.17605/OSF.IO/8PJ26).

Although this systematic review was based on the approach used in a previous study (24), it explored the effectiveness of pedagogical practices used to foster CT/CR/CJ skills and dispositions exclusively in undergraduate medical students while also including other pedagogical outcomes.

2.1 Information sources and search strategy

In April 2021, the literature search was conducted in PubMed, Web of Science, and Scopus using the following Boolean expressions: (“Critical Thinking” OR “Clinical Reasoning” OR “Clinical Judgement”) AND (Skill OR Ability OR Disposition OR Attitude) AND (Strategies OR Interventions OR Educat* OR Teach* OR Practice OR Train OR Develop* Analyse* OR Test* OR Evaluate* OR Assess*)

AND (Student* OR Undergraduate* OR School OR Faculty OR College OR High* Education OR Universities) AND (Medic*). The filters “Article title, Abstract, and Keywords” and “All fields” were used in Scopus, PubMed, and Web of Science databases. In addition, a time filter from January 2010 to April 2021 was set, and the studies were included according to the eligibility criteria.

2.2 Eligibility criteria

Studies were deemed eligible according to the following inclusion and exclusion criteria defined using the PICOS tool (25):

- P (Population) — undergraduate medical students;
- I (Interventions) — pedagogical practices to foster the development of CT/CR/CJ skills and/or dispositions;
- C (Comparison) — not applicable;
- O (Outcomes) — CT/CR/CJ skills and/or dispositions gains;
- S (Study design) — Qualitative, quantitative, and mixed studies.

Articles published from January 2010 to April 2021, referring to CT/CR/CJ pedagogical practices as interventions and undergraduate medical students as its target population, were included. Furthermore, articles with a clear description of the pedagogical practices that were used to foster the development of CT/CR/CJ were also included. For this systematic review, studies with qualitative, quantitative, and mixed-method designs were considered.

Letters, short communications, systematic reviews, reviews, and meta-analysis were excluded. Studies lacking a clear methodological description of the intervention, articles outside the scope of undergraduate medical education or studies with no access to the full text despite the attempt to contact the authors, were also excluded during the screening phase. Although the search terms used were in English, no language restrictions were applied in our research strategy.

2.3 Study selection

Studies were screened and selected by two independent reviewers. After duplicate records were removed, first, studies were screened based on the title and then based on the abstract. The remaining records were eligible for full paper reading based on the inclusion and exclusion criteria. Any disagreement was solved by consensus.

2.4 Data collection process

Data were collected, organized, and synthesized in tables based on the following: author(s), publication year, country, study design, objectives, sample, pedagogical approach (pedagogical practices, curricular context, subject specificity, regime, subject, length of the intervention, interventional and/or control group, format, instructional support, and feedback), assessment tools (pre-intervention and post-intervention), and main findings ([Supplementary material — data collection](#)).

2.5 Quality assessment

The included studies underwent a quality assessment by two independent reviewers using the Standard Quality Assessment

Criteria for Evaluating Primary Research Papers from a Variety of Fields (26). The studies received scores for their compliance (“yes” = 2, “partial” = 1; and “no” = 0), with each of the 10 criteria for qualitative studies, 14 criteria for quantitative studies, and both (24) criteria for mixed studies. For each study, a sum score was calculated by adding the scores for these criteria and dividing them by the total possible score (26).

2.6 Data management

To synthesize the information and enable comparison between pedagogical practices, some studies were characterized and grouped according to the characteristics of the pedagogical approach (i.e., cognitive/visual problem representation when a cognitive knowledge organization strategy was used such as mind map, conceptual mapping, illness script, or case vignettes; simulation when a low-fidelity or high-fidelity patient simulation was used; literature exposure when students were exposed and instructed to reflect based on books, literary excerpts, or papers).

Then, pedagogical practices were also divided into CT, CR, or CJ. In studies where multiple pedagogical practices were mentioned (e.g., debates during problem-based learning), only the most prominent studies focusing on the development of CT/CR/CJ skills and/or dispositions were considered to characterize the intervention except for the innovative curriculums that purposefully combine different pedagogical practices.

The curricular context of the pedagogical approach was categorized as curricular (when the CT/CR/CJ pedagogical practices are implemented during the formal curriculum within the context of a specific year and subject content as part of the objectives of a given curricular unit) or extracurricular (when the CT/CR/CJ pedagogical practices are implemented during an elective course or workshop with a certain group of students regardless of the objectives of any curricular unit of the formal curriculum). According to Ennis (27), subject specificity was characterized as follows: general—an approach that attempts to teach CT/CR/CJ abilities and dispositions regardless of the subject content; infusive—an approach where students are encouraged to think critically in the subject (subject-related), in which general CT/CR/CJ principles are made explicit to the students; immersive—an approach where students are encouraged to think critically in the subject (subject-related), in which general CT/CR/CJ principles are not made explicit to the students; or mixed—an approach that combines the general approach with one of the other two, infusive or immersive.

Additionally, the analysis also considered the length of the intervention, the learning regime (face-to-face vs. e-learning), the work format (individual vs. collaborative), the presence or absence of a control group, the number of group interventions, the presence or absence of instructional support (contextualization, facilitators/tutors' guidance, or guidelines), or feedback during or at the end of the interventions.

To simplify the comparison between CT/CR/CJ pedagogical practices and learning outcomes, the assessment approach was classified according to the tool standard (standardized or non-standardized), the domain specificity (general, health sciences, or medical domain), and the assessment tool specificity to evaluate CT/CR/CJ skills and/or dispositions (tests or rubrics, knowledge tests, self-assessment surveys or questionnaires, and focus group session).

Since some studies used more than one assessment tool, only the most effective in evaluating CT/CR/CJ pedagogical practices was used to describe the assessment approach in the following order:

- 1 Domain-specific standardized tests—any quantitative assessment tools used to measure CT/CR/CJ skills and/or dispositions specifically in the health science domain [e.g., The Health Science Reasoning Test (HSRT); Yoon's Critical Thinking Disposition Instrument (YCTDI); Critical Thinking Disposition Assessment (CTDA); Objective Structured Clinical Examination (OSCE); and the Diagnostic Thinking Inventory (DTI)];
- 2 General standardized tests—any quantitative assessment tools used to measure CT/CR/CJ skills and/or dispositions with no specific domain [e.g., The California Critical Thinking Skills Test (CCTST); the Ennis–Weir Critical Thinking Essay Test; and the California Critical Thinking Disposition Inventory (CCTDI)];
- 3 Domain-specific non-standardized tests or rubrics—a quantitative assessment tool previously developed or adapted specifically to assess CT/CR/CJ skills and/or dispositions or students' performance in CT/CR/CJ regarding the domain and the context of the intervention [e.g., Script Concordance Test (SCT), Key Feature Problem Examination (KFPE); Critical Thinking Skills Rating Instrument (CTSRI) — rubric, the Clinical reasoning performance — assessed with 3 knowledge tests as follows: (1) conceptual knowledge test with multiple choice questions, (2) strategic knowledge test with key feature questions, and (3) conditional knowledge with problem solving tests, clinical reasoning test, or problem solving tests assessed with rubrics];
- 4 Domain-specific non-standardized knowledge tests, self-assessment surveys, or questionnaires—a quantitative non-standardized or self-reported assessment tool previously developed to assess the learning experience as student knowledge retention, self-perception, or satisfaction with the pedagogical approach;
- 5 Domain-specific non-standardized focus group sessions—a qualitative non-standardized self-report assessment tool used to assess the learning experience such as the perception or satisfaction of the students with the pedagogical approach.

Considering the most significant assessment tool for the evaluation of the CT/CR/CJ pedagogical practices, the learning outcomes were classified as follows:

- CT/CR/CJ general gain (++) when a statistically significant gain in terms of a general set of CT/CR/CJ skills or dispositions was verified (based on the general score of domain-specific standardized tests, general standardized tests, or domain-specific non-standardized tests or rubrics);
- CT/CR/CJ specific gain (+) when a statistically significant gain was reported for a specific CT/CR/CJ skill or disposition (based on the individual score related to a specific item of domain-specific standardized tests, general standardized tests, or domain-specific non-standardized tests or rubrics);
- CT/CR/CJ no gain (—) when no statistic gain in terms of a general set of CT/CR/CJ skills or dispositions was verified (based on the general score of domain-specific standardized tests, general

standardized tests, domain-specific, or non-standardized tests or rubrics);

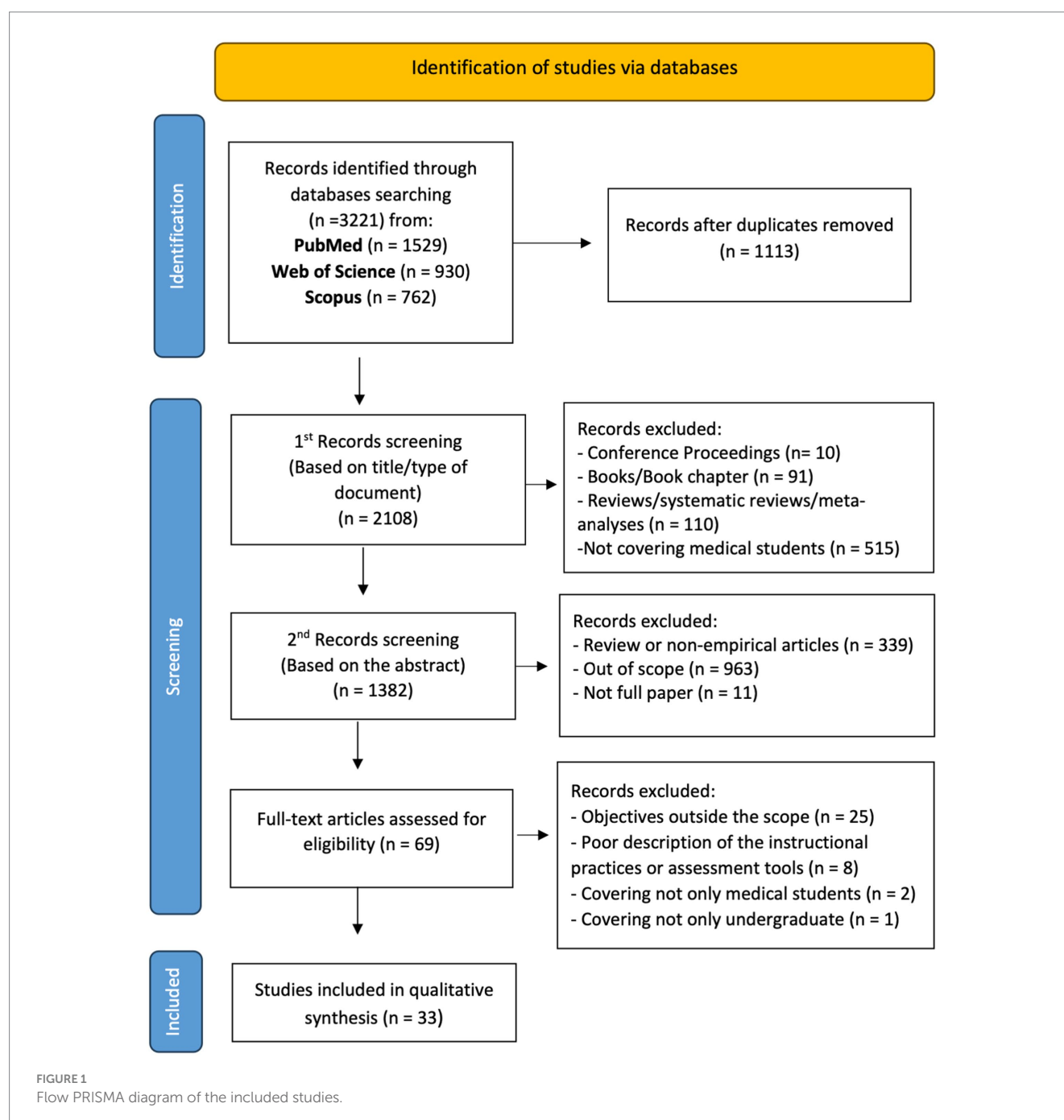
- Other gains (+?) as knowledge, satisfaction, or perception of the development of CT/CR/CJ skills and/or dispositions when a statistically significant gain or a qualitative gain was verified regarding the student learning experiences (based on the general score of domain-specific non-standardized tests, surveys or questionnaires, or a qualitative result of domain-specific focus group sessions).
- No other gains (−?) as knowledge, satisfaction, or perception of the development of CT/CR/CJ skills and/or dispositions when no statistically significant gain or a qualitative gain was verified regarding the student learning experiences (based on the general

score of domain-specific non-standardized tests, surveys or questionnaires, or a qualitative result of domain-specific focus group sessions).

3 Results

3.1 Study selection

A total of 3,221 articles were identified through database screening and were subsequently subjected to a stepwise filtering process (Figure 1). After duplicate records were removed, the remaining 2,108 studies were screened in two phases. In the first



phase, 726 studies were excluded based on the title and document type. In the second phase, 1,314 studies were excluded based on the abstract by applying the exclusion criteria. From the remaining 69 eligible records for full paper reading, 35 were excluded for not meeting the inclusion criteria. The remaining 33 studies were included in this systematic review.

3.2 Quality assessment

The results of the quality assessment are summarized in the Quality Assessment Table ([Supplementary material](#)). The quality rating ranged from 0.62 to 0.95 with a mean of 0.75. Lower ratings were due to a poor description of the sampling strategy and a lack of evidence of both verification procedures and reflexivity in qualitative and mixed studies. In quantitative and mixed studies, lower ratings were due to poor description and appropriateness of the sampling, incomplete baseline/demographic data, and lack of confounding assessment.

3.3 Population

The included studies were markedly different in sample size, the smallest comprising 10 first-year medical students (28) and the largest comprising 214 third-year medical students (29).

The CT/CR/CJ pedagogical practices were applied to students attending the first 5 years of the medical course, with most studies covering the fourth year ($n=8$) followed by the first ($n=6$), second ($n=4$), and third year ($n=4$) ([Tables 1, 2](#)). In addition, of the 33 studies, 8 (40, 41, 46–49, 56, 58) studies recruited students from different academic years. Additionally, 61% of the studies did not mention the students' age and 30% did not mention their gender.

3.4 Pedagogical practices, assessment tools, and learning outcomes

A diversity of pedagogical practices was used to foster the development of CT/CR/CJ skills and/or dispositions in undergraduate medical students ([Tables 1, 2](#)). The most frequently used were cognitive/visual representation approaches (8/33, 24.2%) as mind map (33), concept mapping (14, 40, 41), clinical-anatomical case vignettes (42), and illness script (39, 43, 59). In addition, simulation was also frequent (6/33, 18.2%) through virtual patients—low-fidelity patient simulations (48, 49, 51, 52) and manikins—high-fidelity patient simulations (35, 50).

Literature exposure was also reported (3/33, 9.0%) involving drama-based scenarios (30) and literature and film analysis (31) and integrating a gender perspective into literature studies (32).

Other approaches, such as test-enhancing learning (53, 54), team-based learning (46, 58), clinical case discussion (55, 56), case creation, team-based learning (46), game-based learning (44), error-based learning (47), dialog narrative approach (37), reflective writing (9), journal club (36), debate (34), and the explicit CT instruction approach (38), were mentioned in a few studies. Furthermore, some studies (29, 57) employed an

innovative curriculum that combines different pedagogical approaches.

Pedagogical practices, such as cognitive/visual representation approaches, mind maps (33), illness-script (39, 43, 59), clinical-anatomy case vignettes (42), literature exposure by integrating the gender perspective into literature studies (32), simulation (50), clinical case discussion (55), team-based learning (46), test-enhancing learning (53, 54), explicit CT instruction approach (38), and the innovative curriculums (29, 57), were compared with the traditional format while case creation was compared with case-based learning (28).

Overall, CT/CR/CJ pedagogical practices were employed in both “curricular” (16/33; 48.5%) and “extracurricular” (17/33; 51.5%) contexts, predominantly with an “immersive” approach (25/33; 75.8%) and in a face-to-face regime (27/33; 78.8%).

The subjects/topics covered by the interventions were quite variable. Overall, as described in [Tables 1, 2](#), they were mainly related to the medical curriculum (31/33; 93.9%), with just a few more associated with medical humanities (4/33; 12.1%) or not directly related to the typical medical curriculum topics (2/33; 6.1%) (33, 38).

Regarding other factors of the intervention, it is important to highlight that the length and/or the number of sessions were fairly different, with shorter interventions being implemented in a single session (9/33; 27.3%) and the longest over a year-long course. In addition, most of the studies (23/33; 69.7%) mentioned a collaborative format promoting peer learning with groups ranging from 2—discussion group mate and “think-pair-share” (32, 42)—to 15 students (59). Furthermore, most of the studies (29/33; 87.8%) mentioned the involvement of instructional support during the sessions by contextualizing the pedagogical practice and/or the aims of the session, providing guidelines/worksheets to provide clear and standardized instructions or providing facilitators/tutors' guidance during the session, either by rising and/or answering questions or by moderating discussions. In addition, most studies (24/33; 72.7%) mentioned that debriefing or feedback was provided during or at the end of the interventions.

[Table 3](#) shows the assessment tools that were identified in this review. Overall, the development of CT/CR/CJ skills and/or dispositions or students' performance in CT/CR/CJ was mainly assessed by domain-specific non-standardized tests or rubrics (19 of 26; 73.1%), while student's knowledge, satisfaction, or perceptions about the efficacy of the pedagogical practice to foster the development of CT/CR/CJ skills were mainly assessed through domain-specific knowledge tests, self-assessment surveys, or questionnaires (5 of 7; 71.4%).

Of the 33 articles, 19 (57.6%) reported CT/CR/CJ general gains (++), 3 (9.1%) reported CT/CR/CJ specific gains (+), and 4 (12.1%) reported CT/CR/CJ no gains (—). The remaining seven (21.2%) reported positive effects of the CT/CR/CJ pedagogical practices, but the gains were measured by considering improvements in knowledge or by assessing the students' satisfaction or perceptions of the effectiveness of the intervention ([Figure 2; Table 3](#)).

Learning outcomes were also analyzed according to curricular context, subject specificity, regime, format, presence or absence of instructional support, and feedback ([Tables 1, 2](#)). [Table 4](#) summarizes the methodological characteristics of the articles.

TABLE 1 Characteristics of the studies employing critical thinking pedagogical practices.

Author, data	Sample	Pedagogical approach						Assessment tool	Outcomes
		Pedagogical practice		Specificity	Subject/topic	Length	Regime		
Archila (30)	<i>n</i> = 91 62% (56) F 1st y	Literature exposure	Argument evaluation -drama-based CT classroom scenarios	Immersive	ethics, social responsibility, and scientific work	1 session (60 min)	Individual	Domain-specific non-standardized knowledge tests, self-assessment surveys or questionnaires	Other gains (+?)
Kim (31)	<i>n</i> = 51 25% (13) F 2nd y		Read/watch the material + group discussion sessions + writing a critical essay	Immersive	Social sciences and humanities	15 weeks	Collaborative (groups of 10–11)	Domain-specific standardized tests	CT/CR/CJ Specific gain (+)
Liao and Wang (32)	<i>n</i> = 82 –		Gender perspective into literature + reflection + e-discussion	Immersive	Gender literature studies	15 weeks (2xs/ week)	Individual + Collaborative (with group mate)	General standardized tests	CT/CR/CJ Specific gains (+)
D'Antoni et al (33)	<i>n</i> = 131 52% (68) F 1st y	Cognitive/visual representation	Mind map <i>vs</i> standard notetaking	Mixed	Cacti and other succulent plants	1 session (205 min)	Individual	Domain-specific standardized tests	CT/CR/CJ No gain (–)
Bixler et al. (14)	<i>n</i> = 33 – 4th y		Concept mapping in small groups	Immersive	Pediatric topics	4 sessions; (1 h/ session)	Collaborative (groups of 4–6)	General standardized tests	CT/CR/CJ No gain (–)
Mumtaz and Latif (34)	<i>n</i> = 182 100% F 2nd y	Debate (during PBL)		Immersive	Areas of controversy in medicine	1 year/2 semesters 6–7 sessions	Collaborative (groups of 10–13)	Domain-specific non-standardized knowledge tests, self-assessment surveys or questionnaires	Other gains (+?)
Nguyen et al. (35)	<i>n</i> = 120 – 1st y	Simulation —high-fidelity patient simulations – manikins		Immersive	Physiology	17 weeks (55–60 min each session)	Collaborative (groups of 6)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Banerjee et al. (36)	<i>n</i> = 54 – 1st y	Mentored journal clubs : 6D-approach		Immersive	Molecular biology and principles of genetics	15 weeks (7 sessions)	Collaborative (groups of 3–4)	Domain-specific non-standardized knowledge tests, self-assessment surveys or questionnaires	Other gains (+?)
Sahoo and Mohammed (9)	<i>n</i> = 188 56% (105) F 4th y	Reflective writing : collaborative research protocol writing		Immersive	Ophthalmology	4 weeks	Collaborative (Small groups)	Domain-specific non-standardized knowledge tests, self-assessment surveys or questionnaires	Other gains (+?)
Ghiam et al. (37)	<i>n</i> = 100 – 2nd y	Dialogue narrative approach : storytelling format + question-answer conversational style at regular intervals and flipped classroom		Immersive	Thyroid physiology	1 session (50 min)	Individual	Domain-specific non-standardized focus group sessions	Other gains (+?)
McClintic et al. (29)	<i>n</i> = 214 44% (94) F 3th y	Innovative curriculum : entrustable professional activities ^(a)		Immersive	Surgical clerkship	8 sessions	Individual + Collaborative (small groups)	Domain-specific standardized tests	CT/CR/CJ general gains (++)
Taghinezhad and Riasati (38)	<i>n</i> = 140 Both genders	Explicit CT instructions ^(b)		Infusive	Parking problem in a small town/CT concepts	1 semester; 15 weeks (3 h each)	Individual	General standardized tests	CT/CR/CJ Specific gains (+)

PBL-problem based learning; (a) deliberated-practice structured orientation + small group sessions + online quizzes + extensive didactics + team-based learning and simulation exercises + clinical portfolio; (b) providing CT explicit instruction, teaching students how to make use of those techniques to synthesize, analyze, and evaluate information, presenting support materials in CT classrooms (including leaflets and models) of the instructional techniques, leading Socratic discussions based on the elements and criteria suggested in the instructional techniques, assigning classroom activities and giving them adequate time to practice each skill, using both oral and written techniques, and assessing students' performance.

TABLE 2 Characteristics of the studies employing Clinical Reasoning pedagogical practices.

Author, data	Sample	Pedagogical approach						Assessment tool	Outcomes
		Pedagogical practice		Specificity	Subject/topic	Length	Regime		
Lee et al. (39)	n = 53 51% (33) F 4th y	Cognitive / visual representation	Problem representation + illness script – web-based CR problems	Infusive	Two scenarios: 1) an elderly man with a persistent cough; 2) a middle-aged woman with an acute swollen and painful left leg	1 session (3 h)	Collaborative (small groups)	Domain-specific standardized tests	CT/CR/CJ no gain (–)
Wu et al. (40)	n = 29 66% (19) F 4th y (19) 3rd and 5th y (10)		Computer-based argument mapping + concept mapping	Immersive	Kidney disease	4 weeks (5 h/ week)	Individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Si et al. (41)	n = 95 41.1% (39) F 1st y (44) and 2nd y (51)		Argumentation with the concept map method during PBL – according to Toulmin’s model of argumentation	Immersive	Clinical cases	3 sessions (2 h each) 3 weeks	Collaborative (groups of 7–8)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Kumar et al. (42)	n = 150 – 1st y		clinical-anatomical case vignettes for analyzing clinical cases	Immersive	Varicose veins and thyroid goiter	2 sessions (1 h each)	Collaborative (groups of 2)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gains (++)
Moghadami et al. (43)	n = 100 53% (53) F 4th y		Illness script - > small group discussion (think aloud) - > debriefing	Immersive	Cirrhosis / CHF / Nephrotic Syndrome /leg edema	2 sessions (7 h each) 4 weeks	Individual + collaborative (small group discussion + open discussion)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gains (++)
Middeke et al. (44)	n = 112 56.3% (63) F 5th y	Game-based learning —Serious Game (playing EMERGE) vs small-group PBL		Immersive	internal medicine (Cardiology, pulmonology nephrology, rheumatology hematology, oncology)	6 weeks 10 sessions (90 min each)	individual vs collaborative (groups of 6–8)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Chandrasekar et al. (28)	n = 10 – 1st y	Case creation —“build-a-case” approach vs traditional CBL		Immersive	Dilated cardiomyopathy	1 session (3 h)	Collaborative (groups of 5)	Domain-specific non-standardized focus group sessions	Other gain (+?)
Brich et al. (45)	n = 122 57.4% (70) F 3rd (92) and 4th (30) y	Team-based Learning	(symptom-oriented small-group seminars or sTBL units)	Infusive	Neurology topics (vertigo, acute back pain, first epileptic seizure, and acute altered mental status)	2 weeks	Collaborative (small groups)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Jost et al. (46)	n = 26 58% (15) F 4th (18) and 5th y (8)		TBL vs. non-TBL	Immersive	(vertigo, acute back pain, first epileptic seizure, and acute altered mental status)	4 sessions (90 min/session)	Collaborative (5–7 students) vs. individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Klein et al. (47)	n = 84 67% (56) F clinical semesters	Error-based learning: learning from errors in a clinical case-based online learning environment (text vignettes)		Immersive	arterial hypertension	1 session (no time limit)	Individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)

(Continued)

TABLE 2 (Continued)

Author, data	Sample	Pedagogical approach					Assessment tool	Outcomes	
		Pedagogical practice		Specificity	Subject/topic	Length			Regime
Schubach et al. (48)	n = 56 67% (38) F 4th and 5th y	Simulation	VPs + key feature-based instructions on multiple short cases vs VPs + systematic instruction on a few long cases	Immersive	Acute abdomen gastrointestinal bleeding	3 sessions (90 min/session) 2 weeks	Individual work -> small group discussion -> moderated group discussion	Domain-specific non-standardized tests or rubrics	CT/CR/CJ No gain (–)
Isaza-Restrepo et al. (49)	n = 20 – 1st to 3rd y		Web-based VPs: low-fidelity simulator of clinical cases	Infusive	Abdomen pain of different etiology	16 weeks- 2 sessions per week (2 h per session)	Collaborative (small groups)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Mutter et al. (50)	n = 96 – 4th y		High-fidelity simulation (patient case scenario with vs without manikin)	Immersive	Chest pain	1 session (2 h)	Collaborative (6 students)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gains (++)
Watari et al. (51)	n = 169 37% (63) F 4th y		VPs (*Body Interact, Portugal)	Infusive	Two scenarios: 1) a 55-year-old male with altered mental status; 2) a 65-year-old male with acute chest pain	1 session (2 h)	Individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gains (++)
Kleinert et al. (52)	n = 62 – 3rd year		VPs (ALICE)	Immersive	Esophageal cancer (different tumor stages and different therapeutic options)	-	Collaborative (small groups - <5 MS)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Ludwig et al. (53)	n = 93 64.5% (60) F 4th y	Test-enhanced learning	video-based key feature questions vs repeated testing with text-based on key feature questions	Immersive	cardiology, pulmonology, nephrology, rheumatology, hematology, and oncology	10 weeks 1 session per week (45 min) ^(a)	Individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Raupach et al. (54)	n = 87 58,6% (51) F 4th year		Computer CBL + augmented case presentation + key feature questions vs repeated CBL (long case narratives)	Immersive	Cardiology, pulmonology nephrology, rheumatology hematology, and oncology	10 weeks 1 session per week (45 min) ^(a)	Individual	Domain-specific non-standardized knowledge tests, self-assessment surveys, or questionnaires	CT/CR/CJ General gain (++)
Montaldo Lorca and Herskovic (55)	n = 64 – 3rd y	Clinical case discussion (CCD)	Prototypical clinical cases (lectures and tutorial sessions with patients vs with patients + discussion seminars).	Immersive	Semiology and Internal Medical clerkship cardiac and pulmonary pathology syndromes	6 months	Collaborative (small groups)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Weidenbusch et al. (56)	n = 90 65.5% (59) F 1st to 4th y		(Live-CCD vs Video – CCD vs Paper – cases)	Immersive	paresthesia, fever, and respiratory failure, rapidly progressive respiratory failure	3 weekly—5 sessions (90 min each) ^(b)	collaborative vs. individual	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)
Bonifacino et al. (57)	n = 67 – 3rd year	Innovative curriculum: six interactive online modules – didactic videos, simulated clinical cases, and interactive prompts for open-ended MCQ; and a case-based workshop		Infusive	diagnostic error, cognitive psychology of decision-making, specific CR skills, semantic qualifiers and problem representation, cognitive biases, and heuristics	4 weeks	Individual and collaborative (small groups of 3–4 students + large groups)	Domain-specific non-standardized tests or rubrics	CT/CR/CJ General gain (++)

PBL – problem based learning; CBL – case based learning; TBL – team based learning. VPs – Virtual Patients simulation; CBL – case-based learning; CCD – clinical case discussion; MCQ – multiple choice questions (a) 13th week (exit exam) 9th month (retention test); (b) 3rd week (exit exam), 5th week (retention test).

TABLE 3 Assessment tools mentioned in the selected studies (n = 33).

Classification	Description of the assessment tools
Domain-specific standardized tests (n = 5)	• Health Science Reasoning Test (HSRT) (33) • Diagnostic Thinking Inventory (DTI) (39) • Yoon's Critical Thinking Disposition Instrument (YCTDI) (31) • Observed Structured Clinical Examinations (OSCE) (29) • Critical Thinking Disposition Assessment (CTDA) (32)
General standardized tests (n = 2)	• California Critical Thinking Skills Test (CCTST) (14) • The Ennis-Weir Critical Thinking Essay Test + The California Critical Thinking Dispositions Inventory (CCTDI) (38)
Domain-specific non-standardized tests or rubrics (n = 19)	• Test: students' performance (52) • Test: Key feature problem examination (KFPE) (44, 46, 53, 54) • Test: Script Concordance Test (SCT) (43, 48, 50) • Test: Medical CR performance assessed with three different knowledge tests (47, 56) • Test: CR Test (10 problem clinical cases) (55) • Test: Multiple choice questions to assess knowledge and CR (42, 51) • Rubric: Critical Thinking Skills Ranking Instrument (CTSRI) (35) • Rubric: Dual-mapping scores assessed based on a set predefined rubric (40) • Rubric: Students performance assessed based on a matrix to measure CR skills Isaza-Restrepo et al. (49) • Rubric: Students problem-solving performance with problem solving test and using a scoring rubric (41) • Rubric: Interpretive summary, Differential diagnosis, Explanation of reasoning and Alternatives (IDEA) tool to assess CR skills in student hospital admission notes (57)
Domain-specific knowledge tests, self-assessment surveys, or questionnaires (n = 5)	• Self-assessment surveys or questionnaires: Students' perception questionnaire (34) • Self-assessment surveys or questionnaires: students' responses to a close-ended and open-ended questionnaire (30) • Knowledge test: Students' final grade at the end of the course + self-assessment surveys or questionnaires: students' feedback at the end of the course – MBRU questionnaire (36) • Self-assessment surveys or questionnaires: students' perception of the enhancement of CT + survey questionnaire (9) • Self-assessment surveys or questionnaires: self-perception survey + open-ended comments about the exercise (59)
Domain-specific focus group sessions (n = 2)	• Students focus groups to assess students' perceptions of the effectiveness of the approach (37) • Students and faculty focus groups to compare students' case creation experiences with traditional case-based learning sessions (28)

3.5 Critical thinking versus clinical reasoning pedagogical practices

Of the 33 articles eligible for the review, 21 (63.6%) reported CR pedagogical practices and 12 (36.4%) reported CT pedagogical practices. No article reported interventions related to CJ pedagogical practices. Literature exposure, debate, journal club, reflective writing, dialog narrative approach, explicit CT instructions, high-fidelity patient simulation, and cognitive/visual representation (such as mind map and conceptual mapping) seem to be mainly used as CT pedagogical practices, while clinical case discussion, case creation, team-based learning, game-based learning, error-based learning, test-enhancing learning, low-fidelity simulation, and cognitive/visual representations (such as illness script, concept map, and clinical-anatomical case vignettes) seem to be mainly used as CR pedagogical practices. In addition, CT pedagogical practices were mostly applied to students attending the first (n=4) and second (n=3) academic years, while CR pedagogical practices, although applied in all academic years, were mostly used in the fourth year (n=6), followed by third (n=3), fourth plus fifth (n=2), and first year (n=2). Moreover, the studies that recruited students from different academic years only used CR pedagogical practices.

Regarding the multiple pedagogical practices, cognitive/visual representation and simulation were applied to the first 5 years of the curriculum. In addition, literature exposure, journal club, case creation, debate, and dialog narrative approach were implemented in the first 2 years of the curriculum, while game-based learning, team-based learning, error-based learning, and reflective writing were employed from the third to fifth years. Furthermore, domain-specific knowledge tests, general standardized tests, domain-specific knowledge tests, self-assessment surveys, or questionnaires were mostly used to assess CT development, while domain-specific non-standardized tests or rubrics were mainly used to assess CR.

As for the development of CT/CR/CJ skills and dispositions based on targeted outcomes, CT pedagogical practices reported mostly other gains (+?) and specific gains (+), while CR pedagogical practices reported mainly general gains (++) (Figures 3, 4).

4 Discussion

Due to differences in study designs and methodologies, as well as the diversity of CT/CR/CJ pedagogical practices and respective assessment tools used in the selected studies, it is challenging to

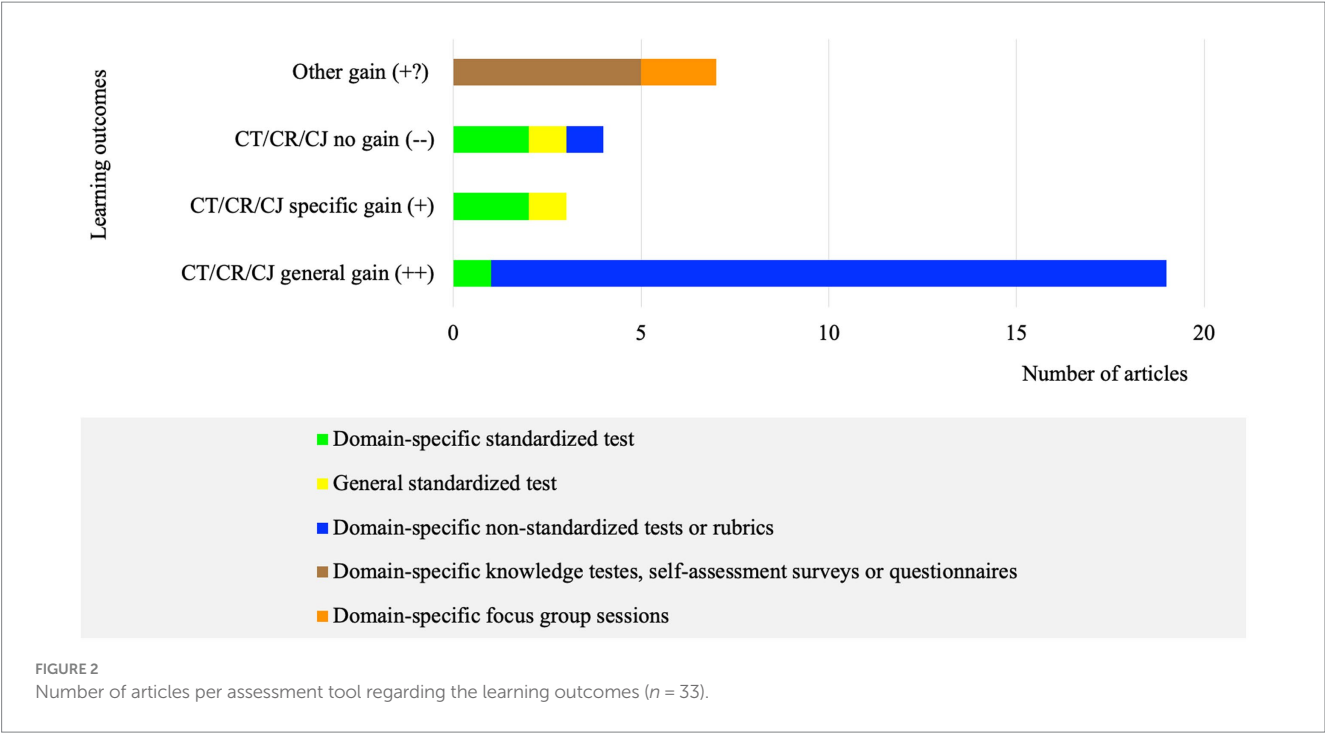


TABLE 4 Methodological characteristics of the selected articles (n = 33).

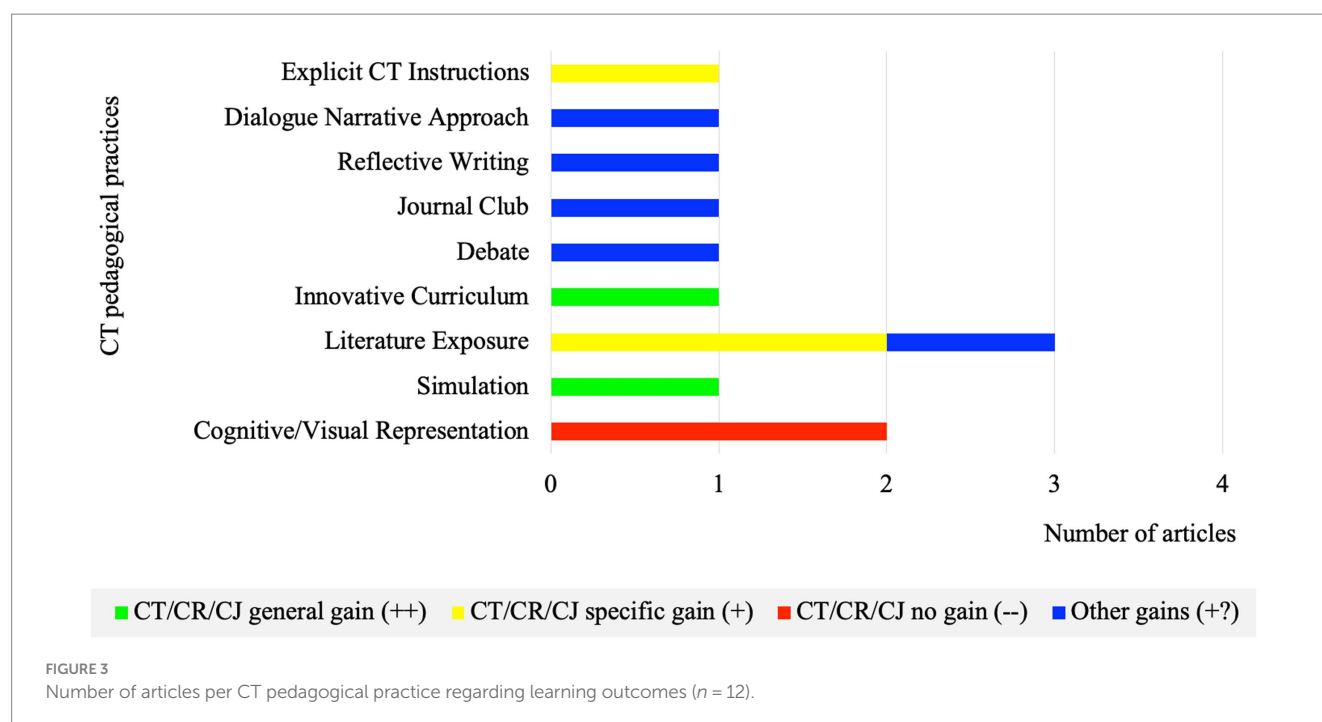
Study	Assessment	n papers	CT/CR/CJ general gain (++)	CT/CR/CJ Specific Gain (+)	CT/CR/CJ no gain (—)	Other gains (+?)
Quantitative	Pre- and post-test	14	Watari et al. (51), Moghadami et al. (43), Klein et al. (47), Ludwig et al. (53), Weidenbusch et al. (56), Si et al. (41), Kleinert et al. (52), and Raupach et al. (54)	Kim (31), Kim (32), and Taghinezhad and Riasati (38)	D'Antoni et al. (33), Bixler et al. (14), and Lee et al. (39)	—
	Only post-test	8	Nguyen et al. (35), Mutter et al. (50), Kumar et al. (42), Jost et al. (46), Middeke et al. (44), Bonifacino et al. (57), and McClintic et al. (29)	—	Schubach et al. (48)	—
Qualitative	Pre- and post-test	0	—	—	—	—
	Only post-test	2	—	—	—	Banerjee et al. (36) and Ghiam et al. (37)
Mixed	Pre- and post-test	4	Montaldo Lorca and Herskovic (55), Wu et al. (40), and Isaza-Restrepo et al. (49)	—	—	Chandrasekar et al. (28)
	Only post-test	5	Brich et al. (45)	—	—	Archila (30), Mumtaz and Latif (34), Sahoo and Mohammed (9), and Levin et al. (59)

identify the most effective pedagogical practices in fostering the development of CT/CR/CJ skills and/or dispositions in medical students.

Nevertheless, pedagogical practices, such as cognitive/visual representation, simulation, literature exposure, test-enhancing and team-based learning, clinical case discussion, error-based learning, game-based learning, explicit CT instructions, and the innovative curriculum approaches, seem to be effective in the development of the CT/CR/CT skills and/or dispositions as most of them reported CT/

CR/CJ general gains (++)

Alternatively, pedagogical practices, such as debate, journal club, reflective writing, dialog narrative approach, and case creation, seem to have a positive impact on students' learning experience, showing improvements in student's knowledge, satisfaction, and perception of the development of CT/CR/CJ. However, these outcomes do not directly express CT/CR/CJ since, knowledge *per se* is not enough for its development (60). In addition, the concept of CT/CR/CJ is still complex and ambiguous, leading to different perceptions among students, teachers, and experts on CT



(17). Therefore, no conclusion can be inferred either from knowledge improvements or students' perceptions of the development of CT/CR/CJ and the effective development of those skills.

When comparing CT and CR pedagogical practices, despite the differences in the medical curriculum length across countries, usually the first years are basic science-oriented while the last years are clinical training-orientated. This may be particularly interesting given that CT pedagogical practices seem to be mainly employed during the first 2 years of the curriculum, and CR pedagogical approaches are more commonly used in the last years, which may be related to the fact that CR constitutes CT application within a clinical context (61).

Most studies seem to foster CT mainly through the development of skills and dispositions. On the other hand, studies fostering CR development, in line with past evidence, were more profession-oriented and focused mainly on the development of diagnostic and decision-making skills (24, 62). In addition, the studies evaluating the effectiveness of pedagogical practices were mainly focused on fostering and assessing both CT and CR skills, whereas studies focused on dispositions were limited ($n=3$) and only used in the context of CT (31, 32). Therefore, this can be related to how doctors conceptualize CT (i.e., in their definitions) (17). As we know, CT "requires mastery of context-specific knowledge to evaluate specific beliefs, claims, and actions" (63). Therefore, it may be useful to apply pedagogical practices that encourage the development of CT dispositions as "open-mindedness, willingness to reconsider, honesty about personal biases and persistence" (15) in the early years of the medical curriculum. As students acquire a deeper knowledge, pedagogical practices that foster CT/CR skills, more oriented to clinical decision and problem solving, could be explored. According to students' needs and learning objectives, CT and CR pedagogical practices could be combined to achieve a more comprehensive development of skills and dispositions.

Regarding the "curricular" and "extracurricular" approaches, we have found contrasting results, with both reporting CT/CR/CJ

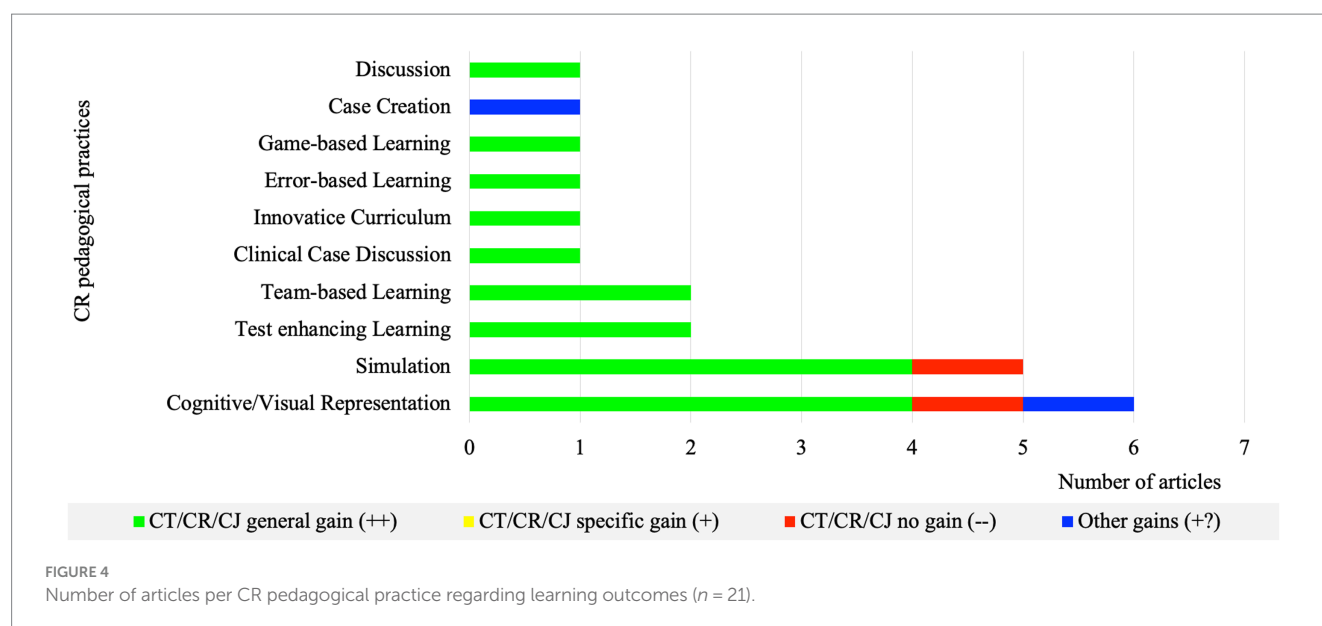
general gains (++) (56.3% versus 58.8%) and CT/CR/CJ no gains (--) (6.3% versus 17.6%). In agreement with previous literature, some authors believe that CT skills can be assessed regardless of the context, while others disagree (64).

Regarding the subject specificity, most studies that adopted an "immersive" followed by an "infusive" approach positively impacted the development of CT/CR/CJ skills and/or dispositions and students' learning experiences (92.0 and 85.7%, respectively). Considering these results, and that the only study with a mixed approach reported CT/CR/CJ no gains (--), it seems that an approach in which students are encouraged to think critically about a subject (subject-related) could be more effective, especially when they have prior knowledge on the topic. In fact, some studies show the need for a sustained specific-knowledge background to enable its application in more complex systems (60, 65). In contrast, when comparing the "immersive" and "infusive" approaches, despite the limited number of studies reporting an "infusive" approach ($n=7$), we can suggest that the most effective are those that make these principles explicit to students. Furthermore, gains resulting from the explicit CT instructions approach corroborate this (38). Therefore, it would be interesting to incorporate CT as a specific subject in the medical curriculum, ideally during the first year.

Regarding the intervention length, some studies report that longer, progressive, and continuous interventions can lead to better outcomes, indicating that length may be an important factor in the development of CT. Although supported by a previous systematic review (24), the results herein presented are not enough to support this association.

Regarding the regime of the approach, most strategies applied the face-to-face approach, although both face-to-face and e-learning methodologies seem to positively impact the development of CT/CR/CJ.

Both individual and collaborative approaches seem to have a positive impact on the development of CT/CR. However, the learning experience can be improved by students' engagement in discussions



with each other (56). In addition, pedagogical practices seem to increase their efficacy in the following order: passive < active < constructive < interactive learning environments (66). This highlights the need to better understand the role of group interaction in the development of CT when assessing the impact of the number of students per group, both large (7 to 8 medical students) and small groups (2 medical students), and induced CT/CR/CJ general gains (++) . In addition to group size, constitution seems to play a pivotal role in the effectiveness of the pedagogical practices. For instance, in heterogeneous groups, the strongest students may end up doing all the work, especially when a limited amount of time is available to perform a task. However, heterogeneity may also breed different backgrounds and perspectives on a subject, thus enriching discussion and aiding group productivity (67). In contrast, a homogeneous group will probably operate on the extremes, with the strongest groups completing their work more quickly, while the weaker groups need more time or do not complete the task by themselves (67).

Additionally, the pedagogical practice effectiveness seems to depend on the tutor and his CT proficiency (44, 68, 69). Although most studies mentioned instructional support (87.9%) or feedback (66.7%), only a minority considered their educational role and impact in fostering CT. Kim (31) highlights that activities based on interactions and feedback (e.g., group discussion and narrative comments) were more advantageous for CT development than those done individually (i.e., book reading and essay writing) (31). In agreement, some authors (28, 46, 70) suggest that the presence of an instructor through students' guidance and prevention of misunderstandings or misconceptions can benefit both the pedagogical approach and learning experiences. Instructional support and feedback ensure that learners with learning difficulties are adequately supported in their learning process (47).

The lack of results regarding the learning outcomes may be related to differences between studies in terms of methodology and design, particularly in the intervention length, subject specificity, and assessment tools. Hence, negative outcomes might have resulted from the unfamiliarity with the approach that may have led to less proficiency (37) or from the short duration of studies with high

cognitive load interventions that may compromise student adjustment to pedagogical practice and consequent performance (48).

Overall, developing CT/CR/CJ skills represents a complex learning challenge. The learner is involved in a combination of meaningful learning activities that promote both knowledge, by integrating new information into pre-existing frameworks (71), and skills such as observation, analysis, evaluation, contextualization, questioning, and, finally, reflection on a subject or a problem (72, 73). A learning environment that promotes individual growth through peer instruction, meaningful learning, and learning by doing, in which students play an active role in the learning process and receive proper feedback, is essential to ensure the efficacy of a pedagogical approach (49).

To the best of our knowledge, this is the first systematic review focused on the effectiveness of CT/CR/CJ pedagogical practices in medical education. Some limitations were identified. Although most studies were mainly focused on pedagogical practice, the presence of other approaches may have had an impact on the results. In addition, gains may also have been influenced by the educational environment, other concurrent courses, or the existence of residual or unmeasured confounders (31). Caution must also be taken with the studies describing assessment tools not previously validated for the study population (e.g., general standardized tests), as well as when a domain-specific non-standardized test or rubric was used, compromising the reproducibility of the intervention in other fields, as well as when studies relied upon a post-test assessment and/or there was no control group or other comparative intervention. Finally, no language restrictions were applied to our research strategy, but the search terms we used were in English. Given that most of the non-English language journals translate the abstract into English, we suggest that the search terms used only in English did not significantly affect the search outcomes.

4.1 Future research

Future studies should be planned and designed detailing the context of the intervention, the subject specificity, objectives,

length, regime (face-to-face vs. e-learning), format (individual vs. collaborative), the role of the facilitators, the presence or absence of feedback, and the teacher's expertise in CT/CR/CJ. Future research should also explore how to assess the long-term impact of the interventions on CT retention over time.

5 Conclusion

Pedagogical practices that actively engage undergraduate medical students in the learning process are likely more effective than traditional lectures in fostering the development of CT/CR/CJ skills and/or dispositions. However, comparison between practices is not easy due to the limited number of studies and diverse methodologies and assessment tools employed. Despite these challenges, our systematic review raises important questions about the timing, length, curricular context, regime (face-to-face vs. e-learning), and format (individual vs. collaborative) of interventions that should be carefully considered to enhance the effectiveness of the pedagogical approaches. Furthermore, it acknowledges the complexity of fostering critical thinking in medical education, recognizing that there is no “one-size-fits-all solution.” Overall, we can conclude that different pedagogical practices should be used and combined throughout the curriculum considering diverse learning environments and student needs to effectively enhance CT skills and dispositions of medical students.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

References

- Patterson F, Zibarras LD. Selecting for creativity and innovation potential: implications for practice in healthcare education. *Adv Health Sci Educ Theory Pract*. (2017) 22:417–28. doi: 10.1007/s10459-016-9731-4
- Guragai M, Mandal D. Five skills medical students should have. *J Nepal Med Ass*. (2020) 58:269–71. doi: 10.31729/jnma.4878
- Kerins J, Smith SE, Phillips EC, Clarke B, Hamilton AL, Tallentire VR. Exploring transformative learning when developing medical students' non-technical skills. *Med Educ*. (2020) 54:264–74. doi: 10.1111/medu.14062
- Revere D, Turner AM, Madhavan A, Rambo N, Bugni PF, Kimball A, et al. Understanding the information needs of public health practitioners: a literature review to inform design of an interactive digital knowledge management system. *J Biomed Inform*. (2007) 40:410–21. doi: 10.1016/j.jbi.2006.12.008
- OECD Health Division team. (2018). Feasibility study on health workforce skills assessment—supporting health workers achieve person-centred care. Available at: <https://www.oecd.org/els/health-systems/Feasibility-Study-On-Health-Workforce-Skills-Assessment-Feb2018.pdf>
- UNESCO. (2017). Education for sustainable development goals: learning objectives. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000247444>
- Straková Z, Cimermanová I. Critical thinking development—a necessary step in higher education transformation towards sustainability. *Sustain For*. (2018) 10:3366. Available at: <https://www.mdpi.com/2071-1050/10/10/3366>
- Kuiper RA, Pesut DJ. Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. *J Adv Nurs*. (2004) 45:381–91. doi: 10.1046/j.1365-2648.2003.02921.x
- Sahoo S, Mohammed CA. Fostering critical thinking and collaborative learning skills among medical students through a research protocol writing activity in the curriculum. *Korean J Med Educ*. (2018) 30:109–18. doi: 10.3946/kjme.2018.86
- Ross D, Loeffler K, Schipper S, Vandermeer B, Allan GM. Do scores on three commonly used measures of critical thinking correlate with academic success of health professions trainees? A systematic review and meta-analysis. *Acad Med*. (2013) 88:724–34. doi: 10.1097/ACM.0b013e31828b0823
- Papp KK, Huang GC, Lauzon Clabo LM, Delva D, Fischer M, Konopasek L, et al. Milestones of critical thinking: a developmental model for medicine and nursing. *Acad Med*. (2014) 89:715–20. doi: 10.1097/acm.0000000000000220
- Shehata GM, Zaki A, Dowidar NL, El Sayed I. Critical thinking and attitude of physicians toward evidence-based medicine in Alexandria, Egypt. *J Egypt Public Health Ass*. (2015) 90:115–20. doi: 10.1097/01.EPX.0000470848.07476.f9
- Zayapragassarazan Z, Vikas M, Sitanshu SK, Gitanjali B. Understanding critical thinking to create better doctors. *J Adv Med Educ Res*. (2016) 1:9–13.
- Bixler GM, Brown A, Way D, Ledford C, Mahan JD. Collaborative concept mapping and critical thinking in fourth-year medical students. *Clin Pediatr*. (2015) 54:833–9. doi: 10.1177/0009922815590223
- Facione P. (1990). Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations. American Philosophical Association. Available at: <http://files.eric.ed.gov/fulltext/ED315423.pdf>
- Zia A, Dar UF. Critical thinking: perception and disposition of students in a medical College of Pakistan. *J Pak Med Assoc*. (2019) 69:968–72.

Author contributions

BA: Conceptualization, Investigation, Methodology, Writing – original draft. SG: Investigation, Methodology, Writing – review & editing. LR: Conceptualization, Supervision, Writing – review & editing, Investigation, Methodology.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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/fmed.2024.1358444/full#supplementary-material>

17. Krupat E, Sprague JM, Wolpaw D, Haidet P, Hatem D, O'Brien B. Thinking critically about critical thinking: ability, disposition or both? *Med Educ.* (2011) 45:625–35. doi: 10.1111/j.1365-2923.2010.03910.x
18. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Searching for and selecting studies In: JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* (2019). 67–107.
19. Li T, Higgins JPT, Deeks JJ. Collecting data In: JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al. (Editors), *Cochrane handbook for systematic reviews of interventions* (2019). 109–41.
20. McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV, Thomas J. Defining the criteria for including studies and how they will be grouped for the synthesis In: JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al. (Editors), *Cochrane handbook for systematic reviews of interventions* (2019). 33–65.
21. Schünemann HJ, Vist GE, Higgins JPT, Santesso N, Deeks JJ, Glasziou P, et al. Interpreting results and drawing conclusions In: JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al. (Editors), *Cochrane handbook for systematic reviews of interventions* (2019). 403–31.
22. Thomas J, Kneale D, McKenzie JE, Brennan SE, Bhaumik S. Determining the scope of the review and the questions it will address In: JPT Higgins, J Thomas, J Chandler, M Cumpston, T Li, MJ Page, et al. (Editors), *Cochrane handbook for systematic reviews of interventions* (2019). 13–31.
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* (2021) 372:n71. doi: 10.1136/bmj.n71
24. Payan-Carreira R, Cruz G, Papathanasiou IV, Fradelos EC, Jiang L. The effectiveness of critical thinking instructional strategies in health professions education: a systematic review. *Stud High Educ.* (2019) 44:829–43. doi: 10.1080/03075079.2019.1586330
25. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv Res.* (2014) 14:579. doi: 10.1186/s12913-014-0579-0
26. Kmet L, Lee R, Cook L. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of FieldsAHFMRHTA Initiative20040213. HTA Initiative, 2. Available at: https://era.library.ualberta.ca/items/48b9b989-c221-4df6-9c35-af782082280e/view/1cfddc-243e-41c3-be98-885f6d4dcb29/standard_quality_assessment_criteria_for_evaluating_primary_research_papers_from_a_variety_of_fields.pdf
27. Ennis RH. "Critical Thinking and Subject Specificity: Clarification and Needed Research." *Educ Res.* (1989) 18:4–10.
28. Chandrasekar H, Gesundheit N, Nevins AB, Pompei P, Bruce J, Merrell SB. Promoting student case creation to enhance instruction of clinical reasoning skills: a pilot feasibility study. *Adv Med Educ Pract.* (2018) 9:249–57. doi: 10.2147/amep.S155481
29. McClintic JA, Snyder CL, Brown KM. Curricular innovation in the surgery clerkship: can assessment methods influence development of critical thinking and clinical skills? [article]. *J Surg Educ.* (2018) 75:1236–44. doi: 10.1016/j.surg.2018.02.012
30. Archila P. Evaluating arguments from a play about ethics in science: a study with medical learners. *Argumentation.* (2018) 32:53–76. doi: 10.1007/s10503-017-9429-7
31. Kim DH. Evaluation of critical thinking course for premedical students using literature and film. *Korean J Med Educ.* (2019) 31:19–28. doi: 10.3946/kjme.2019.115
32. Liao HC, Wang YH. Integrating the gender perspective into literature studies to enhance medical university students' gender awareness and critical thinking [article]. *Int J Environ Res Public Health.* (2020) 17:Article 9245:1–13. doi: 10.3390/ijerph17249245
33. D'Antoni AV, Zipp GP, Olson VG, Cahill TF. Does the mind map learning strategy facilitate information retrieval and critical thinking in medical students? *BMC Med Educ.* (2010) 10:61. doi: 10.1186/1472-6920-10-61
34. Mumtaz S, Latif R. Learning through debate during problem-based learning: an active learning strategy [article]. *Adv Physiol Educ.* (2017) 41:390–4. doi: 10.1152/advan.00157.2016
35. Nguyen K, Khallouq BB, Schuster A, Beevers C, Dil N, Kay D, et al. Developing a tool for observing group critical thinking skills in first-year medical students: a pilot study using physiology-based, high-fidelity patient simulations [article]. *Adv Physiol Educ.* (2017) 41:604–11. doi: 10.1152/ADVAN.00126.2017
36. Banerjee Y, Azar AJ, Tuffnell C, Lansberg PJ, Bayoumi R, Davis D. A novel 6D-approach to radically transform undergraduate medical education: preliminary reflections from MBRU. *BMC Med Educ.* (2018) 18:304. doi: 10.1186/s12909-018-1402-0
37. Ghiam BK, Loftus S, Kamel-ElSayed S. Dialogical narrative approach to enhance critical thinking and student engagement during lecture-based classes. *Focus Health Prof Educ Multidiscipl J.* (2019) 20:30–43. doi: 10.1157/fohpe.v20i3.344
38. Taghinezhad A, Riasati MJ. The interplay of critical thinking explicit instruction, academic writing performance, critical thinking ability, and critical thinking dispositions: an experimental study [article]. *IJERI.* (2020) 13:143–65. doi: 10.46661/ijeri.4594
39. Lee A, Joynt GM, Lee AK, Ho AM, Groves M, Vlantis AC, et al. Using illness scripts to teach clinical reasoning skills to medical students. *Fam Med.* (2010) 42:255–61.
40. Wu B, Wang M, Johnson JM, Grotzer TA. Improving the learning of clinical reasoning through computer-based cognitive representation. *Med Educ Online.* (2014) 19:25940. doi: 10.3402/meo.v19.25940
41. Si J, Kong HH, Lee SH. Developing clinical reasoning skills through argumentation with the concept map method in medical problem-based learning. *Interdiscipl J Prob Learn.* (2019) 13:17763. doi: 10.7771/1541-5015.1776
42. Kumar V, Rajprasath R, Priyadharshini NA, Murugan M, Devi R. Infusing the axioms of clinical reasoning while designing clinical anatomy case vignettes teaching for novice medical students: a randomised cross over study. *Anat Cell Biol.* (2020) 53:151–61. doi: 10.5115/acb.19.199
43. Moghadami M, Amini M, Moghadami M, Dalal B, Charlin B. Teaching clinical reasoning to undergraduate medical students by illness script method: a randomized controlled trial. *BMC Med Educ.* (2021) 21:87. doi: 10.1186/s12909-021-02522-0
44. Middeke A, Anders S, Schuelper M, Raupach T, Schuelper N. Training of clinical reasoning with a serious game versus small-group problem-based learning: a prospective study. *PLoS One.* (2018) 13:e0203851. doi: 10.1371/journal.pone.0203851
45. Brich J, Jost M, Brüstle P, Giesler M, Rijntjes M. Teaching neurology to medical students with a simplified version of team-based learning [article]. *Neurology.* (2017) 89:616–22. doi: 10.1212/WNL.0000000000004211
46. Jost M, Brüstle P, Giesler M, Rijntjes M, Brich J. Effects of additional team-based learning on students' clinical reasoning skills: a pilot study. *BMC Res Notes.* (2017) 10:282. doi: 10.1186/s13104-017-2614-9
47. Klein M, Otto B, Fischer MR, Stark R. Fostering medical students' clinical reasoning by learning from errors in clinical case vignettes: effects and conditions of additional prompting procedures to foster self-explanations. *Adv Health Sci Educ Theory Pract.* (2019) 24:331–51. doi: 10.1007/s10459-018-09870-5
48. Schubach F, Goos M, Fabry G, Vach W, Boeker M. Virtual patients in the acquisition of clinical reasoning skills: does presentation mode matter? A quasi-randomized controlled trial. *BMC Med Educ.* (2017) 17:165. doi: 10.1186/s12909-017-1004-2
49. Isaza-Restrepo A, Gomez MT, Cifuentes G, Arguello A. The virtual patient as a learning tool: a mixed quantitative qualitative study [article]. *BMC Med Educ.* (2018) 18:Article 297. doi: 10.1186/s12909-018-1395-8
50. Mutter MK, Martindale JR, Shah N, Gusic ME, Wolf SJ. Case-based teaching: does the addition of high-Fidelity simulation make a difference in medical Students' clinical reasoning skills? *Med Sci Educ.* (2020) 30:307–13. doi: 10.1007/s40670-019-00904-0
51. Watari T, Tokuda Y, Owada M, Onigata K. The utility of virtual patient simulations for clinical reasoning education. *Int J Environ Res Public Health.* (2020) 17:5325. doi: 10.3390/ijerph17155325
52. Kleinert R, Heiermann N, Plum PS, Wahba R, Chang DH, Maus M, et al. Web-based immersive virtual patient simulators: positive effect on clinical reasoning in medical education. *J Med Internet Res.* (2015) 17:e263. doi: 10.2196/jmir.5035
53. Ludwig S, Schuelper N, Brown J, Anders S, Raupach T. How can we teach medical students to choose wisely? A randomised controlled cross-over study of video- versus text-based case scenarios. *BMC Med.* (2018) 16:107. doi: 10.1186/s12916-018-1090-y
54. Raupach T, Andresen JC, Meyer K, Strobel L, Koziolok M, Jung W, et al. Test-enhanced learning of clinical reasoning: a crossover randomised trial. *Med Educ.* (2016) 50:711–20. doi: 10.1111/medu.13069
55. Montaldo Lorca G, Herskovic LP. Teaching of clinical reasoning to medical students using prototypical clinical cases. *Revista Medica de Chile.* (2013) 141:823–30. doi: 10.4067/S0034-98872013000700001
56. Weidenbusch M, Lenzer B, Sailer M, Strobel C, Kunisch R, Kiesewetter J, et al. Can clinical case discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial. *BMJ Open.* (2019) 9:e025973. doi: 10.1136/bmjopen-2018-025973
57. Bonifacio E, Follansbee WP, Farkas AH, Jeong K, McNeil MA, DiNardo DJ. Implementation of a clinical reasoning curriculum for clerkship-level medical students: a pseudo-randomized and controlled study. *Diagnosi.* (2019) 6:165–72. doi: 10.1515/dx-2018-0063
58. Brich J. Feasibility, acceptance and impact of team-based learning in neurology: a pilot study. *GMS Z Med Ausbild.* (2013) 30:Doc20. doi: 10.3205/zma000863
59. Levin M, Cennimo D, Chen S, Lamba S. Teaching clinical reasoning to medical students: a case-based illness script worksheet approach. *MedEdPORTAL.* (2016) 12:10445. doi: 10.15766/mep.2374-8265.10445
60. Abrami P, Bernard R, Borokhovski E, Wade A, Surkes M, Tamim R, et al. Instructional interventions affecting critical thinking skills and dispositions: a stage 1 Meta-analysis. *Rev Educ Res.* (2008) 78:1102–34. doi: 10.3102/0034654308326084
61. Victor-Chmil J. Critical thinking versus clinical reasoning versus clinical judgment: differential diagnosis. *Nurse Educ.* (2013) 38:34–6. doi: 10.1097/NNE.0b013e318276dfbe
62. Cruz G, Payan Carreira R, Dominguez C. Critical thinking education in the portuguese higher education institutions: a systematic review of educational practices. *Revista Lusófona de Educação.* (2017) 38:43–61. doi: 10.24140/issn.1645-7250.rle38.03
63. Bailin S, Siegel H. (2007). *Critical thinking.* 181–193.

64. Nair GG, Stamler LL. A conceptual framework for developing a critical thinking self-assessment scale. *J Nurs Educ.* (2013) 52:131–8. doi: 10.3928/01484834-20120215-01
65. Hughes C. Theory of knowledge aims, objectives and assessment criteria: an analysis of critical thinking descriptors. *J Res Int Educ.* (2014) 13:30–45. doi: 10.1177/1475240914528084
66. Chi MTH, Wylie R. “The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes.” *Educ Psychol* (2014) 49:219–243.
67. Schullery N, Schullery S. Are heterogeneous or homogeneous groups more beneficial to students? *J Manag Educ.* (2006) 30:542–56. doi: 10.1177/1052562905277305
68. Biggs J. Enhancing teaching through constructive alignment. *High Educ.* (1996) 32:347–64. doi: 10.1007/BF00138871
69. Spencer JA, Jordan RK. Learner centred approaches in medical education. *BMJ.* (1999) 318:1280–3. doi: 10.1136/bmj.318.7193.1280
70. Kassirer JP. Teaching clinical reasoning: case-based and coached. *Acad Med.* (2010) 85:1118–24. doi: 10.1097/acm.0b013e3181d5dd0d
71. Ausubel DP. A subsumption theory of meaningful verbal learning and retention. *J Gen Psychol.* (1962) 66:213–24. doi: 10.1080/00221309.1962.9711837
72. Imwattana K, Dangprapai Y, Ngamskulrungronj P. Active learning classes in a preclinical year may help improving some soft skills of medical students. *Siriraj Med J.* (2020) 72:415–23. doi: 10.33192/SMJ.2020.56
73. Torralba KD, Doo L. Active learning strategies to improve progression from knowledge to action. *Rheum Dis Clin N Am.* (2020) 46:1–19. doi: 10.1016/j.rdc.2019.09.001



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Sharifah Zannierah Syed Marzuki,
Universiti Teknologi MARA Puncak Alam,
Malaysia

Stacy Grau,
Texas Christian University, United States
Carlos Ossa,
University of the Bio Bio, Chile

*CORRESPONDENCE

Nadya Shaznay Patel
✉ nadya.patel@singaporetech.edu.sg

RECEIVED 19 December 2023

ACCEPTED 15 May 2024

PUBLISHED 20 June 2024

CITATION

Patel NS, Puah S and Kok X-FK (2024) Shaping future-ready graduates with mindset shifts: studying the impact of integrating critical and design thinking in design innovation education.

Front. Educ. 9:1358431.

doi: 10.3389/feduc.2024.1358431

COPYRIGHT

© 2024 Patel, Puah and Kok. 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.

Shaping future-ready graduates with mindset shifts: studying the impact of integrating critical and design thinking in design innovation education

Nadya Shaznay Patel*, Shermain Puah and
Xiao-Feng Kenan Kok

Singapore Institute of Technology, Singapore, Singapore

In an era marked by rapid change and complex global challenges, Institutes of Higher Learning (IHLs) are tasked with preparing students to navigate and address these evolving demands. This paper explores the critical role of Higher Education (HE) in equipping students with the necessary skills and mindsets to tackle real-world problems through innovative solutions. Integrating critical thinking and design thinking within a Design Innovation module is central to this exploration. The study is undergirded by a conceptual framework that blends critical, design, and futures thinking, focusing exclusively in this paper on applying critical thinking (CT) and design thinking (DT). The research investigates two primary questions: (1) How do students' DT and CT mindsets change after participation in a Design Innovation module? (2) Is CT a prerequisite for developing DT? This study aims to illuminate the shifts in students' mindsets from before to after the completion of the module, highlighting the importance of developing key dispositions for ethical and socially responsible problem-solving. Results show a statistically significant increase in CT and DT disposition scores from pre- to post-test, suggesting a shift to more positive CT and DT mindsets after going through the Design Innovation module. In addition, a significant moderation effect of pre-test CT mindset on the relationship between pre-test and post-test DT mindset scores was observed, implying that CT was a prerequisite for DT. The findings offer insights into the module's effectiveness in fostering future-ready graduates' thinking capabilities on innovating for real-world challenges and highlight the need for our future-ready students to achieve critical competence and creative confidence. Finally, we conclude the paper with recommendations for educators to integrate CT skill development intentionally and in tandem with DT skill development for a balanced approach to developing critical competence and creative confidence in interdisciplinary courses.

KEYWORDS

critical thinking, design thinking, design innovation, 21st-century skills and dispositions, mindset shifts, interdisciplinary learning

1 Introduction

Higher Education (HE) must transition from the standard 'traditional' discipline-focused curriculum toward one that contributes to sustainable social change by preparing students to be change agents who care for society's well-being. Universities urgently need to prepare graduates and equip young professionals for a myriad of global challenges that lie ahead. From the post-cold war to the coronavirus pandemic, the new world is described as Brittle, Anxious, Nonlinear and Incomprehensible (BANI) (Kraaijenbrink, 2022). To keep abreast with the needs of the future workforce, universities need to rethink their education models and curricula. According to Aoun (2017, p. xviii), "a robot-proof model of higher education is not concerned with topping students off with high-octane facts. Rather, it refits their mental engines, calibrating them with a creative mindset and the mental elasticity to invent, discover or otherwise produce something society deems valuable." To develop those dispositions, scholars have referred to the 4C meta-skills necessary for the success of future professionals: creativity, critical thinking, collaboration, and communication (Pearson, 2018). Bridging the gap between specialized (vertical) and generalised (horizontal) education (e.g., Epstein, 2019), there is a push toward transdisciplinary thinking – an approach to thinking and working that transcends disciplinary boundaries.

Furthermore, developing transferable skills like critical and design thinking is crucial for graduates and the workforce. Critical thinking skills enable individuals to analyze information, evaluate arguments, and make sound judgments, all of which are necessary for success in various fields (Sosu, 2013). Design thinking, on the other hand, is a problem-solving approach that emphasises empathy, creativity, and iterative prototyping (Jakobsone, 2017). This approach is beneficial for addressing complex, multidisciplinary problems that require innovative solutions. Both transferable skills can be applied across disciplines and contexts, enabling students to approach issues from multiple perspectives and develop creative solutions to complex challenges (Suligoj et al., 2020). By integrating critical and design thinking into HE curricula, universities can ensure graduates can innovate in the 21st century.

This paper posits that developing transferable skills, particularly critical thinking and design thinking (CTDT), is essential for preparing students for the challenges of the 21st century. Critical thinking equips students with the ability to analyze, evaluate, and make informed decisions, while design thinking fosters empathy, creativity, and an iterative approach to problem-solving. These skills are crucial for addressing complex, multidisciplinary challenges and fostering innovation across various fields. Our study focuses on a university-wide Design Innovation course underpinned by a conceptual framework that blends critical design, and futures thinking. This course aims to equip students with the necessary skills and knowledge to engage effectively in the design process, from problem definition to ideation and prototyping. The research questions guiding this study are: (1) How do students' DT and CT mindsets change after participation in a Design Innovation module? (2) Is critical thinking a prerequisite for developing design thinking? By investigating these questions, the study seeks to understand the impact of the Design Innovation module on students' mindset shifts and their preparedness to innovate and solve real-world problems. The findings will contribute to the discourse on the importance of transdisciplinary knowledge and skills in HE, particularly in

Singapore. They will provide valuable insights for educators and policymakers on effectively preparing future-ready graduates.

2 A review of literature

2.1 Trends of developing transferable skills and their challenges

Today, IHLs face a critical challenge: preparing graduates for success in a workforce that demands a broad set of transferable skills. Among these skills, critical thinking and design thinking have emerged as essential for navigating the complexities of the modern workplace. However, developing these skills in HE presents several challenges.

One challenge is the need for faculty and instructors to understand and be able to teach these skills effectively. Critical thinking and design thinking require a different approach to teaching and learning than traditional lecture-based instruction (Turnali, 2016). For example, critical thinking involves analysing and evaluating information, synthesising ideas from multiple sources, and making informed judgments. This requires a shift away from simply providing students with the knowledge and toward creating opportunities to practice critical thinking skills in various contexts (e.g., Patel, 2021). Similarly, design thinking involves a problem-solving approach emphasising empathy, creativity, and iterative prototyping. This requires shifting from linear problem-solving methods to a more iterative and flexible approach (Mosely et al., 2018). Thus, much evidence-based research has called for universities to spearhead intentional faculty development programmes and training workshops to demystify the development of these thinking skills in students (e.g., Patel and Wong, 2022). These help instructors develop the knowledge and skills necessary to teach critical thinking and design thinking effectively as transferable skills across degree specializations. For example, practical strategies using dialogic scaffolding to model these thinking skills so that instructors can make them explicit and visible to students (Patel, 2022).

Another challenge is the need for appropriate assessment measures to evaluate these skills. Traditional assessment methods, such as multiple-choice exams, may only partially capture the application of critical thinking and design thinking. Therefore, instructors must develop performance-based assessments and portfolio evaluations or adopt a students-as-partners approach as an appropriate assessment method aligned with these skills' learning outcomes. A further challenge is the need for students to understand the value of these skills and be motivated to develop them. Students may not see the immediate relevance of critical thinking and design thinking skills to their future careers or may need to be made aware of how to develop these skills (Ramírez-Montoya et al., 2022). Therefore, IHLs need to rethink and redesign their curricula so that these key transferable skills are embedded intentionally across all specializations. This gives students ample opportunities to develop and internalize them before applying them in their future work.

2.2 Future skills and mindsets

The concept of 21st-century skills emerged as a response to the changing nature of the world and the workforce. It emphasises the

need for graduates to possess a broad set of transferable skills, like critical thinking, creative thinking, communication, and collaboration, to succeed. However, as we move into the 21st century, the landscape continues to evolve, and new skills will likely emerge as essential for success (Dilekçi and Karatay, 2023). The Organization for Economic Co-operation and Development (OECD) launched a project to identify the skills individuals need in the 21st century (Pedro, 2006). The project aimed to define the characteristics of the century and determine how individuals could acquire these skills. Pedro (2006), who participated in the project, grouped the skills that should be incorporated into the educational system into three categories: alternative cognitive skills, changes in cultural practices and social values, and expectations regarding teaching and learning. Another researcher, Wagner (2008), studied the topic and proposed a framework under the concept of “survival skills,” which included seven skills: problem-solving and critical thinking, cooperation and leadership, adaptation and agile intelligence, entrepreneurship, effective communication, access to information and analysis of information, and imagination and curiosity. The World Economic Forum, UNESCO, the European Commission, and the OECD have recently focused on researching 21st Century or Future Skills. These skills are the attributes that graduates need to thrive in an increasingly globalized and digitized world in a socially creative, responsible, sustainable way and by the Millennium and Sustainable Development Goals. Despite years of discussion and research, effectively embedding and integrating these skills still needs to be improved (Osmani et al., 2019).

The discourse on skills for the future often emphasises technical competencies and cognitive abilities. However, key dispositions, attributes, or mindsets are equally vital, particularly for critical and design thinkers. These mindsets or dispositions are crucial for navigating the complexities of the 21st century. A critical thinker in this era must embody a disposition of open-minded scepticism – a balance between being open to new ideas and rigorously questioning and evaluating information (Paul and Elder, 2012). This mindset fosters the ability to navigate through misinformation and rapidly evolving knowledge landscapes. For design thinkers, a mindset of empathetic engagement is paramount. This involves profoundly understanding and connecting with the needs and experiences of others, which is essential for human-centred design (Brown, 2009). Empathy enables design thinkers to create solutions that are not only innovative but also resonate deeply with the end-users needs and contexts. Another key disposition is adaptability, flexibility, and resilience in the face of change. This is crucial as the future is marked by rapid technological advancements and shifting socio-economic landscapes (Dweck, 2006). Adaptability allows individuals to thrive in diverse environments and embrace continuous learning.

The literature review on the Design Thinking (DT) Mindset in Dosi et al.'s (2018) study focuses on identifying and categorising the elements that constitute this mindset. The authors conducted a structured research analysis, drawing from the scientific literature on Google Scholar, and applied specific criteria to select relevant publications. This process led to identifying 19 latent constructs of the DT mindset, derived from a comprehensive analysis of 17 selected papers (Table 1). These constructs encompass a range of attitudes and cognitive approaches essential for effective design thinking, such as

tolerance for ambiguity, embracing risk, human-centeredness, empathy, mindfulness, holistic view, problem reframing, teamwork, multidisciplinary collaboration, openness to different perspectives, learning orientation, experimentation, experiential intelligence, critical questioning, abductive thinking, envisioning new things, creative confidence, desire to make a difference, and optimism (Schweitzer et al., 2016; Dosi et al., 2018).

This study aims to capture a comprehensive and multifaceted understanding of the DT mindset by selecting representative items from each construct. This approach is grounded in the belief that a holistic representation of the DT mindset is crucial for accurately assessing and understanding student mindset shifts. The authors believe that the rigorous literature review and categorization process by Dosi et al. (2018) ensures that the identified constructs and their representative items are grounded in academic research and practice. For instance, the importance of constructs like empathy and human-centeredness in design thinking is well-documented in the works of authors like Tim Brown (Brown, 2009) and Jeanne Liedtka (Liedtka et al., 2017).

Similarly, the emphasis on critical questioning and abductive thinking aligns with the principles outlined in design thinking literature by scholars such as Nigel Cross (Cross, 2011) and Richard Buchanan (Buchanan, 1992, 2015). The selection of items from each construct is not arbitrary but is informed by the frequency and significance of these elements in the analyzed literature. This methodological approach ensures that the selected items represent the broader concepts and themes within the DT mindset. The empirical backing for this approach is supported by the scholars shared earlier, whose research described their analysis and selection process based on the prevalence and importance of these elements in the design thinking discourse. This forms a deliberate strategy to ensure a comprehensive and nuanced understanding of the design thinking mindset. This approach also reflects the depth and breadth of the design thinking field as documented in academic literature.

TABLE 1 Dosi et al.'s (2018) 19 constructs of design thinking mindset.

A. Tolerance for - being comfortable with ambiguity - uncertainty	J. Open to different perspectives / diversity
B. Embracing risk	K. Learning oriented
C. Human centeredness	L. Experimentation or learn from mistake or from failure
D. Empathy / empathic	M. Experiential intelligence / Bias toward action
E. Mindfulness and awareness of process	N. Critical questioning (“beginners mind,” curiosity)
F. Holistic view / consider the problem as a whole	O. Abductive thinking
G. Problem reframing	P. Envisioning new things
H. Team Working	Q. Creative confidence
I. Multi-/ inter-/ cross-disciplinary collaborative teams	R. Desire to make a difference
	S. Optimism to have an impact

2.3 Design thinking and design innovation

Design Innovation is a multidisciplinary approach that integrates design and innovation management principles to address complex challenges and deliver user-centric solutions. Scholars argue that the interplay between innovation and design processes comes from a perspective in which value is defined by and co-created with the user rather than embedded in the output (Concilio et al., 2019). It emphasises applying design principles not only to product aesthetics but also to enhance the functionality and accessibility of solutions, thereby meeting and often anticipating user needs (Verganti, 2009). This approach includes various methodologies, such as user research, prototyping, and usability testing, which are critical for developing successful innovations.

Design Thinking (DT), a core component of design innovation, is an iterative process that seeks to understand users, challenge assumptions, redefine problems, and create innovative solutions to prototype and test. This methodology effectively tackles complex problems that are ill-defined or unknown by re-framing these in human-centric ways, encouraging multiple iterations, and fostering a learning mindset by doing (Brown, 2009). Dosi et al.'s (2018) study on DT Mindset offers valuable insights into the cognitive processes that underlie effective design thinking. The 19 identified latent constructs of the design thinking mindset (See Table 1 earlier), which include empathy, tolerance for ambiguity, risk-taking, openness to experience, and collaborative spirit, are crucial for fostering an environment conducive to innovation and for empowering individuals and teams to approach complex problems creatively (Dosi et al., 2018).

Moreover, embracing a DT mindset involves cultivating certain attitudinal qualities such as optimism and the willingness to fail. These attributes enable practitioners to persist in the iterative cycles of prototyping, testing, and refining ideas. Thus, this study emphasises the need for educational programmes in higher education to integrate these mindset elements to better prepare students for real-world challenges.

2.4 Critical thinking definitions, dispositions and application in design innovation

Facione (1990, p. 2) describes Critical Thinking (CT) as a deliberate and self-regulating judgment encompassing analysis, evaluation, and inference, grounded on various evidential and contextual factors. This comprehensive definition underscores the complexities of educating CT, reflecting teachers' challenges in fostering CT skills (Tsui, 2002). Research consistently recommends teaching methods that emphasise "how to think" rather than "what to think" to effectively cultivate CT skills (Cloete, 2019; Puig et al., 2019). The need for enhanced CT instruction is recognized in higher education, especially as students often struggle to grasp and apply CT in their work, and educators frequently misunderstand how to teach these skills effectively (Abrami et al., 2008, 2015; Janssen et al., 2019). Duro et al. (2013) highlight this issue, noting a discrepancy between student and faculty perceptions of CT, and suggest structured interactive activities to improve students' critical and metacognitive abilities.

Critical thinking (CT) is characterized by a duality of skills and dispositions fundamental for academic and professional success. CT

skills involve the ability to analyze, evaluate, and synthesize information, whereas CT dispositions involve the inclination to engage in those cognitive activities (Facione et al., 1992). Recent literature underscores that while CT skills are about performing cognitive tasks, dispositions relate to the willingness to apply such skills in varying contexts, including the readiness to think critically which is pivotal in effective decision-making and problem-solving (Dwyer et al., 2012; Abrami et al., 2015). Simonovic et al. (2022) differentiate between critical thinking skills and dispositions in a research studying the effect of online students' perceptions and attitudes toward CT. It points out that while CT skills relate to the ability to carry out critical thinking functions (such as analysis and evaluation), dispositions involve the willingness to engage in and the attitudes toward such cognitive activities. It stresses that skills and dispositions are important for effective CT, with cognitive reflection and student attitudes and beliefs playing significant roles in their academic performance.

In exploring the relationship between CT readiness and DT tasks, studies have demonstrated that engaging in DT activities enhances students' ability to think critically by fostering an environment that requires iterative learning, problem redefinition, and solution-focused thinking (Johansson-Sköldberg et al., 2013). These design tasks promote a mindset that is open to exploration and encourages a reflective and iterative approach to problem-solving (Carroll et al., 2010). Additionally, research by Wrigley and Straker (2017) suggests that DT activities can significantly impact students' CT dispositions by placing them in real-world problem-solving scenarios that necessitate active engagement and critical appraisal of information. This interaction between CT dispositions and skills through DT tasks has led to improved cognitive processes and outcomes, particularly in complex and ambiguous situations (Koh et al., 2015).

Paul and Elder (2012) define it as the *art* of thinking to improve it. It is essential in design innovation as it provides the analytical backbone for convergent and divergent thinking processes. Convergent thinking involves narrowing down the options discovered in divergent thinking to identify the best solution. It is through this synthesis of creativity and systematic problem-solving that innovative solutions are born (Runco, 2014). CT aids designers in various stages of the design process. For example, during the empathy phase of DT, CT is used to question assumptions about user needs and experiences. This rigorous questioning ensures that solutions are developed based on deep, insightful understandings rather than superficial observations.

CT also plays a crucial role in the testing and implementation phases of design innovation. It helps teams evaluate the effectiveness of their prototypes and refine them based on feedback and critical analysis. This iterative process, underpinned by strong critical thinking, leads to the development of innovative solutions that are viable and sustainable in the long term (Ennis, 2018). Integrating critical thinking into design innovation practices and curricula can significantly enhance students' abilities to engage with complex problems creatively and effectively. It cultivates a mindset that values rigorous evaluation, ethical consideration, and the sustainability of design solutions, thereby fostering a more holistic approach to design innovation (Brookfield, 2012).

2.5 Blending critical and design thinking

Designers have utilized Design Thinking (DT) for decades. However, the truth of DT is that it could need more depth. Although

in the initial stage of employing empathy, where the designers get a feel of the users' problems, the process runs the risk of being solely based on the designers' understanding of the issue. Scholars highlight current applications of DT as lacking criticality (e.g., [Turnali, 2016](#)). According to [Loewe \(2019\)](#), DT continues to fail at tackling the most 'wicked' kinds of problems, which are defined as complex problems that lead to unwanted outcomes or even unsolvable problems that include social, technological, economic, and political ones. This is partly because of DT's superficial understanding of the 'wicked' factors to effectively address the multitude of often intertwined contextual factors of the problem ([Loewe, 2019](#)). Thus, while DT itself may be able to explore the issue at hand, solutions might still fall short without Critical Thinking (CT). In fact, without unbiased, critical thought, DT is ineffectual. It becomes a design process where all ideas are good, and best intentions overshadow best practices ([Bezhnar, 2019](#)). Therefore, CT should be applied to each step of the design process to create the solution carefully without losing the sole purpose and vision (see [Figure 1](#)).

While some scholars have posited that design thinking may support critical thinking, empirical research examining the relationship between these two modes of thinking is lacking because their shared conceptual structure has not been articulated in detail, and they have remained siloed in practice. A recent study ([Ericson, 2021](#)), reporting on a detailed comparison of the two modes of thinking, suggests that design thinking methods can support and augment traditional critical thinking practices and that design thinking frameworks could be modified to incorporate critical thinking more explicitly. Moreover, scholars have posited that products are often designed to take advantage of human psychology in ways that do not always align with end users' interests ([Nodder, 2013](#)).

Blending design and critical thinking work with the premise that design thinking can offer educators creative new approaches to engage students in critical thinking. Second, critical thinking must be more explicitly integrated into product design and development methods. Hence, this research supports other evidence-based studies (e.g., [Ericson, 2021](#)) first to clarify the relationship between these two modes of thinking to help educators leverage design thinking methods to support and augment various approaches to critical thinking and then help prepare students for the workplace where they can explicitly integrate essential practices of thinking into their design processes. Furthermore, as design thinking is increasingly used to design the products of the modern world, the environments people inhabit, and even the social interactions that people engage in, it is indeed crucial for educators to understand the relationship between critical thinking and design thinking ([Ericson, 2021](#)). Also, critical thinking must be more explicitly integrated into the design thinking process to ensure that design efforts maximize potential benefits and minimize potential harm to society.

2.6 Creative thinking and critical design approach to innovation

The new world facing us demands skilful and creative thinkers to develop new ways of living and solutions to the world's most significant problems. According to [Torrance \(1966\)](#), creative thinking involves generating ideas that demonstrate fluency, flexibility, and

originality in response to a problem, event, or situation. Fluency refers to producing many ideas, flexibility entails considering different perspectives, and novelty involves presenting new ideas. Competence in these areas is used to evaluate creativity ([Mourtzis et al., 2018](#)).

Contrary to the belief that creativity is a rare talent, creative thinking is a multi-faceted cognitive process in everyone at different levels. Creative thinking and critical thinking are often presented as two distinct approaches to problem-solving. Creative thinking involves expansive and innovative thinking that is free of constraints and is focused on generating new ideas and exploring possibilities ([Treffinger and Isaksen, 2005](#)). In contrast, critical thinking is a more disciplined and logical approach focusing on practical and realistic solutions to problems ([Brookfield, 2012](#)). The key difference between creative and critical thinking is that generating alternatives is a creative activity while selecting among them requires a critical approach. Although creativity and criticalness are often perceived as opposite ends of a spectrum, they are not mutually exclusive and can often complement one another in problem-solving processes ([Nickerson, 1999](#), p. 397). In fact, in this paper, intentionally blending critical and creative (design) thinking could develop better thinkers among students as they achieve creative confidence and critical competence when ideating solutions to solve real-world problems. The premise for this is based on the evolving debate on the practice of Critical Design ([Dunne and Raby, 2001](#); [Malpass, 2013](#)) and critical design thinking ([Loewe, 2019](#)).

Critical Design is a design practice that uses provocative designs to stimulate critical thinking and open up discussions about themes and issues raised by the design work ([Loewe, 2019](#)). It differs from developing necessary design thinking skills, as critical design is an affective practice that generates questions rather than providing solutions or answers to design problems ([Malpass, 2015, 2017](#)). The primary goal of critical design is to provoke debate and engage its audience rather than to simplify or explain complex design issues. Critical design seeks to diversify how people understand design problems and ideas by appealing to their emotions and senses and opening up specific lines of inquiry that lead to a more informed and critical understanding of a given problem ([Malpass, 2017](#)). Using this approach, critical design challenges traditional design processes and opens up new ways of thinking about design and its impact on society. [Loewe \(2019\)](#) highlighted the key questions in this discussion: How can critical design, which appeals to emotions, still be critical without rational analysis? How can it transform affect into an analytical assessment? How can critical design achieve more than an accidental emotional connection based on people's tastes, fixed opinions, or moral convictions? *In our research, we asked how HE educators, informed by the literature on critical thinking, design thinking and critical design approach to innovation, can design an instructional approach to develop critical design thinking skills in students so that they will ideate innovative and sustainable solutions for the world's most complex problems. In particular, how would such an instructional approach influence their mindset shifts?*

Critical design is an affective design practice that aims to provoke debate and engage its audience by using provocative designs to stimulate critical thinking ([Malpass, 2017](#)). However, critical thinking in this notion is predominantly based on the alleged criticality of the object itself, and the interdependency of deep reflection and critical thinking remains an aporia ([Loewe, 2019](#)). While critical design may stimulate debates that lead to broader

Stages	Design Thinking	Critical Thinking
Empathise	Empathy is used to fully understand the context of the users experiencing the problem. Divergent thinking is used at this stage to explore data.	Proper time and resources must be given to research, collecting all relevant data, and applying critical thinking to fully understand those who are in need of a solution. If designers cannot relate to the customer's state of mind, empathy is impossible.
Define	From the data collected earlier, designers will distil insights with convergent thinking to formulate a problem statement.	Critical thinking is used to define the problem from the users' point of view, rather than designers'. Ensure relevance and benefit to the user first.
Ideate	Ideation, where multiple solutions are being ranked, is crucial in the design process. The evaluation and elimination processes, using both divergent and convergent thinking, will lead to an optimal solution selected.	Critical thinking is vital to ensure out-of-the-box ideation that goes beyond creative ways to solve a problem. While it is intended to be inspirational and collaborative, it is easy to become distracted in the process and lose sight of the empathic and well-defined vision.
Prototype	Designing an early version of the solution to reveal how users think and feel. Creating multiple prototypes, low or high fidelity, takes the risk out of innovation by allowing designers to fail quickly.	By including all stakeholders, designers will achieve a more efficient and robust experimental process. With critical thinking, any acceptance, improvements, and even rejections will be driven by user experiences to determine solution constraints and challenges.
Test	As the last stage of the process, tests tend to perform together with the prototyping stage. Through testing, designers can learn more about users, improve the prototype and even refine the problem statement.	The testing stage redefines, informs, and drives changes and refinements. Here, critical thinking—based on perpetual learning—is applied to understand how end-users think, behave, and feel in order to (re)empathize and optimize the experience.

FIGURE 1

Critical thinking for each of the 5-stage design thinking process (Bezhnar, 2019).

critical knowledge, more is needed to gain deep insights and understanding of complex problems. To do this, HE educators must adopt an instructional approach to develop critical design thinking skills in students. This approach should focus on ideating innovative and sustainable solutions to complex problems by emphasising the importance of deep reflection, broad inquiry, and critical analysis. By doing so, students can become more aware of the complex issues and problems that contextualise the perception and utility of design, leading to more informed and practical solutions to real-world problems.

Criticism of design thinking is common among design scholars, who seek to ensure critical rigor as they scrutinize approaches to design decision-making. Rodgers (2017) advocates for more critical design thinking to provide a means of addressing the complexities of the world. Kimbell (2011) criticises the lack of criticality and the focus on the designer in current design thinking applications. However, few offer practical guidance on implementing a critically informed design thinking approach. Loewe (2019) notes that design thinking cannot substitute for deep critical reflection on the complexity of a design challenge. This reveals a common misconception among those who use design thinking. Critical design thinking should be introduced to

guide decision-making, merging critical design practices with design thinking to address 21st-century problems.

However, the biggest challenge for HE to produce better thinkers is, as Halpern (2010) recognized, that “the enhancement of critical and creative thinking is still more of a desirable vision than an empirical outcome” (p. 381). Design thinking requires curiosity, imagination, and creativity to generate, explore, and develop possible solutions, and it might also depend on the skill level of critical thinking (Mosely et al., 2018). As a creative approach, design thinking has become an evolving field in HE, connecting students of various disciplines to solve complex problems as a team (Wrigley and Straker, 2017). Furthermore, IHLs have increasingly incorporated design thinking into the undergraduate curriculum, exposing non-design students to design thinking skills (Avsec and Ferk Savec, 2019).

Therefore, integrating Design Thinking, Critical Thinking, and Critical Design forms a holistic approach that enriches innovation. This integration enhances the ability to address complex challenges by ensuring that solutions are innovative, user-centred, critically evaluated, and socially relevant. The convergence of these methodologies fosters a comprehensive problem-solving framework that is adaptive, reflective, and

inclusive. This approach supports the creation of solutions that are deeply reflective of user needs and societal contexts, thus enhancing the overall effectiveness and sustainability of innovations. The blend of these perspectives encourages a culture of continuous learning and adaptation, which is crucial in today's rapidly changing world (Kolko, 2015).

3 Context: design innovation at the Singapore Institute of Technology

In Singapore, many IHLs have embarked on an ambitious endeavour to foster an interdisciplinary curriculum with identified transferable skills like critical thinking, creative thinking and design thinking. At our university, we developed a Design Innovation university-wide module. For our university, an applied learning autonomous university, up to 3,000 students, regardless of their programmes, would read the module every academic year.

3.1 UDE1001: Introduction to design innovation

This module aims to introduce design innovation and the application of user-centric innovation. Students will learn about design thinking as an ideology and approach that seeks to solve complex real-world problems in any specialized discipline area. You will learn to apply critical thinking tools and a human-centred mindset to understand the key challenges of an identified authentic problem. Throughout the module, students will use several iterations of Critical-Design Thinking (CTDT) exercises to critically question the problem and gain insights. Additionally, students will learn about synthesising preliminary ideas and solutions to the identified problem. Students can look forward to practising CTDT skills in many engaging exercises and applying them later in the four-credit Interdisciplinary Design Innovation module (UDE2001). Students will continue working in teams to ideate further and prototype solutions.

We endeavour to drive creative innovation with the university-wide Design Innovation modules (UDE1001 and UDE2001). Undergirded by a Critical Design Futures Thinking (CDF™) conceptual model, students collaboratively explore real-world sustainability issues and propose creative solutions. In the next section, we will share findings from a research study, where students' needs are at the centre of the enquiry, to propose a conceptual framework of critical design thinking and the tools we develop to develop students' creative confidence and critical competence. Focusing on developing key dispositions for ethical and socially responsible students, we recommend playful learning as the adopted signature pedagogy and using prototypes as critical debates.

3.2 Adopting a critical design futures thinking conceptual framework

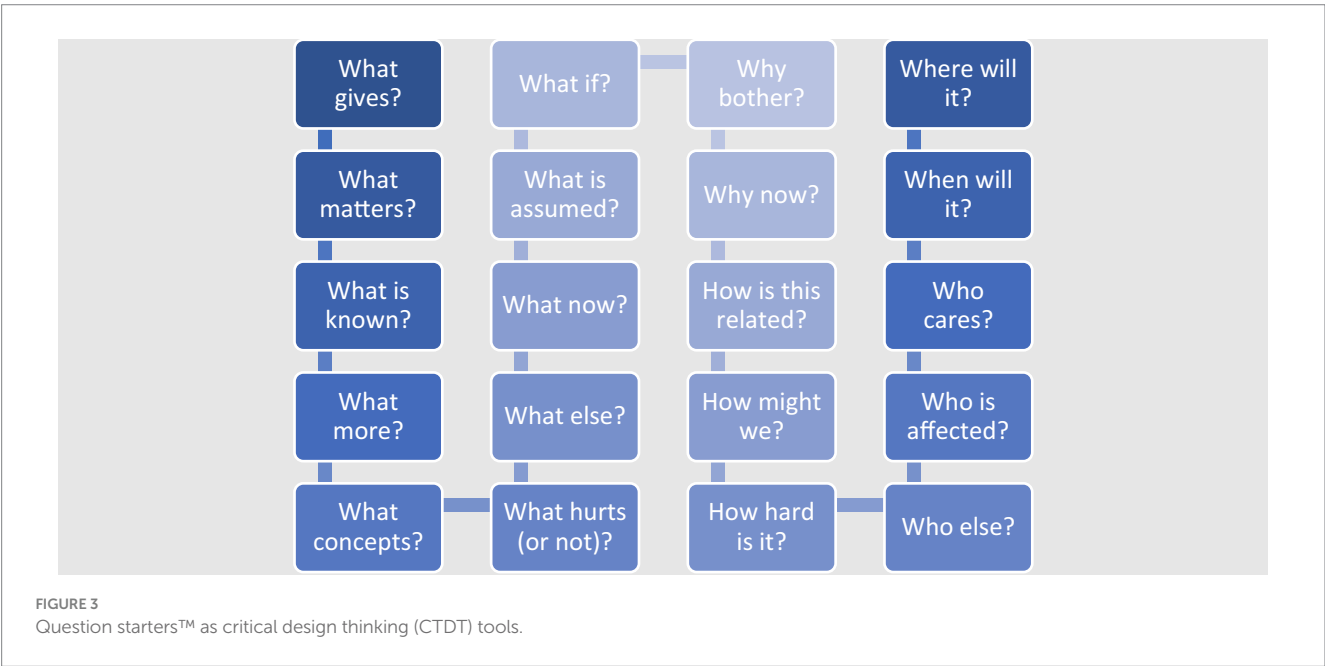
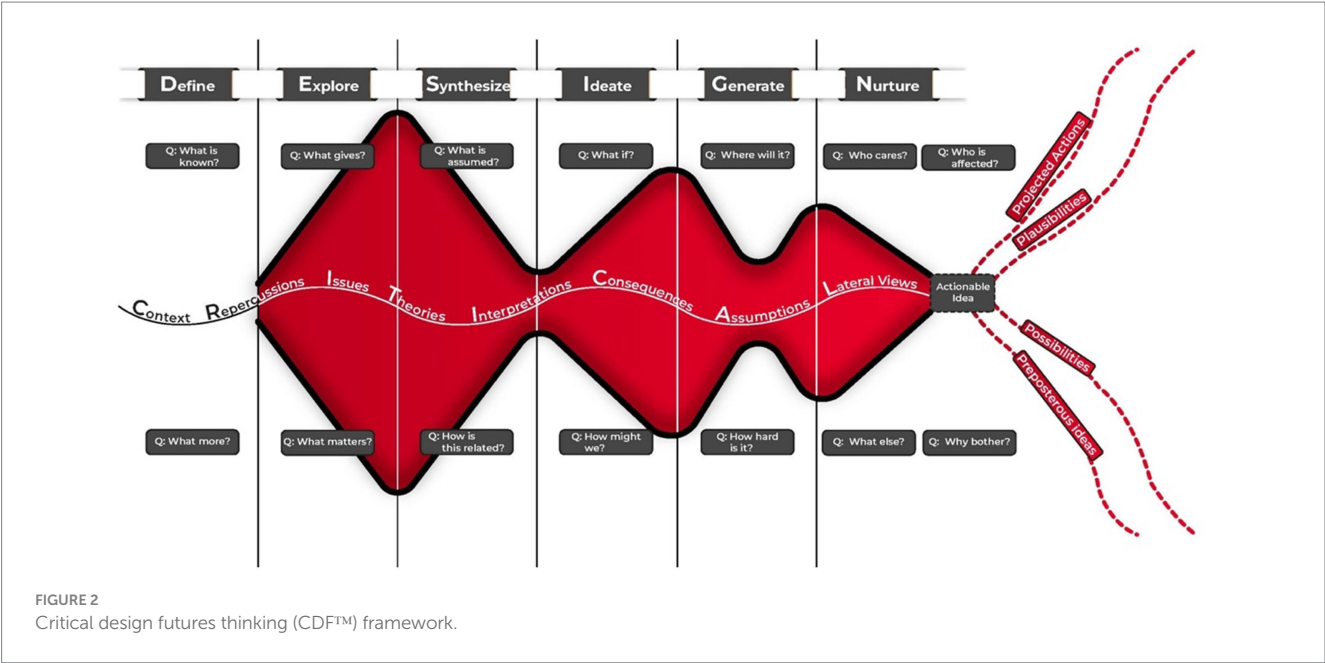
The wicked problems we increasingly face require a radically different form of thinking that rigorously questions ideas and

assumptions. It demands skilful thinkers to develop new ways of living and solutions. The intentional integration of critical design thinking in our adopted framework allows for an explicit application of a structure. This effectively guides convergent and divergent thinking in the design innovation approach. Thus demystifying the process of creating innovative solutions. When good questions are asked, we reframe how problems are viewed and solved. It is with critical questions that we get to evaluate and achieve the best of our thinking.

In our university, we proposed a conceptual framework to blend critical thinking, design thinking and futures thinking (CDF™) to undergird the design of the Design Innovation module (Figure 2). At the point of this study, we had only piloted the application of critical and design thinking into the curriculum. Thus, the mindset shifts surveys for critical and design thinking were administered. The intentional integration of Critical Design Futures (CDF™) thinking in our framework allows for an explicit application of a structure to help empower students today to innovate for a better world. This effectively guides their convergent and divergent thinking in the design innovation approach. Thus demystifying the process of creating innovative solutions. It is with critical questions that we get to evaluate and achieve the best of our thinking. When good questions are asked, we reframe how problems are viewed and solved.

Undergirded by the model, students are guided by their instructors. They are introduced to "Question Starters" (Figure 3) to ask good questions with a CDF™ mindset to explore multifaceted real-world issues in a humanity-centred approach. This highlights the need for our future-ready students to achieve creative confidence and critical competence when ideating solutions for a sustainable future. We endeavour to design a curriculum incorporating transdisciplinary approaches to ensure students have the necessary skills like curiosity, confidence, creativity, criticality and collaboration to succeed. We posit that undergirded by the CDF™ conceptual framework, using Question Starters will allow students to internalize the skills of critical thinking, design thinking, and futures thinking. This will help guide the iteration of divergent and convergent thinking to ideate for innovative solutions.

In achieving our aims to develop future-ready graduates with the necessary competencies to innovate as they tackle complex real-world problems with transdisciplinary collaborations, we reflected on our curriculum design for the design innovation module. We wanted to ensure that the development of transferable skills of critical design thinking was explicit and visible for students and instructors in the curricular and instructional strategies. Hence, with the CDF™ conceptual framework, we develop the first iteration of a series of *question starters* for students to internalize the critical-design thinking tools, from conceptualising their projects at the discovery stage to testing their prototypes. (Figure 3) These question starters correspond to Elements of Thought and Intellectual Standards from the Paul and Elder (2019) critical thinking framework and the five stages of the Design Thinking process. Instructors first modeled the questions when engaged in discussions and dialogs with students before encouraging them to use them in group discussions and peer reviews.



Integrating Question Starters within the design innovation modules was pivotal in fostering a learning environment of deep, reflective, and innovative thinking among students. This approach aligns with the ethos of CDF™ by encouraging critical engagement with present complexities and shaping proactive thinkers and innovators capable of navigating future uncertainties. Students' design innovation projects explicitly used these Question Starters when conceptualising, framing the problem to be addressed, ideating and evaluating/testing their prototypes. Industry leaders who served as guest judges for students' final presentations gave encouraging feedback on the quality of the students' work. The Question Starters such as "How else?," "What now?," "Why bother?," "What gives?" and "What drives?" were instrumental in making the cognitive processes explicit and visible to students. They serve as cognitive prompts, encouraging students to delve deeper into their thought processes, exploring various dimensions of a problem or concept. By posing such questions, students can articulate their reasoning, assumptions, and conceptual understandings, which might remain implicit or unexamined. This was evident in the quality of their proposed design innovation projects, based on challenge briefs relating to UN Sustainable Developmental Goals. This explicit articulation of thought promotes metacognitive awareness and helps internalize complex thinking skills.

3.3 Sample students' design innovation projects

3.3.1 Project 1: Eat what

Using games to resolve youths' indecisiveness when going out for a meal with peers while at the same time raising awareness on the issue of food security globally (Figure 4).

3.3.2 Project 2: Parenthoodease

Using mobile technology to help low-income first-time parents navigate the challenges of parenthood, particularly on the increased financial responsibility of raising their child/children and receiving emotional support from peers (Figure 5).

4 Methods

4.1 Participants

A longitudinal study was conducted with two cohorts of first-year undergraduate students at a Singapore-based university over two consecutive trimesters. Data for each cohort was collected across two-time points: students completed the survey at the beginning of the trimester (pre-test) and again at the end of the trimester

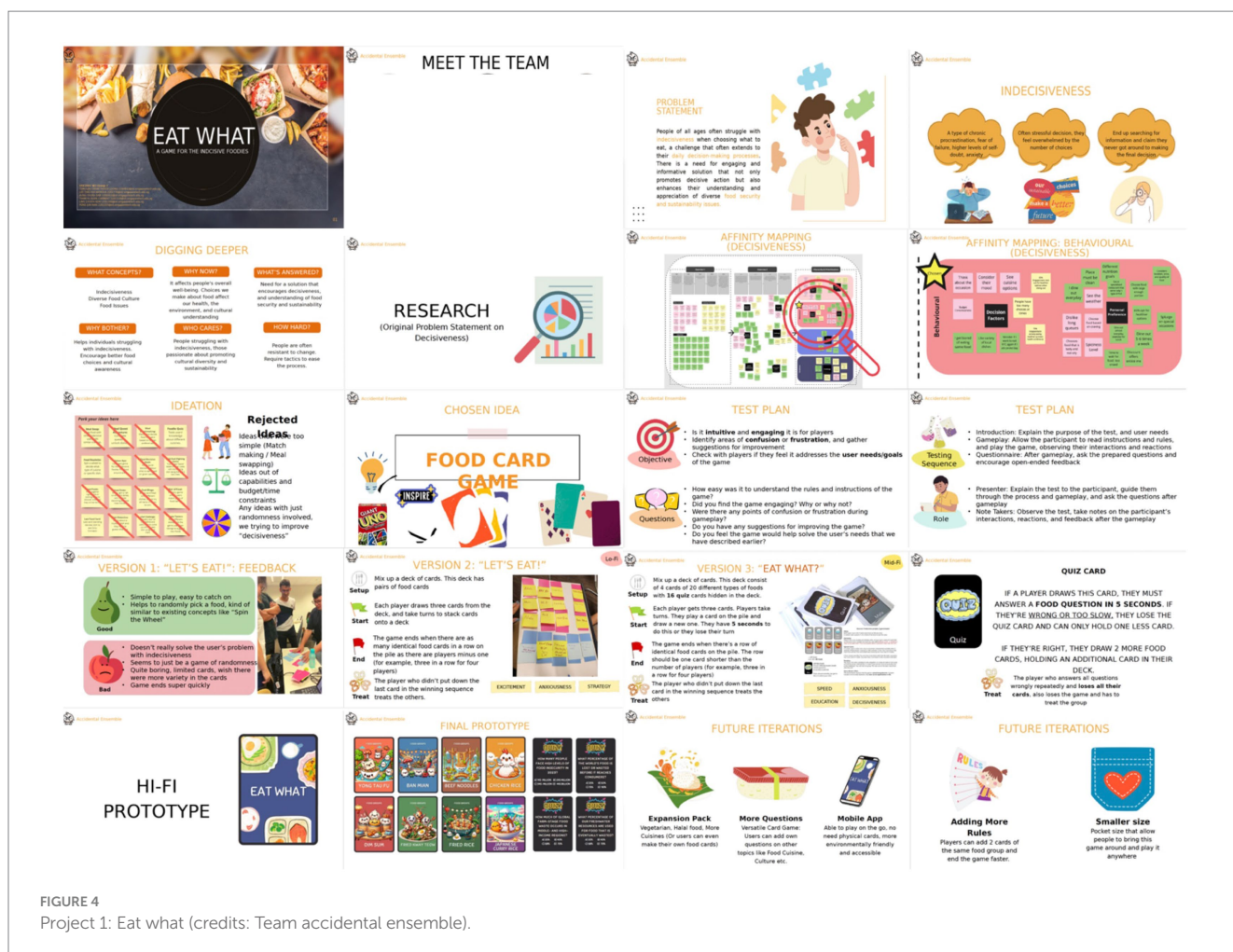
(post-test). The final participant sample matched for completing the pre-test and post-test were 373 students in Cohort 1 and 625 students in Cohort 2. Students were enrolled in various degree programmes (e.g., Accountancy, Diagnostic Radiography, Air Transport Management, Nursing, Electrical Power Engineering).

4.2 Measures and procedures

An online survey was administered via Qualtrics to all students of the "Introduction to Design Innovation" elective module. Ethical approval for data collection was sought and obtained from the Institutional Review Board at the authors' institution (IRB: 2021154). Students were encouraged to complete the pre-test survey in the first week of the study and the post-test survey in the final study week of the trimester. After responding to demographic questions, students completed two survey instruments: a (1) 24-item Critical Thinking Disposition Scale and (2) 21-item Design Thinking Mindset tool.

4.3 Critical thinking disposition

The Critical Thinking Disposition Scale (CTDS; [Sosu, 2013](#)) was used to measure characteristics that illustrate a disposition toward



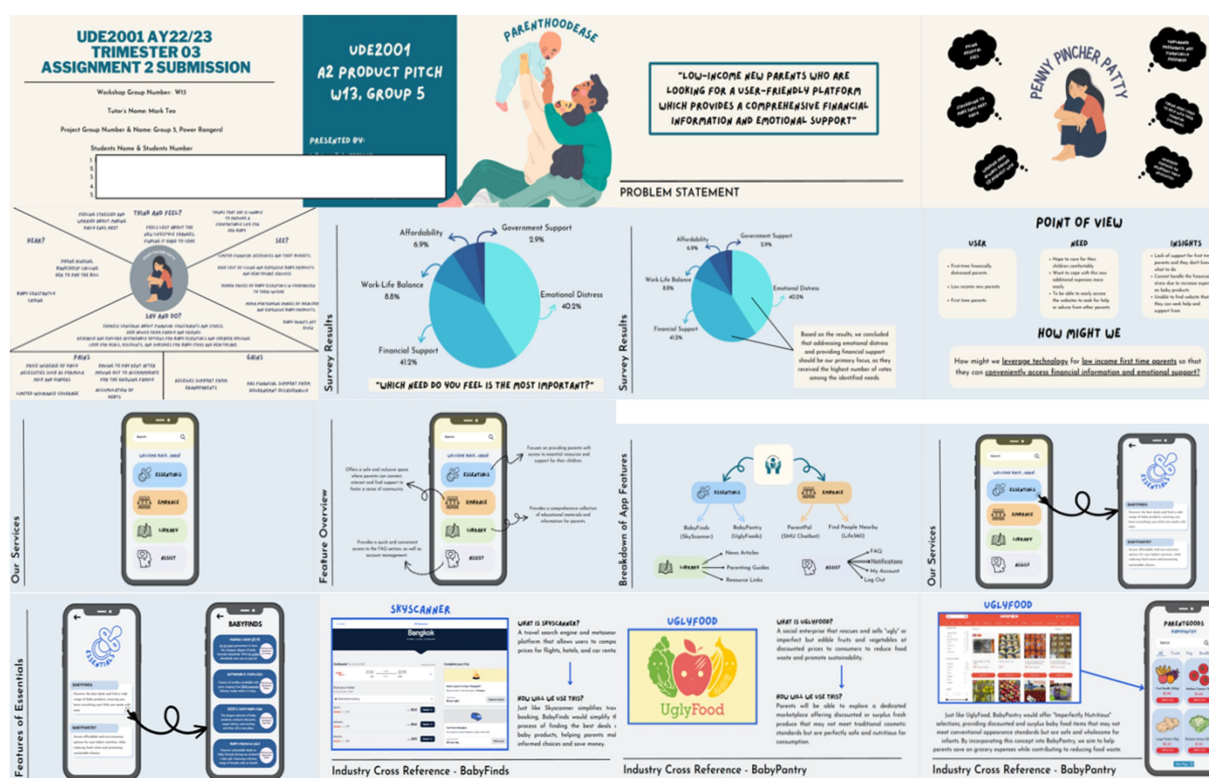


FIGURE 5
Project 2: Parenthoodease (credits: Team W13Grp5).

critical thinking (e.g., “I usually check the credibility of the source of information before making judgments”). The 24 items measuring students’ critical thinking disposition were rated on a five-point Likert-type rating scale (1 = *Strongly Disagree*, 5 = *Strongly Agree*). The internal reliability of the 24 items in this study across cohorts and time points ranged from 0.93–0.97, as assessed using Cronbach’s alpha.

4.4 Design thinking mindset

The Design Thinking Mindset (DT Mindset; Dosi et al., 2018) tool was used to measure students’ metacognitive self-awareness of DT mindset attributes (e.g., “I am comfortable transforming ideas into something tangible”). Of the 71 items clustered in 19 DT mindset constructs of the reported tool, we carefully selected 21 items with at least one item from each construct (Refer to Appendix 1). To ensure that the selected items are representative of their respective latent constructs, we examined their content validity by (a) referring to the conceptual definitions of each element of the DT mindset, (b) identifying the characteristics that define each element of the DT mindset, and (c) seeking inputs from experts in the DT mindset research field. The internal reliability of the 21 items in this study across cohorts and time points ranged from 0.92 to 0.96, as evaluated using Cronbach’s alpha.

4.5 Semi-structured focus group discussions (FGDs)

To better understand the changes in design thinking mindset after going through the Design Innovation module, we asked

students about the dispositions they developed after the module, valuable takeaways or skills from the module that they could transfer to other settings, how the module sharpens their intellectual muscle, and how design thinking overlaps with critical thinking. We conducted three FGD sessions with 23 students, each lasting 60 min on average. The FGDs were audio-recorded and transcribed verbatim.

4.6 Data analyses

To answer RQ1 (“How do students’ DT and CT mindsets change after participation in a Design Innovation module?”), a paired samples *t*-test was used to assess the change in CT disposition and DT mindset from pre-test to post-test for both cohorts. This was followed up with linear regression to examine the degree of change in CT disposition and DT mindset from pre-test to post-test for both cohorts, where appropriate. To answer RQ2, a linear regression of post-test DT mindset on pre-test DT mindset was conducted with pre-test CT disposition as a moderator.

The FGD data were qualitatively coded according to the 19 DT mindset elements (e.g., human centeredness, empathy/empathic, mindfulness and awareness of process) used to develop the items in the DT mindset tool (Dosi et al., 2018). The third author performed the coding, and all authors discussed the results in depth. Although agreement among the authors was high, the outcome was only agreed upon after several rounds of discussion. Table 2 illustrates the number of verbatim quotes extracted for each category of DT mindset element and examples of quotes for each element.

5 Results

5.1 Changes to critical thinking disposition and DT mindset

Prior to conducting the analyses, Levene's test was performed and we found no violation to the assumption of homogeneity of variances between the paired observations ($p > 0.05$). The skewness and kurtosis of each dependent variable also fell within acceptable range. With both cohorts, findings from a paired samples t -test showed a significant increase in students' CT disposition from pre-test to post-test, Cohort 1: $t(372) = 7.45$, $p < 0.01$, $d = 0.39$; Cohort 2: $t(624) = 9.52$, $p < 0.01$, $d = 0.38$. Similarly, with both cohorts, findings from a paired samples t -test also showed a significant increase in students' DT mindset from pre-test to post-test, Cohort 1: $t(372) = 5.57$, $p < 0.01$, $d = 0.29$; Cohort 2: $t(624) = 9.24$, $p < 0.01$, $d = 0.37$. See Figure 6.

As CT disposition and DT mindset were found to have significantly increased from pre-test to post-test for both cohorts, a linear regression conducted on the amount of change in students' CT disposition and DT mindset showed that no significant difference between the degree of change in CT disposition and DT mindset from pre-test to post-test, Cohort 1: $F(1, 744) = -0.05$, $p = 0.16$; Cohort 2: $F(1, 1,248) = 0.04$, $p = 0.84$. See Figure 7.

In line with the quantitative results, the students also described the development of a DT mindset after going through the module. They reported developing the following 13 DT dispositions: (1) the confidence of being more creative to think "out-of-the-box" (creative confidence), (2) critically questioning to understand the problem deeper, (3) the desire to make an impact on the people around them (desire to make a difference), (4) being more open to treading the unknown (embracing risk), (5) empathising with the concerns of the users (empathy/empathic), (6) being open to ideas and not being fixated with a single idea (envisioning new things) (7) considering the problem from a broader perspective beyond the pros and cons (holistic view/consider the problem as a whole), (8) understanding users' needs (human centeredness), (9) being aware of the design thinking processes (mindfulness and awareness of practice), (10) being open to working with people with diverse perspectives (multi-/inter-/cross-disciplinary collaborative teams), (11) being comfortable with diverse perspectives (open to different

perspectives/diversity), (12) becoming more opportunity focused (optimism to have an impact), and (13) valuing the opinions of teammates in the design thinking process (team working). We refer readers to Table 1 for some examples of the quotes relating to the above DT dispositions.

5.2 Critical thinking as a moderator

To answer RQ2, "Is critical thinking a pre-requisite for developing design thinking?" a moderated regression analysis was conducted for each cohort. Results of the linear regression of post-test DT mindset on pre-test DT mindset, with pre-test CT disposition as a moderator, found a significant moderation effect in both cohorts, Cohort 1: $F(3, 369) = 2.78$, $p = 0.04$; Cohort 2: $F(3, 621) = 47.5$, $p < 0.01$. See Table 3 for more details. In other words, in line with our hypothesis, the change in DT mindset from pre-test to post-test significantly depends on students' level of critical thinking disposition measured before the commencement of the module.

Regarding critical thinking dispositions interacting with design thinking mindsets, students reported in the FGDs that critical thinking skills were part and parcel of the design thinking process. One student described how critical thinking and design thinking blended in the following manner:

"It is something that makes us not only use our thinking skills but also we can learn how to come up with a certain prototype. Therefore, I feel that design and critical thinking blends together very well."

In addition, students described critical thinking as being embedded in the lived human experiences, suggesting that such a skill is generally essential for daily living, regardless of the module taken.

"I think everyone throughout their entire life has been applying critical thinking, whether it's in school or whether they are gossiping with their friends to find out more information, so it's just about learning how to apply that same concept of finding out more and seeing whether that information you get is credible and then applying it to different aspects of life."

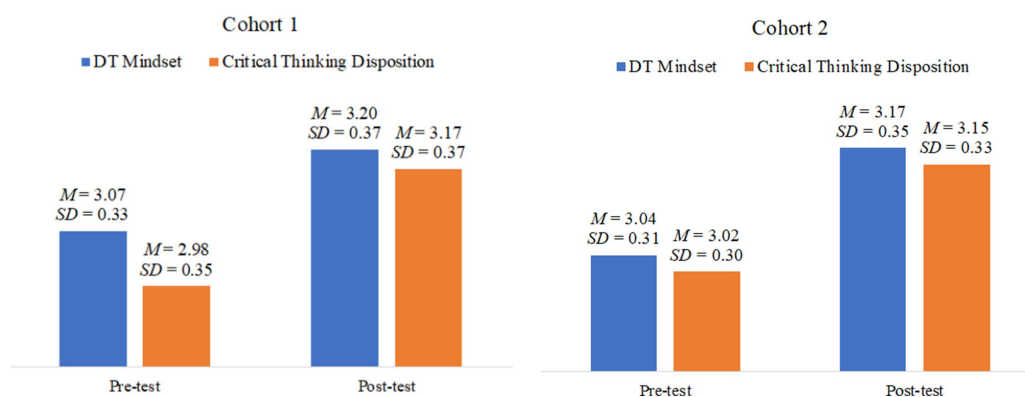


FIGURE 6

Descriptive statistics of changes in critical thinking disposition and DT mindset from pre-test to post-test in Cohort 1 and Cohort 2.



6 Findings

The findings from this study are consistent with previous studies (e.g., [Suligoj et al., 2020](#)). The study reveals that students' critical thinking, measured across three subscales, significantly influences their creative design ability. Higher critical thinking perception correlates with increased fluency and flexibility in ideas and greater usefulness in design outcomes. Notably, understanding misconceptions strongly predicts design usefulness and idea fluency, while a high trust in critical thinking negatively impacts design originality. It provides compelling evidence that participation in the Design Innovation module significantly enhances students' critical thinking disposition and design thinking (DT) mindset. The paired samples t-test results indicate a substantial increase in these areas for both cohorts, underscoring the module's effectiveness in fostering these essential skills. Notably, the linear regression analysis revealed no significant difference in the degree of change between critical thinking disposition and DT mindset, suggesting that the module equally contributes to developing both skill sets. This balance is crucial, as it aligns with the emerging educational paradigm that emphasises the integration of critical and design thinking for comprehensive problem-solving and innovation. Furthermore, the qualitative data from focus group discussions support these findings. Students' reflections on their experiences highlight the development of key DT dispositions, such as creative confidence, empathy, and the ability to embrace risk and diversity. These dispositions are critical for navigating the complexities of modern challenges and innovating in a rapidly changing world. The students' ability to articulate their growth in these areas is a testament to the module's impact on their cognitive and affective development.

Moreover, the moderated regression analysis provides an intriguing insight into the relationship between critical thinking and design thinking. The significant moderation effect observed in both cohorts suggests that a student's initial level of critical thinking disposition influences the extent of their development in DT mindset. This finding addresses the second research question, indicating that critical thinking may be a foundational element that enhances the cultivation of design thinking skills. This

interdependence highlights the importance of fostering critical thinking as a prerequisite or concurrent skill to maximize the benefits of design thinking education. We have not determined any studies examining such a moderation effect. The focus group discussions further illuminate this interplay, with students perceiving critical thinking as an integral part of the design thinking process. Their descriptions of how these skills blend and apply to real-life scenarios underscore the practical relevance of the module. This real-world applicability is crucial for preparing students to be effective problem-solvers and innovators in their future careers.

Our findings show that the interplay between critical thinking mindset and design thinking dispositions underscores a dynamic teaching and learning experience where both skill sets mutually reinforce each other. The moderated regression results suggest that students predisposed to engage critically are likelier to excel in design thinking instruction that prioritises iterative problem-solving and creativity ([Carroll et al., 2010](#); [Johansson-Sköldberg et al., 2013](#)). This synergy is critical in educational settings where the goal is to solve problems and redefine them innovatively. Our study corroborates findings by [Wrigley and Straker \(2017\)](#), who observed that involvement in DT tasks enhances critical thinking by immersing students in situations where they must critically evaluate and navigate complex scenarios. This environment fosters a CT disposition characterized by an active and reflective engagement with problems, potentially leading to more innovative outcomes and a deeper understanding of the problem space.

However, a limitation of the present study is that we have yet to examine what aspects of the module enhance specific elements of DT or CT more significantly than others. Hence, we recommend educators integrate CT skill development intentionally into the curriculum and in tandem with DT skill development so that there is a balanced approach to developing critical competence and creative confidence. We also encourage interdisciplinary learning and real-world problem-solving in collaborative projects that value diverse perspectives. Integrating critical debates and iterative prototyping exercises can also enhance students' application of critical thinking in creative design. Through such practical, applied learning experiences, educators of interdisciplinary design innovation can aim to improve students' CT and DT mindsets.

TABLE 2 Number of verbatim quotes extracted and examples of quotes for each DT mindset element.

DT mindset element	No. of quotes	Examples of quotes
Creative confidence	7	"When it comes to nursing, in a mental health setting, we create activities for the patient. This requires innovation and creativity. Therefore, the concept of thinking out of the box could be applied to brainstorming of the activities for the patients."
Critical questioning	5	"I think we should not jump to a conclusion. We should try to understand better and ask more questions, instead of coming up with the solution immediately after hearing about the problem."
Desire to make a difference	3	"For me right now, I'm looking at certain things at different perspective. I tend to think about how to make things better for both my patients as well as myself."
Embracing risk	1	"For myself, I feel that this module has offered me in terms of pushing the boundaries, and to challenge the norms being set. For example, let us say you have this set of problem, you must solve it in this manner, but you can also approach it at another angle. I find that this module has drilled me in terms of being open to unknowns, and how I can further learn my terrain in a way I can step out of my comfort zone."
Empathy / empathic	3	"From this course, I learnt to be more empathetic towards the users by putting ourselves in the shoes of other."
Envisioning new things	1	"One of the things that impacted me was that time when we created our first prototype, then we had to tear the paper when the profession told us to. It really helped us to not be so stuck on that single idea. Sometimes we would think that the idea we came up with is the perfect idea, but others may think otherwise. So, this really helped us to get over it and not see the initial idea as final."
Holistic view / consider the problem as a whole	4	"This module taught us to not jump straight to solutions. Usually when you encounter a problem, you would try to think of a solution straight away. However, this module makes you go through the whole process of finding out what other people think, how they feel, and they want to see being changed. By taking all these voices, you make a better holistic solution. It's not about me answering the question but going through the journey to find the answer."
Human centeredness	5	"It allows me to be in the perspective of someone who creates product. We need to think about how the users would use the product by putting ourselves in their perspective and where we can find problems that can be improved upon."
Mindfulness and awareness of practice	2	"In this module, for every slide that are given to us, it makes me think and it really strengthens the muscle, and the brain power. In life, for every step of your journey, for instance, when you go down the stairs, or climb up the stairs, those movements require certain brain muscle to tell you that you have to do this, do that. And with this module, I realize that I start to think why I do this."
Multi-/inter-/cross-disciplinary collaborative teams	4	"I think it's also good to have different people talking about it and sharing ideas and throwing it around is easier to figure out what is the best because everyone has different perspective and background where they grow up in. So, from there, we kind of have more perspective to work with."
Open to different perspectives / diversity	4	"I think regarding DT, I'm quite fixated on the idea itself, but during one of the lessons, the instructor said to let our ideas go wild. I think that's quite good, because sometimes you come up with solutions during the process of coming up with weird ideas. I've become more open minded to view things from various perspective using different lens."
Optimism to have an impact	1	"I feel like this module helps me to be more opportunity focused. I have the opportunity to find any challenges or problems faced by the community. And it helps me to become more of myself as I realize that I can come up with any of the crazy ideas when brainstorming."
Team working	2	"Group discussion allows me to open towards another person and acknowledge that their ideas are better than mine. As a human, we only think that our idea is so much better than anyone else. So, when we have this kind of group discussion, it makes me feel that you have to put yourself down and be willing to listen to other thoughts, like 'Why I did not think of that?'"

7 Discussion

Meta-skills, global competence, and digital literacy are potential directions for future skills development. According to Ehlers (2020), *Future Skills* are defined as competencies that allow individuals to solve complex problems in highly emergent contexts of action in a self-organized way and enable them to act

(successfully). They are based on cognitive, motivational, volitional and social resources, are value-based and can be acquired in a learning process (Ehlers, 2020, p. 53). Over the years, approaches to 21st-century skills and Future Skills have often focused on designing policy framework recommendations and may only sometimes be empirically based. Existing approaches to Future Skills generally consist of lists of essential

TABLE 3 Moderated regression analysis results.

Cohort 1		
	β (SE)	p
(Intercept)	4.42 (0.68)	<0.01
Pre-test DT mindset	−0.46 (0.23)	0.04
Pre-test critical thinking disposition	−0.45 (0.24)	0.06
Pre-test DT mindset * Pre-test critical thinking disposition	0.16 (0.07)	0.02
Cohort 2		
	β (SE)	p
(Intercept)	2.96 (0.56)	<0.01
Pre-test DT mindset	−0.17 (0.19)	0.38
Pre-test critical thinking disposition	−0.19 (0.19)	0.32
Pre-test DT mindset * Pre-test critical thinking disposition	0.14 (0.06)	0.02

competencies but are based on something other than sound competence theory approaches (Barrie, 2004). While it is necessary to understand the characteristics of students when designing a curriculum so that HE educators can tailor the educational approach to meet their specific learning needs, more still needs to be done to explore the effectiveness of learning processes based on 21st-century skills that cater to the interests and needs of digital natives (Dilekçi and Karatay, 2023).

Today’s education system caters to 21st-century students who believe in continually producing new products and using technology to maintain their lives (Duran and Ertan Özen, 2018). If we prioritize future skills in HE, it becomes clear that we need to rethink the traditional roles of research, teaching, and learning in IHLs. In this context, it is essential to recognize that anything easily teachable and assessable is also easily digitized and prone to automation (Mourtzis et al., 2018). On the other hand, future skills like creativity, self-organization, self-reflection, and design thinking require innovative and sophisticated approaches to learning, teaching, and development (Miranda et al., 2020). To successfully integrate future skills into HE, educators must focus on implementing active and creative teaching methods and learning objectives that require complex assessments of competence beyond mere knowledge transfer (Erpenbeck, 2012). This will require a shift from traditional, passive approaches to education toward a more dynamic and engaged learning experience that focuses on developing specific transferable skills.

Furthermore, with the emergence of advanced technologies like AI and automation, graduates need to develop crucial digital literacy skills. Digital literacy refers to the ability to use technology effectively and responsibly, to understand the ethical implications of technology, and to leverage technology to solve problems and create opportunities (Khan et al., 2022). Thus, it is vital for IHLs to remain adaptable and responsive to these changing needs and to continue to prioritize the development of well-prepared graduates for the challenges and opportunities of the future. However, while 21st-century skills continue to be essential, the evolving nature of the world and the workforce means that new skills will likely emerge as critical for success.

8 Conclusion

Investigating mindset shifts is essential because it provides insights into how individuals adapt and evolve thinking processes, which is crucial for developing future-ready graduates. Our data shows significant increases in mindset shifts toward critical and design thinking, indicating a transformative educational impact. These shifts suggest that students are not only acquiring knowledge but are also reshaping their cognitive frameworks to approach problems more innovatively and critically. This transformation is vital for innovation in real-world problem-solving. Future research can explore a complementary framework (e.g., Cross, 2011) for understanding how these mindset shifts can lead to innovative and effective problem-solving in real-world contexts. We plan to continue the present study and investigate students’ future thinking mindsets and the changes that have occurred since they participated in the Design Innovation module. Similar to this study, we will administer pre- and post-test surveys for students to self-report their futures thinking mindset shift. The survey adopts the five factors of 33 items of the revised Foresight Styles Assessment (FSA) from Chen et al. (2021). In conclusion, the findings on increased mindset shifts in critical and design thinking among students is a promising indicator of the potential to develop graduates with key dispositions and are equipped with the cognitive frameworks necessary to address the complex challenges of the future.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Singapore Institute of Technology. Ethics approval was granted on 08 October 2021 (IRB: 2021154). We acknowledge two other co-investigators of the research study who will be named later. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NP: Conceptualization, Data curation, Project administration, Visualization, Methodology, Writing – original draft, Writing – review & editing. SP: Methodology, Writing – original draft, Writing – review & editing. X-FK: Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research project “Blending Design and Critical Thinking in Design Innovation Modules”

was awarded an Applied Learning and Innovation Grant (ALIGN) by the Singapore Institute of Technology from Jan 2022 to July 2023.

Acknowledgments

We acknowledge two other co-investigators of the research study, Mark Teo and Jawn Lim.

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

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, A., and Persson, T. (2015). Strategies for teaching students to think critically: a meta-analysis. *Rev. Educ. Res.* 85, 275–314. doi: 10.3102/0034654314551063
- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., et al. (2008). Instructional interventions affecting critical thinking skills and dispositions: a stage 1 meta-analysis. *Rev. Educ. Res.* 78, 1102–1134. doi: 10.3102/0034654308326084
- Aoun, J. E. (2017). *Robot-proof: Higher education in the age of artificial intelligence*. Cambridge, MA: The MIT Press.
- Avsec, S., and Ferk Savec, V. (2019). Creativity and critical thinking in engineering design: the role of interdisciplinary augmentation. *Glob. J. Eng. Educ.* 21, 30–36.
- Barrie, S. C. (2004). A research-based approach to generic graduate attributes policy. *High. Educ. Res. Dev.* 23, 261–275. doi: 10.1080/0729436042000235391
- Bezhnar, S. (2019). *The critical standard for design thinking*. Austin, TX: Softserve.
- Brookfield, S. D. (2012). *Teaching for critical thinking: Tools and techniques to help students question their assumptions*. San Francisco, CA: Jossey-Bass.
- Brown, T. (2009). *Change by design: How design thinking transforms organizations and inspires innovation*. New York, NY: Harper Business.
- Buchanan, R. (1992). Wicked problems in design thinking. *Des. Issues* 8, 5–21. doi: 10.2307/1511637
- Buchanan, R. (2015). Worlds in the making: design, management, and the reform of organizational culture. *She Ji* 1, 5–21. doi: 10.1016/j.sheji.2015.09.003
- Carroll, M., Goldman, S., Britos, L., Koh, J., Royalty, A., and Hornstein, M. (2010). Destination, imagination and the fires within: design thinking in a middle school classroom. *Int J Art Design Educ.* 29, 37–53. doi: 10.1111/j.1476-8070.2010.01632.x
- Chen, K. H., Hsu, L. P., and Hoffman, J. (2021). Reevaluating the foresight styles assessment: a measurement of futures competency for university students. *J. Futu. Stud.* 26, 19–31. doi: 10.6531/JFS.202109_26(1)
- Cloete, M. B. (2019). The impact of an integrated assessment on the critical thinking skills of first-year university students. *Acc. Educ.* 27, 479–494. doi: 10.1080/09639284.2018.1501717
- Concilio, G., de Götzén, A., Molinari, F., Morelli, N., Mulder, I., Simeone, L., et al. (2019). “Innovation and design” in *Innovation capacity and the City*. eds. G. Concilio and I. Tosoni (Cham: Springer Briefs in Applied Sciences and Technology: Springer).
- Cross, N. (2011). *Design thinking*. Oxford, England: Berg.
- Dilekçi, A., and Karatay, H. (2023). The effects of the 21st century skills curriculum on the development of students’ creative thinking skills. *Think. Skills Creat.* 47:101229. doi: 10.1016/j.tsc.2022.101229
- Dosi, C., Rosati, F., and Vignoli, M. (2018). Measuring design thinking mindset, Proceedings of International Design Conference, DESIGN, Faculty of Mechanical Engineering and Naval Architecture, 1991–2002.
- Dunne, A., and Raby, F. (2001). *Design noir: The secret life of electronic objects*. Bloomsbury Publishing, UK: Springer Science & Business Media.
- Duran, E., and Ertan Özen, N. (2018). Digital literacy in Turkish lessons. *Tur. Educ. J.* 3, 31–46.
- Duro, E., Elander, J., Maratos, F. A., Stupple, E. J. N., and Aubeeluck, A. (2013). In search of critical thinking in psychology: an exploration of student and lecturer understandings in higher education. *Psychol. Learn. Teach.* 12, 275–281. doi: 10.2304/plat.2013.12.3.275
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.
- Dwyer, C. P., Hogan, M. J., and Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacogn. Learn.* 7, 219–244. doi: 10.1007/s11409-012-9092-1
- Ehlers, U.-D. (2020) in *Future skills: the future of learning and higher education*. eds. U. D. Ehlers, P. Bonaudo and L. Eigbrecht (Wiesbaden: Springer).
- Ennis, R. H. (2018). Critical thinking across the curriculum: a vision. *Topoi* 37, 165–184. doi: 10.1007/s11245-016-9401-4
- Epstein, D. (2019). *Range: Why generalists triumph in a specialized world*. New York: Riverhead Books.
- Ericson, J. D. (2021). Mapping the relationship between critical thinking and design thinking. *J. Knowl. Econ.* 13, 406–429. doi: 10.1007/s13132-021-00733-w
- Erpenbeck, J. (2012). “Führungskompetenz” in *Kompetenz. Festschrift Prof. Dr. John Erpenbeck zum 70. Geburtstag*. ed. W. G. Faix, vol. Band 4 (Stuttgart: Steinbeis-Edition), 109–142.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. USA: California Academic Press.
- Facione, P. A., Facione, N. C., and Giancarlo, C. (1992). *The California critical thinking skills test*. USA: California Academic Press.
- Halpern, D. F. (2010). *Undergraduate education in psychology: A blueprint for the future of the discipline*. PsycInfo Database Record (c) 2023 APA, USA: American Psychological Association.
- Jakobsone, L. (2017). Critical design as an approach to next thinking. *Des. J.* 20, S4253–S4262. doi: 10.1080/14606925.2017.1352923
- Janssen, E. M., Meulendijks, W., Mainhard, T., Verkoijen, P. P. J. L., Heijltjes, A. E. G., van Peppen, V., et al. (2019). Identifying characteristics associated with higher education teachers’ cognitive reflection test performance and their attitudes towards teaching critical thinking. *Teach. Teach. Educ.* 84, 139–149. doi: 10.1016/j.tate.2019.05.008
- Johansson-Sköldberg, U., Woodilla, J., and Çetinkaya, M. (2013). Design thinking: past, present and possible futures. *Creat. Innov. Manag.* 22, 121–146. doi: 10.1111/caim.12023
- Khan, N., Abdullah, S., Chen, T. B., and Khan, S. (2022). Connecting digital literacy in higher education to the 21st century workforce. *Knowl. Manag. E-Learn.* 14, 46–61. doi: 10.34105/j.kmel.2022.14.004
- Kimbell, L. (2011). Rethinking design thinking: part 1. *Design Cult.* 3, 285–306. doi: 10.2752/175470811X13071166525216
- Koh, J. H. L., Chai, C. S., Wong, B., and Hong, H.-Y. (2015). *Design thinking for education: Conceptions and applications in teaching and learning*. Singapore: Springer.
- Kolko, J. (2015). Design thinking comes of age. *Harv. Bus. Rev.* Available at: <https://hbr.org/2015/09/design-thinking-comes-of-age>
- Kraaijenbrink, J. (2022). *What BANI really means (and how it corrects your world view)*. Forbes.
- Liedtka, J., Salzman, R., and Azer, D. (2017). *Design thinking for the greater good*. New York: Columbia Business School.
- Loewe, S. (2019). Towards a critical design thinking. *Dialectic*, 2, 132–156. Michigan, USA: AIGA Design Educators Community (DEC), Michigan, USA. Michigan Publishing.
- Malpass, M. (2013). Between wit and reasoning: defining associative, speculative, and critical Design in Practice. *Design Cult.* 5, 333–356. doi: 10.2752/175470813X13705953612200

Publisher’s note

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

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2024.1358431/full#supplementary-material>

- Malpass, M. (2015). Criticism and function in critical. *Design Pract.* 31, 59–71. doi: 10.1162/DESI_a_00322
- Malpass, M. (2017). *Critical design in context: History, theory, and practice*. New York, NY, USA: Bloomsbury Academic.
- Miranda, J., Rosas-Fernandez, J. B., and Molina, A. (2020). Achieving innovation and entrepreneurship by applying education 4.0 and open innovation. Proceedings of the 2020 IEEE international conference on engineering, technology and innovation (ICE/ITMC), New York, NY, USA, 1–6.
- Mosely, G., Wright, N., and Wrigley, C. (2018). Facilitating design thinking: a comparison of design expertise. *Think. Skills Creat.* 27, 177–189. doi: 10.1016/j.tsc.2018.02.004
- Mourtzis, D., Vlachou, E., Dimitrakopoulos, G., and Zogopoulos, V. (2018). Cyber-physical systems and education 4.0 –the teaching factory 4.0 concept. *Procedia Manuf.* 23, 129–134. doi: 10.1016/j.promfg.2018.04.005
- Nickerson, R. S. (1999). “Enhancing creativity” in *Handbook of creativity*. ed. R. J. Sternberg (New York, NY: Cambridge University Press), 392–430.
- Osmani, M., Weerakkody, V., Hindi, N., and Eldabi, T. (2019). Graduates employability skills: a review of literature against market demand. *J. Educ. Bus.* 94, 423–432. doi: 10.1080/08832323.2018.1545629
- Patel, N. S. (2021). Development of criticality in thought: a conceptual framework for online student discussion forums. *Int. J. TESOL Stud.* 3, 22–40. doi: 10.46451/ijts.2021.09.02
- Patel, N. S. (2022). Making critical thinking explicit and visible with dialogic scaffolding. Workshop facilitation at the Redesigning Pedagogy International Conference: Educating for Innovation, Nurturing for Society.
- Patel, N. S., and Wong, J. (2022). Understanding critical thinking, reading and writing: a minimal English approach. Paper presentation at the Redesigning Pedagogy International Conference: Educating for Innovation, Nurturing for Society.
- Paul, R., and Elder, L. (2012). *Critical thinking: Tools for taking charge of your learning and your life*. Boston: Pearson.
- Paul, R., and Elder, L. (2019). *The miniature guide to critical thinking concepts and tools*. 8th Edn. UK: Rowman & Littlefield Publishers.
- Pearson (2018). *Partnership for 21st-century learning, 4C research series*: Battelle for Kids.
- Pedro, F. (2006). The new millennium learners: challenging our views on ICT and learning oecd-ceri. 1–17. Available at: <http://www.oecd.org/dataoecd/1/1/38358359.pdf>.
- Puig, B., Blanco-Anaya, P., Bargiela, I. M., and Crujeiras-Pérez, B. (2019). A systematic review on critical thinking intervention studies in higher education across professional fields. *Stud. High. Educ.* 44, 860–869. doi: 10.1080/03075079.2019.1586333
- Ramírez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., and Miranda, J. (2022). Complex thinking in the framework of education 4.0 and open innovation—a systematic literature review. *J. Open Innov.: Technol. Mark. Complex.* 8:4. doi: 10.3390/joitmc8010004
- Rodgers, P. (2017). Paradoxes in design thinking. *Des. J.* 20, S4444–S4458. doi: 10.1080/14606925.2017.1352941
- Runco, M. A. (2014). “Convergence and divergence in creativity” in *Convergence of creativity and innovation*. ed. M. A. Runco (New York, NY: Cambridge University Press).
- Schweitzer, J., Groeger, L., and Sobel, L. (2016). The design thinking mindset: an assessment of what we know and what we see in practice. *J. Design Busi. Soci.* 2, 71–94. doi: 10.1386/dbis.2.1.71_1
- Simonovic, B., Vione, K. C., Fido, D., Stupple, E. J. N., Martin, J., and Clarke, R. (2022). The impact of attitudes, beliefs, and cognitive reflection on the development of critical thinking skills in online students. *Online Learn.* 26, 254–274. doi: 10.24059/olj.v26i2.2725
- Sosu, E. M. (2013). The development and psychometric validation of a critical thinking disposition scale. *Think. Skills Creat.* 9, 107–119. doi: 10.1016/j.tsc.2012.09.002
- Suligoj, V., Zavbi, R., and Avsec, S. (2020). Interdisciplinary critical and design thinking. *Int. J. Eng. Educ.* 36, 84–95.
- Torrance, E. P. (1966). *Torrance test of creative thinking technical manual*. Princeton: Personnel Press Inc.
- Treffinger, D. J., and Isaksen, S. G. (2005). Creative problem solving: the history, development, and implications for gifted education and talent development. *Gifted Child Quart.* 49, 342–353. doi: 10.1177/001698620504900407
- Tsui, L. (2002). Fostering critical thinking through effective pedagogy. *Res. High. Educ.* 40, 740–763.
- Turnali, K. (2016). *Innovation with design thinking demands critical thinking*. Forbes.
- Verganti, R. (2009). *Design-driven innovation: Changing the rules of competition by radically innovating what things mean*. Harvard Business Press.
- Wagner, T. (2008). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need and what we can do about it*. New York: Basic Books.
- Wrigley, C., and Straker, K. (2017). Design thinking pedagogy: the educational design ladder. *Innov. Educ. Teach. Int.* 54, 374–385. doi: 10.1080/14703297.2015.1108214



OPEN ACCESS

EDITED BY

Pedro Gil-Madrona,
University of Castilla-La Mancha, Spain

REVIEWED BY

Francisco D. Fernández-Martin,
University of Granada, Spain
Slavica Šimić Šašić,
University of Zadar, Croatia

*CORRESPONDENCE

P. Robledo
✉ probr@unileon.es

RECEIVED 16 February 2024

ACCEPTED 20 June 2024

PUBLISHED 02 July 2024

CITATION

Robledo P and Álvarez ML (2024) Strategic and self-regulated instruction in synthesis tasks in the university context. *Front. Psychol.* 15:1386907. doi: 10.3389/fpsyg.2024.1386907

COPYRIGHT

© 2024 Robledo and Álvarez. 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.

Strategic and self-regulated instruction in synthesis tasks in the university context

P. Robledo* and M. L. Álvarez

Department of Psychology, Sociology and Philosophy, University of León, León, Spain

Introduction: The aim of the study was to improve student skills in writing good-quality synthesis texts through a strategic, self-regulated instruction program aimed at ensuring that students properly activated reading and writing strategies required by the synthesis task.

Methods: The sample consisted of 84 university students who were randomly assigned to experimental or control conditions. The experimental group received an instructional program based on the development and self-regulated implementation of reading and writing strategies for producing synthesis texts. The control group received a program involving metacognitive knowledge of various academic text types. Both programs involved eight 60-min sessions, taught by teachers in a compulsory degree subject. For the evaluations, students produced synthesis texts from different source texts. The syntheses were graded considering text product measures: information selection, idea connection, text organization, and holistic quality; and measures of reading (underlining and note-taking) and writing (planning and review) strategies.

Results: The results show that the experimental group exhibited greater improvements in synthesis quality and greater improvements in activation of information organization processes, note-taking while reading, and text planning.

Discussion: In conclusion, university students can, following implementation of a strategic instructional procedure in the context of a study plan, adapt and re-work their own reading and writing strategies and apply them in a self-regulated manner to synthesis tasks, improving text quality and some of the cognitive processes involved.

KEYWORDS

reading strategies, strategic and self-regulated instruction, synthesis task, text quality, university, writing strategies

1 Introduction

Nowadays in higher education, students are often expected to demonstrate their communicative competency when given tasks in which they have to manage information from various sources and produce written work based on that information (Boscolo et al., 2007; Roldán et al., 2011). These tasks follow the typical format of synthesis tasks (ST), which teachers routinely ask their students to do in the classroom as part of their teaching-learning activities (Perin, 2013).

STs are hybrid tasks that require students to use a combination of reading comprehension skills and writing skills (Granello, 2001; Solé et al., 2005, 2013). They involve analysing different sources of information about a topic and producing a new text from that information, appropriately comparing, transforming, and combining ideas in an organized way. Students have to select information, connect it, and organize it in a new text structure to produce new,

original written discourse (Spivey and King, 1989; Spivey, 1997). Successfully completing STs demands selecting, organizing, and connecting information. Students must select information because they cannot retain all of the information from what they read, they must choose information to deal with based on relevance or importance. They also have to organize what they have selected within some structure in order to construct a mental representation of the texts, and they must connect the information at a local level to form a new cohesive, coherent discourse (Spivey and King, 1989). STs require students to use recursive reading and writing strategies (Flower et al., 1990; Spivey, 1997; Segev-Miller, 2004, 2007; Boscolo et al., 2007; Mateos and Solé, 2009; Solé et al., 2013). They need to read and re-read the source texts and their own writing, re-examining the source texts to identify relevant information and incorporate it, and re-examining their own text to correct it in a recurrent process allowing the author to transform the source information (Segev-Miller, 2004). ST is also a demanding writing task requiring the mediation of planning, textualization, and monitoring throughout the whole process (Flower and Hayes, 1980), as well as the use of review strategies ensuring a good fit between the student's on text and the source texts. Previous studies have demonstrated differences in the reading and writing strategies used by students with different levels of expertise in STs, as well as a certain correspondence between the holistic quality of the written product and the strategies used. They also indicate that the more successful ST products are associated with using a range of strategies, and that failed texts are associated with students adopting simple, direct, linear processes that are inadequate for the complexity of STs (Solé et al., 2005, 2013; Mateos and Solé, 2009; Du and List, 2020; Nadal et al., 2021).

The inherent cognitive complexity of STs means that they are highly demanding tasks at the level of cognitive processing and activity, which leads to some university students finding them difficult. Some students approach writing STs in a linear, fixed manner, limiting themselves to repeating the content from the sources and not incorporating or properly organizing the information, as well as applying strategic approaches that are less than effective (Boscolo et al., 2007; Mateos and Solé, 2009; Cumming et al., 2016). These students seem to need specific training that will enable them to do synthesis tasks, and there is empirical evidence of how guiding students to perform these types of tasks produces very satisfactory results. Some studies have confirmed the efficacy of ST instruction in the university context. For example, Zhang (2013) examined the effect of instruction in ST in 190 students split into a control and an experimental group. In the experimental group, five iterations of instruction about writing syntheses were incorporated into the existing course curriculum, which was built on reading, writing, and vocabulary instruction from two textbooks. The students in the experimental group were guided through the discourse synthesis writing process and a range of methods such as reading guides, connection exercises, and peer and teacher feedback on first drafts were used to scaffold the students' synthesis writing practice. The control group had an equivalent amount of reading and writing practice in other writing types within the same curriculum. The experimental group had significantly better performance in writing synthesis texts, and the study also confirmed the feasibility of incorporating ST instruction in the classroom without significantly altering the study plan, albeit with a notable amount of scaffolding by teachers, materials or peer support. Boscolo et al. (2007) performed

an intervention in which 52 students were asked to analyse the characteristics of good and poor synthesis texts, and to write and revise their own. The intervention included 10 sections, with both theoretical lectures on the writing models and practical workshops. Participants practiced academic writing by composing texts which were revised by the teacher, and by analysing and discussing examples of good and poor academic texts. The results showed an improvement in the students' abilities to write synthesis texts. The intervention was mostly focused on analysis of text characteristics and had an impact on the students' ability to select information based on structural cues in the text. However, the intervention did not affect the students' ability to integrate this information in a coherent representation and text structure, probably because it did not include training in procedural or strategic knowledge.

In the Spanish context some studies have included strategic or procedural training. For example, Mateos et al. (2018) evaluated the effectiveness of two types of interventions aimed at improving performance in synthesis tasks, including procedural training. One used a graphical guide with peer practice, the other added explicit instruction with modelling and explanation of the processes involved in synthesis tasks (procedural training). The results showed that the students who had the additional explicit instruction demonstrated substantial improvements in the quality of their work. Granado-Peinado et al. (2019) analysed the impact of interventions involving collaborative writing of argumentative syntheses, which combined explicit instruction with video modelling of writing and collaboration processes, along with a guide and collaborative practice. The effects of this intervention were compared with three other programs in which the amount of help was progressively reduced. Their results showed that in order for students to achieve the desired levels of competency, interventions should include explicit instruction with video modelling. When this instruction brings together help aimed at improving production of argumentative syntheses with help designed to promote collaboration, students can assimilate a greater amount and wider range of information. However, for identifying a high level of arguments, explicit instruction focused solely on helping students write argumentative syntheses turns out to be as effective as help aimed at collaboration. In addition, after interventions encouraging collaborative work, students successfully transfer their newly developed skills to their own, individual writing tasks. In both studies, teachers demonstrated how to perform ST, teaching specific procedures and strategies. Students usually learn to emulate these procedures but do not produce their own strategies.

All of these previous studies have demonstrated the efficacy of instruction in ST in the university context and included content or components of strategic, self-regulated instruction, such as declarative or procedural metacognitive knowledge linked to the text product or the writing process, the use of cognitive strategies, direct teaching about procedural and discourse knowledge writing strategies, modelling, and collaborative or individual practice. The strategic instructional approach has been shown to be one of the most effective ways to improve students' writing competence at various educational levels (McArthur and Philippakos, 2013, 2022; Graham et al., 2018; Graham and Harris, 2018; van Ockenburg et al., 2019), including university (McArthur et al., 2015, 2022). That said, its complex, multicomponent nature makes it difficult for teachers to use in the real university classroom, it needs a lot of time to implement, along with specific teacher training that normally requires professional

development programs (Harris et al., 2022). Furthermore, implementation of strategic instruction programs aimed at encouraging mastery of ST at the university level must consider the characteristics of the students, who usually exhibit relatively defined or mature strategic profiles (Kieft et al., 2007; Robledo et al., 2018; Arias-Gundín and Robledo, 2023). This means that rather than focusing on giving students specific strategies to produce their syntheses, instruction in ST at university should focus on guiding and prompting students so that, as part of the natural learning process in their subjects, they develop and implement their own reading and writing strategies in a self-regulated manner when doing ST. This would allow them to manage their cognitive and motivational processes and keep on with a task even if they become discouraged or find it difficult, which would contribute to a successful approach and consequent learning or mastery (Butler, 1998; van Ockenburg et al., 2019). This is exactly where the present study sits, with the aim of assessing whether it is possible to improve university students' competencies for synthesis tasks, via strategic instructional processes which, as part of the context of the university subjects, promote self-regulated application of reading and writing strategies by the students. A line that, as far as we are aware, previous research has not addressed.

1.1 Present study

The study looked at an intervention to improve university students' ability to compose good-quality synthesis texts. The intervention implemented a program of strategic and self-regulated instruction so that students would use suitable reading and writing strategies consistent with the demands of synthesis tasks. The instruction was based on the Strategy Content Learning model (SCL, Butler, 1992) theoretically based on self-regulated learning, which is crucial for students' academic performance in all education levels (Alonso-Tapia, 2012; Rosário et al., 2017). According to that model, when students face an academic task, they enter a recursive cycle of cognitive and metacognitive activities that involve analysing the demands of the task, selecting, adapting, and employing suitable strategic approaches, monitoring the application of these strategies, producing feedback, and readjusting (see revision in Robledo and García, 2017). In the context of written work specifically, the SCL model has proven effective when applied to students in further and higher education. Students improve the quality of their compositions, increase their metacognitive knowledge about this kind of writing, and improve their ability to adapt and effectively activate their own writing strategies for various writing tasks, taking a self-regulated approach to composing texts. They also have an improved sense of self-efficacy and they improve their attributional patterns (Butler, 1994, 1995, 1996, 1997, 1998).

For the present study, the experimental group were instructed according to the SCL model through a program focused on promoting self-regulated application of their own reading and writing strategies to the production of a synthesis text. The study also included a control group, who received instruction on metacognitive knowledge around different types of academic texts at university, including synthesis. The question the study addresses is whether the students in the experimental condition—instructed strategically according to the SCL model—effectively adapt and activate their reading and writing strategies in the production of synthesis texts, and consequently produce better quality texts than

the control group. Based on the evidence about the efficacy of strategic, self-regulated instruction in the general field of writing (Graham et al., 2018; Graham and Harris, 2018), about the SCL model in particular (Butler, 1994, 1995, 1996, 1997, 1998), and about university students' instruction in synthesis tasks (Boscolo et al., 2007; Zhang, 2013; Mateos et al., 2018; Granado-Peinado et al., 2019), our initial hypothesis is that we expect an increase in the effective activation of the reading and writing strategies involved in producing synthesis texts, as well as an improvement in the quality of the written product from the students in the experimental group compared to the control. In other words, we anticipate that the students in the experimental group will produce or adapt their own reading and writing strategies to the execution of synthesis tasks so that they improve their performance at the level of the written product, producing better quality syntheses than the control group.

2 Methodology

2.1 Sample

The sample comprised 84 students who were doing a degree in Early-childhood Education at a Spanish university. They were aged between 18 and 35 years old ($M = 20.18$ years, $SD = 2.74$). Just over a tenth (13.1%) were men, 86.9% were women ($\chi^2_{(4, N=84)} = 85.058$, $p < 0.001$). The students were selected through a non-probabilistic sampling process from four class groups in an obligatory subject in the first year of the degree (group 1, $n = 21$; group 2, $n = 20$; group 3, $n = 22$; group 4, $n = 21$). Two class groups were randomly assigned to the control condition (groups 1 and 2, $n = 41$, 12.2% men) and two to the experimental condition (groups 3 and 4, $n = 43$, 13.9% men). The experimental group received the program of strategic, self-regulated instruction on synthesis tasks. The control group received a program of instruction based on developing their metacognitive knowledge of the structure of various types of scientific or academic texts, including synthesis.

2.2 Measures

At each assessment timepoint (pre-test, post-test, and follow-up) students were given a writing task consisting of producing a synthesis text. The evaluator explained to the students that they had to compose a synthesis text after reading two source texts. The source texts were given to the students by the evaluator and dealt with the corresponding topic in each task. The synthesis text could be a maximum of two pages long. The tasks dealt with various topics (e.g., Metacognition, Self-concept, ICT and education, and Leisure and free time), counterbalanced by condition and timepoint. The source texts were created by the researchers using various psychology and education manuals, ensuring similar characteristics in terms of length (870–970 words), structure (the same number of paragraphs), and content (25–30 main ideas and 6–8 topic units, in both cases identified by three researchers independently). The students were asked to write the best synthesis text possible. They were given a blank draft page to use if they chose, and a blank sheet for their final text. The students were also free to markup the source texts as they wished (underline, annotate, highlight, etc.). There was no time limit for the tasks.

Each task was used to assess students' writing competence with two types of measures (see Table 1). The first type were text product measures, assessed via the synthesis texts themselves: selection of information, connecting ideas, text organization, and holistic quality of the text. These were assessed blind by two reviewers who had been trained beforehand, with a kappa inter-rater index over 0.70 in each of the dimensions analysed (Selection of information: 0.83, Connection: 0.71, Organization: 0.88, Quality: 0.70). The other type of measure was reading and writing strategies measures. Reading strategies, underlining and note-taking, were assessed by looking at how the students treated the source texts. In this case, a reviewer analysed the students' source texts and identified the number of underlined texts (none, one or two texts underlined), the type of underlining (undifferentiated: same colour, format; or differentiated: different colour, format), the number of texts with notes (0, one text or two text with notes), and the type of notes (without elaboration: only one word or symbol; with elaboration: sentences, arrows joining content sentences, explanations). Writing strategies, planning and revising, were assessed by examining the students' drafts (planning strategies) and final texts (revising strategies). For the planning strategies, the reviewer analyzed how students used the draft sheet to plan their synthesis texts, evaluating that based on the parameters given in Table 1. For the revising strategies, two reviewers examined the final synthesis text and identified whether corrections had been made or not, and if there were corrections, whether they were mechanical (spelling, handwriting, grammar) or substantive (adding, changing, deleting, reordering or adding to ideas, etc.).

2.3 Design and procedure

The study followed an experimental pre-test, post-test, follow-up design with a control group. The independent variable was the instruction condition (experimental-control) and the dependent variables were the reading and writing strategy measures and the text product measures. The students gave their informed consent and were then randomly assigned to each condition. The experimental condition was given the strategic program (summarized in Table 2), designed based on SCL. This followed a sequence of three phases in which the students, guided by the instructor, worked on identification, development, and effective, self-regulated application of their strategies. In the first phase, the instructor offered support and assistance to the students so that they could analyse the task, identify objectives, and define criteria for correctly completing it. In the second phase, the instructor guided students so that, starting with identifying their own strategies and analysing their effectiveness, they developed suitable strategic approaches to the task. In the final phase, the instructor supported the students while they applied their strategies, encouraging monitoring and guiding them in self-feedback, finally suggesting strategic reconsideration.

The control group also had 8 sessions of intervention about writing which involved working on structural analyses of different academic and scientific text types, including synthesis (see Table 3).

Both interventions focused on analysing and producing scientific-academic texts, including synthesis texts, and comprised eight teaching sessions lasting one hour each, held every two weeks during a term. Both had the same number of practices of synthesis tasks: all students had to write two ST about the same content—Motivation and Learning Strategies (content from the degree course the students were

doing)—following the same routine, first in pairs and then individually. The interventions were applied by (2) teachers who taught courses in the degree the study took the sample from (a subject in Developmental Psychology which addressed content related to optimizing child development). The teachers had prior experience of instructional research. This meant that they understood the differences between the two treatment conditions and were able to implement them properly.

The sessions were applied in parallel, so that each teacher gave four intervention sessions to the experimental groups and four to the control, each to a different subgroup of students to avoid possible biases related to teachers' preferences for one program or the other. In addition, the content for each session was detailed in an application protocol that teachers had to follow (see outline in Tables 2, 3). Compliance with the protocol was examined by a non-participating observer to ensure it was faithfully applied in each session. Finally, the sessions required students to do specific practical or applied tasks. These were collected in an individual portfolio and reviewed to verify the proper delivery of the sessions. Thus, all of the students were trained over the same number of sessions, with the same time constraints, the same number of written practices, and the same teacher characteristics. Students were assessed before (pre-test) and immediately after instruction (post-test), and one year later (follow-up). All of the students in the sample attended all of the evaluation and intervention sessions, meaning that there was no missing data.

2.4 Data analysis

Data analysis was performed using SPSS software, version 24. The statistics for kurtosis and skewness showed that the text product variables followed a normal distribution while the strategy variables did not (see Table 4). Because of that, for the text product variables, parametric repeated-measure MANOVAS time by condition were performed. To determine the immediate differential effects of the instruction on the text product measures, we performed an analysis of variance with 2×2 repeated measures, with time as the repeated measure (pre-/post-test) and the group as inter-subject factor (experimental or control). In addition, to determine whether, a year after the intervention, the students' situation in the different text product variables had changed from the initial situation (pre-test) or the post-test evaluation, we performed a 3×2 repeated measures analysis.

For the strategy variables, to analyse the immediate intra-subject effects of the treatments, we performed Wilcoxon's non-parametric analysis for two related samples independently for the two groups. Inter-subject effects were assessed using the Mann-Whitney U test for independent samples, testing both groups (experimental and control) at pre-test, post-test, and at follow-up.

3 Results

3.1 ST product measures

Beginning with the 2×2 repeated measures analysis, the multivariate tests of variance gave statistically significant results, with a moderate-large effect size, for the measures of text product: overall

TABLE 1 Measures of students' writing competence.

Measure	Scale
Product measures	
1. Selection of information	
1.1. Idea selection: Main ideas (MI) from source texts appearing in the synthesis	0 MI from only one source text 1 Only MI common to both texts 2 Almost all MI, both common to both texts and otherwise 3 All MI
1.2. Informative capacity: Topic units (TU) from source texts appearing in the synthesis	0 Lacking any TU 1 Less than half of the TU 2 Half of the TU 3 More than half of the TU 4 All of the TU
2. Connecting ideas	
2.1. Cohesion: incorporation of ideas from the source texts in the synthesis	0 Two separate summary texts of the sources juxtaposed. 1 Alternating ideas from the two sources, using connectors, subordinate phrases, etc. 2 Presents its own structure and incorporates ideas from both texts in new discursive units, combines ideas.
2.2. Creation-plagiarism: level of elaboration present in ideas in the synthesis	0 Literal copy of ideas in the source texts 1 Copies some ideas, paraphrases others 2 Copies and paraphrases, but also has some elaboration 3 Demonstrates elaboration
3. Text organization	
3.1. Overall structure of text: has an introduction, paragraphs, conclusion. Displays hierarchy	1 point each for presence of: - Introduction: summarizes, anticipates or presents content - Paragraphs: the information is organized in paragraphs - Conclusions: the writer's opinion, final considerations. - Hierarchy: organization of information according to importance and reciprocal relationships
3.2. Local structure: cohesion of paragraphs and ideas through discourse markers	Number of indicators: meta-structural, structural, reformulating, and argumentative.
4. Holistic quality	Rate the quality of the synthesis (1 low, 5 high) based on indicators of: introduction, organized details, fluid discourse, clear sequence of ideas, no digression, structural markers, conclusion, vocabulary, syntax, spelling, and presentation.
Reading strategy measures	
5. Underlining	
5.1. Underlining	0 Not present 1 Undifferentiated 2 Differentiated
5.2. N° of underlined texts	0 None of the source texts underlined 1 One source text underlined 2 Both source texts underlined
6. Note-taking	
6.1. Type of notes	0 No notes taken 1 Notes without elaboration 2 Notes with elaboration
6.2. N° of texts with notes	0 No notes on either text 1 Notes on one text 2 Notes on both texts
7. General reading strategy	0 Text is un-annotated 1 Only underlining 2 Only notes 3 Underlining and notes

(Continued)

TABLE 1 (Continued)

Measure	Scale
Strategy measures	
8. Planning	
8.1. Overall planning	The level of general planning activity shown by the draft 0 No draft 1 Draft poorly developed 2 Draft partly developed 3 Draft well-developed
8.2. Planning ideas	How the writer produces and organizes ideas: 0 Ideas developed in the text 1 Ideas organized according to a plan or structure 2 Ideas in a list, with no order, apparently unconnected
8.3. Type of planning	Allusion to or presence in the draft of: introduction, development, conclusion, objective, audience, and strategy. 0 No draft 1 No reference to this aspect 2 Graphical reference to this aspect (list, word) 3 Explicit allusion to this aspect, written or drafted
9. Revision	Number of effective revisions shown in the synthesis: mechanical, substantive, total

Measures collected in the “Synthesis assessment protocol,” produced by the authors’ team of experts on writing competency, following review and adaptation of various of their own tools and those from the scientific literature (Spivey and King, 1989; Boscolo et al., 2007; Martínez et al., 2011).

structure [$F_{(1, 82)} = 22,093, p < 0.001, \eta^2 = 0.276$], local structure: meta-structuring [$F_{(1, 82)} = 12,468, p = 0.001, \eta^2 = 0.177$] and reformulators [$F_{(1, 82)} = 7,366, p = 0.009, \eta^2 = 0.113$], and quality [$F_{(1, 82)} = 29,847, p < 0.001, \eta^2 = 0.340$]. The multivariate tests of variance in the 3×2 repeated measures test produced statistically significant results, with a moderate-large effect size, for the text product measures: overall structure [$F_{(2, 81)} = 10,561, p < 0.001, \eta^2 = 0.310$], local structure: meta-structuring [$F_{(2, 81)} = 6,349, p = 0.004, \eta^2 = 0.213$] and reformulators [$F_{(2, 81)} = 3,058, p = 0.056, \eta^2 = 0.115$], and quality [$F_{(2, 81)} = 24,647, p < 0.001, \eta^2 = 0.517$]. Table 5 shows the statistically significant results from the tests of inter-subject effects.

As Table 5 shows, the inter-subject effect of the condition shows that the treatment led to greater improvement in the experimental group than the control group. The improvement was in local structure (meta-structuring and reformulators) and in overall synthesis quality. This improvement continued at follow-up and extended to include the general structure dimension. There were no statistically significant differences in the other measures.

3.2 Results relating to students’ reading and writing strategies when producing STs

Wilcoxon’s test indicated statistically significant differences between pre-test and post-test in both the experimental and control conditions (see Table 6). In the experimental group, following instruction, the students improved in 3 variables related to activating reading strategies, 5 related to activation of text planning strategies, and 2 related to revising the texts. The control group improved activation of revision in general, and substantive revision in particular, with no statistically significant differences in any other variable.

The Mann-Whitney U test indicated that there were no statistically significant differences between the groups at pre-test for any variable. At post-test there were differences in the type of notes variable

($U = 288.000, Z = -2.734, p = 0.006$) and in planning: introduction ($U = 175.000, Z = -4.029, p < 0.000$), development ($U = 277.500, Z = -2.308, p = 0.021$), conclusion ($U = 253.000, Z = -3.055, p = 0.002$), and strategy ($U = 287.528, Z = -2.528, p = 0.011$). The experimental group had higher (mean rank) scores at post-test, as shown in Figure 1.

At the follow-up timepoint, there were statistically significant differences between the experimental and control groups in overall planning ($U = 192.000, Z = -2.648, p = 0.008$, experimental mean rank score = 30.9 vs. Control mean score = 20.50), Planning: development ($U = 158.000, Z = -2.381, p = 0.017$, experimental mean rank = 26.7 vs. Control mean rank = 18.4), and type of planning ($U = 161.500, Z = -2.084, p = 0.037$, experimental mean rank = 26.54 vs. Control mean rank = 18.6). In each case, the experimental group scored higher than the control.

4 Discussion and conclusions

The objective of this study was to improve university students’ competency in composing STs following implementation of a strategic, self-regulated instruction program designed to encourage students to effectively activate the reading and writing strategies demanded by synthesis tasks. To that end, we compared an experimental group, given instruction through a strategic program that was based on Strategy Content Learning (SCL, Butler, 1992), with a control group, who were taught metacognitive knowledge of different text types in the academic context. Our hypothesis, based on the literature in this field, was that compared to the control group, the students in the experimental group would exhibit greater effective activation of their reading and writing strategies during the synthesis tasks and would demonstrate more improvement in their performance in terms of text product, producing better-quality syntheses.

TABLE 2 Instructional program experimental condition.

Sequence	Focus	Instructor	Activities
Phase 1	Session 1 (60 min. February, week 2). Developing metacognitive knowledge of ST and its demands (product and strategies)	Teacher 1: group 3 Teacher 2: group 4	After producing an ST (pre-test): - Students describe what they felt the task demanded and the strategies they used. They answer two questions: - What should you do? - How did you do it? - ST analysis by the teacher, examining the true product demands and activation of reading-writing strategies and processes. - The teacher explains: What are synthesis tasks. What are their characteristics in terms of selection, connection, and organization processes. The demands of ST at the level of writing and reading processes - The teacher presents a self-monitoring sheet to assess the quality of the syntheses at the product level - Pool ideas about ST and strategic procedures used by students in the pre-test ST.
	Session 2 (60 min. February, week 4). Set suitable goals for doing STs	Teacher 1: group 4 Teacher 2: group 3	- Students set individual goals related to STs in terms of product and the cognitive processes required. - Using the self-monitoring sheet from session 1 as a starting point, students make their own self-monitoring sheet; it includes the individual goals about product quality and assessing it to facilitate self-regulation for the process of producing STs. - Pool ideas
Phase 2	Session 3 (60 min. March, week 2). Develop understanding of strategies for producing synthesis and examine efficacy of students' own strategies	Teacher 1: group 3 Teacher 2: group 4	- The teacher analyses and explains examples of effective strategies for synthesis. - The teacher presents correct ST analysis, looking at product quality and the analysis of strategic processes used to do synthesis using the self-monitoring sheet, to which they add assessment of the activated reading and writing processes and strategies. - As a group, students and teachers recognize effective and ineffective aspects of students' own strategies (compare model ST and procedure and students' own STs).
	Session 4 (60 min. March, week 4). Define students' own strategies	Teacher 1: group 4 Teacher 2: group 3	- Students reflect on their (pre-test) STs and examine whether they met the criteria of a quality synthesis according to the teachers' self-monitoring sheet. They identify what fell short in the product and activation of reading and writing strategies and processes. - Students set individual and definitive goals related to STs in terms of product and the cognitive processes required. - Students define their own strategies and adjust their own self-monitoring sheet, with the teacher's help. - Pool ideas.
Phase 3	Session 5 (60 min. April, week 2). Learn types of feedback and how to apply them. Produce self-reinforcement.	Teacher 1: group 3 Teacher 2: group 4	- The teacher presents types of feedback and models applying them and strategies to STs. - Students define self-reinforcement.
	Session 6 (60 min. April, week 4). Apply students' own strategies to peer texts. Evaluate their efficacy and redefine them.	Teacher 1: group 4 Teacher 2: group 3	- Students do an ST task in pairs: One student does the ST task applying the strategies and thinking aloud. The other follows the whole process of creating the synthesis, playing an active role and participating in the process, analyzing possible mistakes and offering their own ideas about strategies and help to guide or re-direct the writing process as needed. - Students analyse the synthesis applying their own self-monitoring sheets. - Individually, students evaluate the efficacy of their own strategies and redefine them according to their weaknesses. - Teachers support students during session.
	Session 7 (60 min. May, week 2). Apply students' own strategies and self-reinforcement to individual texts.	Teacher 1: group 3 Teacher 2: group 4	- Students write individual STs, applying their own strategies and self-reinforcement, and evaluate efficacy with the self-monitoring sheet. - As a group, students and teachers read and evaluate ST quality.
	Session 8 (60 min. May, week 4). Evaluate strategic efficacy and define it definitively.	Teacher 1: group 4 Teacher 2: group 3	- Students compare the ST from session 7 with the pre-test, identifying the differences in the product and in the process and strategies, using the self-monitoring sheets. - Final evaluation of students' own strategies and adjustment. - Review and general reflections on generalizing strategies.

TABLE 3 Summary of the intervention for the control condition.

Session	Focus	Instructor	Activities
1	Identify students' own ideas about synthesis and its requirements (60 min. February, week 2).	Teacher 1: group 1 Teacher 2: group 2	■ Students describe perceived requirements of syntheses After producing an ST (pre-test), students analyse the ST based on students' ideas ■ Teacher explains synthesis and its requirements: what are synthesis tasks and their characteristics in terms of selection, connection, and organization. They also explain the demands of STs at the level of reading and writing processes.
2	Identify the procedure for writing a synthesis and its relationship with the type of representation task (60 min. February, week 4).	Teacher 1: group 2 Teacher 2: group 1	■ Students describe the procedure for crafting a synthesis (in pre-test). ■ Individual written reflection about the efficacy of the procedure followed, clarify task requirements
3	Analyse text types, developing declarative knowledge: definition, characteristics, structure, and discourse strategies (60 min. March, week 2).	Teacher 1: group 1 Teacher 2: group 2	■ Individual reading about the document entitled "Los tipos de texto en español" [<i>Types of texts in Spanish</i>] and individual identification of: definition, characteristics, internal structure and discourse strategies for descriptive, narrative, explanatory and argumentative texts. Make a content table. ■ Pool ideas as a group
4	Look at different types of text in the university academic-scientific context (60 min. March, week 4).	Teacher 1: group 2 Teacher 2: group 1	■ Individual reading of document "La escritura académica en la Universidad" [<i>Academic writing at university</i>]: students identify and summarize academic discourse characteristics, text types, characteristics ■ Pool ideas together
5–6	Apply knowledge of text structures to a real text in pairs (60 min. April, weeks 2 and 4).	Teacher 1: group 1 Teacher 2: group 2	■ Review previous content, characteristics of a synthesis ■ Students write a synthesis in pairs. Students do an ST task in pairs: One student produces the ST applying knowledge about this type of task (characteristics, structure, discourse strategies) while thinking aloud. The other student follows the process, playing an active role and participating, analysing possible mistakes, making their own contributions and offering help and redirection to the process as needed. ■ As a group, students and teacher analyse the resulting synthesis, only modifying the structure (session 6)
7	Apply knowledge of synthesis structure and characteristics to a real text individually (60 min. May, week 2).	Teacher 1: group 2 Teacher 2: group 1	■ The teachers review characteristics of syntheses (presented in session 1) ■ Students write a synthesis individually, applying knowledge about ST characteristics, structure, discourse strategies.
8	Evaluate efficacy of knowledge gained about synthesis (60 min. May, week 4).	Teacher 1: group 1 Teacher 2: group 2	■ As a group, students and teachers read the syntheses from session 7 and analyse it based on structural characteristics. ■ Students compare the text from session 7 to the pre-test and evaluate the text quality. ■ General review, rate efficacy of the program

The results confirmed this hypothesis, and with it, the efficacy of the strategic, self-regulated instruction program for promoting improvements in the quality of university students' synthesis texts, and in the activation of reading and writing strategies. Compared to the control group, the students in the experimental group demonstrated greater improvements in the overall quality of their synthesis texts, producing work with connected, hierarchical ideas and details, demonstrating clear sequences of ideas and fluid discourse. Their texts also exhibited appropriate text organization, with paragraph structures for introductions, development and conclusions that were suitably cohesive. In addition, their texts displayed suitable levels of vocabulary, syntax, spelling, and presentation.

Following the program, the experimental group also exhibited greater improvements than the control group in effective activation of reading strategies, such as taking notes while reading. They also showed greater improvement in writing strategies—especially those related to planning their texts—considering structural aspects of the

text when planning, such as the introduction and conclusion, and explicitly demonstrating the use of strategies. The greater improvement in the experimental group's text quality was maintained over the long term, shown by the follow-up assessment one year later, although it is worth emphasizing that both groups demonstrated improvements in text product scores at the follow-up evaluation. This shows some consolidation of synthesis competencies in university students, presumably driven by their progress in their degree courses. These are courses that generally use active methodologies requiring increasingly active, strategic, and self-regulated learning from the student (Robledo et al., 2015; García et al., 2017).

Hence, in light of the scientific literature in this field, our study confirms that—with proper guidance from teachers via implementing strategic instructional processes (Graham et al., 2018; Graham and Harris, 2018) within the context of the curriculum study plan (Zhang, 2013)—students can adapt and re-work their own strategies (Butler, 1994, 1995, 1996, 1997, 1998), in this case for

TABLE 4 Kurtosis and skewness.

Dimension	Measure	Pre-test		Post-test		Follow-up	
		Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Selection of information	Idea selection	0.279	−0.055	0.803	2.767	−0.734	0.279
	Informative capacity	0.317	−1.968	−1.451	3.740	−0.771	0.728
Connecting ideas	Cohesion	−0.540	−1.770	−1.641	6.955	0.111	−0.091
	Plagiarism	0.229	−0.667	0.551	2.583	0.872	0.076
Text organization	Overall structure	0.828	1.860	0.123	−1.191	−1.294	0.291
	Meta-structural	2.011	2.114	1.322	1.317	1.395	1.134
	Structural	1.393	2.776	0.970	0.399	1.238	0.918
	Reformulating	0.369	−0.569	1.231	2.154	0.357	−0.570
	Argumentative	2.636	8.291	1.706	2.983	1.507	2.466
	Local structure	0.777	0.477	1.334	1.551	0.217	−1.205
Quality	Overall quality	0.335	−0.172	0.262	−0.625	0.382	−1.080
Reading strategies	General read. Strateg.	−1.919	1.807	−3.093	7.826	−0.132	−2.073
	Underlining	−3.408	12.091	−2.213	2.996	−1.813	1.344
	N° underlining texts	−7.746	60.000	.	.	.	.
	Type of notes	−0.193	1.448	0.018	−0.084	0.557	−0.992
	N° of texts with notes	−1.382	0.289	−2.725	5.919	0.218	−1.807
Planning	Overall Plan.	−1.458	0.684	−2.518	5.228	0.235	−1.120
	Plan. Ideas	−0.225	−1.034	−0.100	−0.660	0.764	−1.108
	Introduction	.	.	1.195	0.279	−0.569	−0.583
	Development	−1.828	2.407	−1.877	2.645	1.811	1.827
	Conclusion	4.841	22.331	2.011	3.439	.	.
	Audience	.	.	7.483	56.000	6.708	45.000
	Objective	.	.	.	.	6.708	45.000
	Strategies	4.841	22.331	2.098	2.489	0.962	0.430
Revision							
	Mechanical	1.477	1.824	1.409	2.018	1.945	6.202
	Substantive	2.912	10.733	1.209	0.760	2.419	7.475
	Total	2.184	5.596	0.907	0.427	1.184	1.628

reading and writing, and apply them in a self-regulated manner when doing complex academic tasks, such as synthesis, improving the quality of their texts and of some of the cognitive processes involved (Boscolo et al., 2007; Mateos et al., 2018; Granado-Peinado et al., 2019). Despite that, the program implemented in this study did not contribute to any significant improvements in the processes of selecting information required by ST and related to selecting ideas from source texts. These processes seem to be among the most well-developed in university students (Robledo et al., 2018), which is why the range of improvement is smaller. The program also failed to produce significant improvements in processes for connecting information, including cohesion or incorporating ideas from source texts into the ST. These processes for connecting information require students to deeply read the source texts in combination, which in this case are complex academic texts. They also involve rewriting and connecting the ideas taken from the texts into a new written discourse. This may represent an additional challenge for students who, faced with the complexity of the materials they have to read,

only engage with them superficially and fail to achieve the in-depth understanding that would allow them to write about the new ideas (Beck et al., 1996; van Ockenburg et al., 2019). Processes related to reviewing writing also failed to show any significant improvement, possibly because students activated them when they were cognitively drained after having focused their efforts on detailed planning or even after drafting the text. Students need greater cognitive flexibility and to recursively activate all of the cognitive processes involved in producing STs to achieve optimum levels of functionality (Mateos et al., 2008). Finally, the lack of a proper metacognitive representation of synthesis tasks often conceptualized by students as summaries (Mateos and Solé, 2009; Robledo et al., 2018), may also explain the poor activation of some core processes for these tasks. In any case, it is clear that our results call for future studies that seek deeper explanations and that can provide more specific guidance to teachers in terms of functional activation of all of the processes involved in producing synthesis texts through effective development and deployment of specific strategies linked to each process and through

TABLE 5 Statistically significant differences between groups in the comparison between pre-test, post-test, and follow-up in the text product measures.

	Experimental						Control						Pre-post			Pre-post-follow		
													Time*condition			Time*condition		
	Pre		Post		Follow-up		Pre		Post		Follow-up		F	P-value	η^2	F	p-value	η^2
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD						
Overall structure	2.12	0.59	2.62	0.55	3.65	0.84	2.27	0.45	2.08	0.27	2.92	1.13	3.361	0.072	0.055	6.293	0.016	0.116
Meta-structural	0.21	0.41	0.82	0.79	1.12	1.50	0.08	0.27	0.12	0.32	0.88	1.29	15.351	<0.001	0.209	8.707	0.005	0.154
Reformulating	2.59	1.72	4.68	2.91	5.77	3.46	2.42	1.79	2.58	1.81	4.96	2.97	6.778	0.012	0.105	4.421	0.041	0.084
Quality	1.53	0.74	2.74	0.75	3.88	0.95	1.62	0.57	1.92	0.68	2.57	0.72	4.949	0.030	0.079	18.105	0.000	0.278

TABLE 6 Differences between pre- and post-test in reading and writing strategies in each condition.

	Pre-test		Post-test		Z	p-value
	M	SD	M	SD		
Experimental						
General reading strategy	2.56	0.824	2.88	0.478	−1.960	0.050
Type of notes	0.88	0.537	1.35	0.597	−3.578	0.000
N° of texts with notes	1.41	0.821	1.82	0.521	−2.636	0.008
Overall planning	1.53	0.825	1.85	0.436	−2.271	0.023
Planning ideas	1.65	1.098	2.00	0.739	−2.109	0.035
Introduction	1.00	0.000	1.73	0.719	−3.542	0.000
Conclusion	1.04	0.192	1.36	0.549	−2.309	0.021
Strategy	1.15	0.534	1.48	0.870	−2.449	0.014
Mechanic rev.	0.76	0.955	1.74	2.020	−2.935	0.003
Total rev.	2.56	3.518	3.91	3.995	−2.609	0.009
Control						
Substantive rev.	0.92	1.521	2.35	2.682	−2.646	0.008
Total rev.	2.15	2.679	4.04	3.736	−2.503	0.012

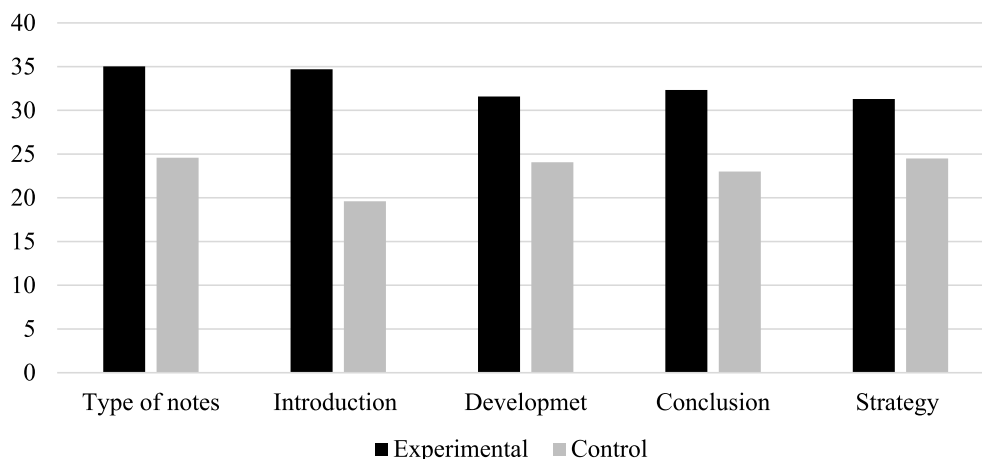


FIGURE 1

Mean ranks for both groups in the variables with significant differences between condition following the program (post-test).

spending more time on their practical application (Merrill, 2002; Van Ockenburg et al., 2019).

Despite that, we can conclude that instruction based on stimulating students' own strategies and encouraging self-regulated application of them in synthesis tasks is effective and contributes to students producing better quality texts, organizing information better, and more effectively activating some of the reading and writing processes these types of tasks call for. This may be because this type of instruction allows students more control over their cognitive and emotional processes, meaning that they approach these tasks with greater motivation and are able to keep at them and complete them successfully, even with the challenges they present (Butler, 1998; Navea-Martín and Suárez-Riveiro, 2017; van Ockenburg et al., 2019). To sum up from a theoretical-empirical perspective, we can conclude that implementation of strategic instruction programs aimed at encouraging ST mastery at the university level should consider how strategically mature the students are and their potential to create and modify their own strategies (Kieft et al., 2007; Robledo et al., 2018; Arias-Gundín and Robledo, 2023). Teachers must focus on guiding and prompting students so that they develop and implement their own strategies in a self-regulated manner when doing ST. Therefore, it seems that in the university context, sequences of strategic instruction programs can be simplified, as can their component parts. Students, with the teacher's guidance, can develop their own strategies and apply them appropriately to perform synthesis tasks.

Nonetheless, these conclusions should be understood considering the limitations of the study. One limitation is the small sample size, which was taken from a specific course that was part of one degree at one university. This limits the possibilities of generalizing the results, and means that it would be interesting for future studies to use a larger sample and also to do more complex statistical analysis, such as mixed-effects models, with observations grouped within individual students and students nested within classes, given that the present study is not immune to the clustering effect. This was an early pilot study in this field, which used a similar sized sample to other international studies in the same area (reviewed in the introduction). In addition, having 20–25 students in a class was ideal for learning. Given that, more studies, and more robust statistical analyses, are

needed to test whether synthesis writing instruction is effective for a wider range of students. It would also be useful to consider other types of intrinsic student variables that might affect the results. In this case, because the instruction was based on the students readjusting their own strategies, it would be useful to determine a general strategic profile for the students and a profile of their reading and writing strategies, which seem to vary between students and affect how they face learning activities in general and written tasks in particular (Kieft et al., 2007; van der Loo et al., 2018; Arias-Gundín and Robledo, 2023). In addition, it would be interesting to use more evaluation instruments, including ones which would allow measurement of qualitative changes in critical writing skills, which would guide interventions more accurately (Deane and Song, 2014). Another limitation to highlight is that the teachers tasked with implementing the program had experience in teaching written composition. This makes it difficult to say whether other teachers with less expertise in teaching processes would be able to apply the program with similar success. In line with this, on a practical note, it is worth emphasizing that the instructional program tested in this study could be applied in different subjects and may be an interesting methodological tool for teachers because, in addition to helping improve students' syntheses, it provides an active method to access content for various learning goals (van Ockenburg et al., 2019), facilitating students' autonomous learning and skill development.

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 humans were approved by Comité de ética University of León for the studies involving humans. The studies were conducted in accordance with the local legislation and institutional requirements. All students agreed to participate voluntarily after

verbal informed consent was obtained. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PR: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MÁ: Conceptualization, Funding acquisition, Resources, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was partially supported by University of León.

References

- Alonso-Tapia, J. (2012). Teorías de autorregulación educativa: una comparación y reflexión teórica. *Psic. Educ.* 20, 11–22. doi: 10.1016/j.pse.2014.05.002
- Arias-Gundín, O., and Robledo, P. (2023). An analysis of the psychometric properties of the writing-specific cognitive strategies questionnaire for undergraduate students. *Front. Psychol.* 14:1274478. doi: 10.3389/fpsyg.2023.1274478
- Beck, I., McKeown, M., Sandora, C., Kucan, L., and Worthy, J. (1996). Questioning the author: a yearlong classroom implementation to engage students with text. *Elem. Sch. J.* 96, 385–414. doi: 10.1086/461835
- Boscolo, P., Arfe, B., and Quarisa, M. (2007). Improving the quality of students' academic writing: an intervention study. *Stud. Hig. Educ.* 32, 419–438. doi: 10.1080/03075070701476092
- Butler, D. (1992). Promoting strategic learning by learning disabled adults and adolescents. *Excep. Educ. Can.* 2, 109–128.
- Butler, D. (1994). From learning strategies to strategic learning: promoting self-regulated learning by post-secondary students with learning disabilities. *Can. J. Spec. Educ.* 4, 69–101.
- Butler, D. (1995). Promoting strategic learning by post-secondary students with learning disabilities. *J. Lear. Disab.* 28, 170–190. doi: 10.1177/002221949502800306
- Butler, D. (1996). The strategic content learning approach to promoting self-regulated learning. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Butler, D. (1997). The roles of goal setting and self-monitoring in students' self-regulated engagement in tasks. Paper presented in the annual meeting of the American Educational Research Association, Chicago.
- Butler, D. (1998). The strategic content learning approach to promoting self-regulated learning: a report of three studies. *J. Educ. Psych.* 90, 682–697. doi: 10.1037/0022-0663.90.4.682
- Cumming, A., Lai, C., and Cho, H. (2016). Students' writing from sources for academic purposes: a synthesis of recent research. *J. Eng. Acad. Purp.* 23, 47–58. doi: 10.1016/j.jeap.2016.06.002
- Deane, P., and Song, Y. (2014). A case study in principled assessment design: designing assessments to measure and support the development of argumentative reading and writing skills. *Psic. Educ.* 20, 99–108. doi: 10.1016/j.pse.2014.10.001
- Du, H., and List, A. (2020). Researching and writing based on multiple texts. *Learn. Instr.* 66, 1–25. doi: 10.1016/j.learninstru.2019.101297
- Flower, L., and Hayes, J. R. (1980). "The dynamics of composing: making plans and juggling constraints" in Cognitive processes in writing. eds. L. W. Gregg and E. R. Steinberg (New Jersey: Lawrence Erlbaum Associates), 31–49.
- Flower, L., Stein, V., Ackerman, J., Kantz, M. J., McCormick, K., and Peck, W. C. (1990). Reading to write: Exploring a cognitive and social process. New York: Oxford University Press.
- García, T., Arias, O., Rodríguez, C., Fidalgo, R., and Robledo, P. (2017). Metodologías activas y desarrollo de competencias en estudiantes universitarios con diferentes estilos de pensamiento. *Rev. d'Inn. Doc. Univ.* 9, 66–80. doi: 10.1344/RIDU2017.9.6
- Graham, S., and Harris, K. R. (2018). "Evidence-based writing practices: A meta-analysis of existing meta-analysis" in Design principles for teaching effective writing: Theoretical and empirical grounded principles. eds. R. Fidalgo, K. R. Harris and M. Braaksma (Leiden: Brill Editions), 13–37.
- Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., et al. (2018). Effectiveness of literacy programs balancing reading and writing instruction: a meta-analysis. *Read. Res. Quar.* 53, 279–304. doi: 10.1002/rq.194
- 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. doi: 10.1007/s11145-019-09939-6
- Granello, D. (2001). Promoting cognitive complexity in graduate written work: using Bloom's taxonomy as a pedagogical tool to improve literature reviews. *Coun. Educ. Super.* 40, 292–307. doi: 10.1002/j.1556-6978.2001.tb01261.x
- Harris, K. R., Camping, A., and McKeown, D. (2022). "A review of research on professional development for multicomponent strategy-focused writing instruction: knowledge gained and challenges remaining" in Conceptualizing, designing, implementing, and evaluating writing interventions. eds. F. DeSmedt, R. Bouwer, T. Limpo and S. Graham (Leiden: Brill).
- Kieft, M., Rijlaarsdam, G., Galbraith, D., and Bergh, H. (2007). The effects of adapting a writing course to students' writing strategies. *Br. J. Educ. Psychol.* 77, 565–578. doi: 10.1348/096317906X120231
- Martínez, I., Martín, E., and Mateos, M. (2011). Enseñar a leer y escribir para aprender en la Educación Primaria. *Cult. Educ.* 23, 399–414. doi: 10.1174/113564011797330306
- Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., and González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cogn. Instr.* 36, 119–138. doi: 10.1080/07370008.2018.1425300
- Mateos, M., Martín, E., Villalón, R., and Luna, M. (2008). Reading and writing to learn in secondary education: online processing activity and written products in summarizing and synthesizing tasks. *Read. Writ.* 21, 675–697. doi: 10.1007/s11145-007-9086-6
- Mateos, M., and Solé, I. (2009). Synthesizing information from various texts: a study of procedures and products at different educational levels. *Eur. J. Psychol. Educ.* 24, 435–451. doi: 10.1007/BF03178760
- McArthur, C. A., and Philippakos, Z. A. (2013). Self-regulated strategy instruction in developmental writing: a design research project. *Community Coll. Rev.* 41, 176–195. doi: 10.1177/0091552113484580
- McArthur, C. A., and Philippakos, Z. A. (2022). "Supporting Students' writing from sources at college entry" in Handbook of Research on Writing Instruction Practices for Equitable and Effective Teaching. ed. T. S. Hodge (IGI Global), 338–358.

Acknowledgments

We would like to thank students who participated voluntarily in the studies, and the researchers who rated all measures under blind conditions.

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.

- McArthur, C., Philippakos, Z. A., and Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *J. Educ. Psych.* 107, 855–867. doi: 10.1037/edu0000011
- McArthur, C. A., Philippakos, Z. A., May, H., and Compello, J. (2022). Strategy instruction with self-regulation in college developmental writing courses: results from a randomized experiment. *J. Educ. Psychol.* 114, 815–832. doi: 10.1037/edu0000705
- Merrill, M. (2002). First principles of instruction. *Educ. Technol. Res. Dev.* 50, 43–59. doi: 10.1007/BF02505024
- Nadal, E., Miras, M., Castells, N., and De la Paz, S. (2021). Intervención en escritura de síntesis a partir de fuentes. Impacto de la comprensión. *Rev. Mex. Invest. Educ.* 26, 95–111.
- Navea-Martín, A., and Suárez-Riveiro, J. M. (2017). Estudio sobre la utilización de estrategias de automotivación en estudiantes universitarios. *Psic. Educ.* 23, 115–121. doi: 10.1016/j.pse.2016.08.001
- Perin, D. (2013). Literacy skills among academically underprepared students. *Community Coll. Rev.* 41, 118–136. doi: 10.1177/0091552113484057
- Robledo, P., Arias-Gundín, O., Palomo, M. P., Andina, E., and Rodríguez, C. (2018). Perfil escritor y conocimiento metacognitivo de las tareas académicas en los estudiantes universitarios. *Rev. Publ.* 48, 197–217. doi: 10.30827/publicaciones.v48i1.7335
- Robledo, P., Fidalgo, R., Arias, O., and Álvarez, M. L. (2015). Percepción de los estudiantes sobre el desarrollo de competencias a través de diferentes metodologías activas. *Rev. Invest. Educ.* 33, 369–383. doi: 10.6018/rie.33.2.201381
- Robledo, P., and García, J. N. (2017). “Description and analysis of strategy-focused instructional models for writing” in Design principles for teaching effective writing. eds. R. Fidalgo, K. Harris and M. Braaksma (Leiden: Brill), 38–65. doi: 10.1163/9789004270480_004
- Roldán, C., Vázquez, A., and Rivarosa, A. (2011). Mirar la escritura en la educación superior como un prisma. *Rev. Iber. Educ.* 55, 1–13. doi: 10.35362/rie5531596
- Rosário, P., Núñez, J. C., Rodríguez, C., Cerezo, R., Fernández, E., Tuero, E., et al (2017). “Analysis of Instructional Programs in Different Academic Levels for Improving Self-Regulated Learning srl through Written Text” in Studies in Writing Series for Teaching Effective Writing. In: R. Fi-dalgo, and K. R. Harris (Series Eds), R. Fi-dalgo, K. Harris and M. Braaksma (Vol. Eds.). (pp. 201–230). Leiden: Brill.
- Segev-Miller, R. (2004). Writing from sources: the effect of explicit instruction on college students’ processes and products. *L1-Educ. Stud. Lang. Liter.* 4, 5–33. doi: 10.1023/B:ESLL.0000033847.00732.af
- Segev-Miller, R. (2007). “Cognitive processes in discourse synthesis: the case of intertextual processing strategies” In Studies in writing. Writing and cognition: research and applications, ed. G. Rijlaarsdam, M. Torrance, WaesL. Van and D. Galbraith (The Netherland: Elsevier), 231–250.
- Solé, I., Mateos, M., Miras, M., Martín, E., Castells, N., Cuevas, I., et al. (2005). Lectura, escritura y adquisición de conocimientos en Educación Secundaria y Educación Universitaria. *Inf. Apr.* 28, 329–347. doi: 10.1174/0210370054740241
- Solé, I., Miras, M., Castells, N., Espino, S., and Minguela, M. (2013). Integrating information: an analysis of the processes involved and the products generated in a written synthesis task. *Writ. Commun.* 30, 63–90. doi: 10.1177/0741088312466532
- Spivey, N. (1997). The constructivist metaphor: Reading, writing, and the making of meaning. San Diego: Academic Press.
- Spivey, N., and King, J. (1989). Readers as writers composing from sources. *Read. Res. Quart.* 24, 7–26. doi: 10.1598/RRQ.24.1.1
- Van der Loo, J., Krahmer, E., and Van Amelsvoort, M. (2018). Learning how to write an academic text: the effect of instructional method and writing preference on academic writing performance. *J. Writ. Res.* 9, 233–258. doi: 10.17239/jowr-2018.09.03.01
- Van Ockenburg, L., van Weijen, D., and Rijlaardam, G. (2019). Learning to writing synthesis texts: a review of intervention studies. *J. Writ. Res.* 10, 401–428. doi: 10.17239/jowr-2019.10.03.01
- Zhang, C. (2013). Effect of instruction on ESL students’ synthesis writing. *J. Sec. Lang. Writ.* 22, 51–67. doi: 10.1016/j.jslw.2012.12.001



OPEN ACCESS

EDITED BY

Linda Saraiva,
Polytechnic Institute of Viana do Castelo,
Portugal

REVIEWED BY

Shadi Dehghanzadeh,
Islamic Azad University, Rasht Branch, Iran
Allison Mari Dering-Anderson,
University of Nebraska Medical Center,
United States

*CORRESPONDENCE

Chaya Gopalan
✉ cgpala@siue.edu

RECEIVED 22 February 2024

ACCEPTED 11 July 2024

PUBLISHED 25 July 2024

CITATION

Gopalan C, Bruno EL, Daugherty SE and
Nguyen K-LT (2024) Flipping the anatomy
classroom: a comparative analysis of 16-week
and 8-week courses in a community college.
Front. Educ. 9:1389825.
doi: 10.3389/feduc.2024.1389825

COPYRIGHT

© 2024 Gopalan, Bruno, Daugherty and
Nguyen. 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.

Flipping the anatomy classroom: a comparative analysis of 16-week and 8-week courses in a community college

Chaya Gopalan^{1,2*}, Elizabeth L. Bruno³, Sheyenne E. Daugherty⁴
and Kim-Leiloni T. Nguyen⁵

¹Department of Exercise, Sport, Nutrition Sciences, School of Education, Health and Human Behavior, Southern Illinois University Edwardsville, Edwardsville, IL, United States, ²Department of Nurse Anesthesiology, School of Nursing, Southern Illinois University Edwardsville, Edwardsville, IL, United States, ³Psychology Department, Southern Illinois University Edwardsville, Edwardsville, IL, United States, ⁴Biology Department, Southern Illinois University Edwardsville, Edwardsville, IL, United States, ⁵Division of Natural Sciences, Department of Biology, Mt. San Antonio College, Walnut, CA, United States

Using flipped teaching modality, a student-centered approach, in community college courses remains an understudied area. This study explored the effectiveness of unflipped (UF), partially flipped (PF), and fully flipped (FF) teaching formats within the confines of an accelerated 8-week Introduction to Human Anatomy course at a community college. The purpose of this study was to examine flipped teaching in an accelerated anatomy course by comparing final examination scores between the UF, PF, and FF modalities, compare the effectiveness of flipped teaching between a regular (16-week) and an accelerated anatomy course, evaluate the effect of different teaching modalities on male and female students between a regular and an accelerated course; and examine UF, PF, and FF among the above-median and the below-median students. Students in the FF sections were required to read relevant chapter(s) in the textbook and review slides before class. Students in the PF sections were required to read the textbook and slides for 25% of the chapters. All students took the same final exams and similar section exams. Final exam scores were higher in the regular 16-week course than in the 8-week course for UF ($p = 0.0219$) and PF ($p = 0.0183$) modalities. The 16-week course had higher final examination scores ($p = 0.0492$; $n = 65$) than the 8-week course. Male students scored lower in the 8-week FF course (49.42 ± 13.72) than in the 8-week PF ($p = 0.006$) and 16-week FF ($p = 0.0008$) formats and also compared to female students in the 8-week FF course ($p = 0.0121$). Above-median students in the 8-week course had significantly lower scores in the FF modality (71.35 ± 7.01) compared to PF (80.92 ± 5.30) ($p = 0.009$). Below-median students in the 8-week course had lower scores in the FF modality (52.25 ± 11.48) compared to UF ($p = 0.0113$) and PF ($p = 0.04$). In conclusion, FF in a 16-week anatomy course resulted in higher final exam scores than in the 8-week course. The 8-week FF format affected male student scores. Both above- and below-median groups experienced similar effects when exposed to FF within the accelerated course. Further investigations are essential to inform the refinement of flipped teaching methodologies in community colleges, particularly in the context of accelerated courses.

KEYWORDS

gross anatomy education, undergraduate education, flipped classroom, student success, course duration

1 Introduction

Evidence-based pedagogies for student engagement and learning are crucial in supporting community colleges. Active learning strategies, such as group discussions, problem-solving activities, and peer teaching, have been found to improve student engagement and knowledge retention (Freeman et al., 2014; Faucette, 2023; Loyens et al., 2023). The flipped classroom model is one such approach that embodies active learning and enhances student achievement and engagement (Talbert, 2017; Dehghanzadeh and Jafaraghaee, 2018; Doğan et al., 2023).

The flipped teaching method is an instructional approach that reimagines the traditional classroom setup by relocating lectures outside the classroom. Instead, in-class time is reallocated to foster student-centered activities, thereby facilitating a more profound comprehension of the subject matter (Al-Samarraie et al., 2020; Strelan et al., 2020; Özbay and Çınar, 2021; Baig and Yadegaridehkordi, 2023). The flipped classroom consists of two fundamental components: pre- and in-class phases. During the pre-class stage, students acquire foundational knowledge through pre-recorded lecture videos and assigned readings. This phase is pivotal for establishing the fundamental concepts that serve as a springboard for the subsequent in-class activities. In contrast, the in-class segment incorporates a diverse range of activities, including succinct review sessions in the form of mini-lectures, self-assessment exercises, valuable instructor feedback, and a deliberate emphasis on delving into advanced content (Lo and Hew, 2019; Al-Samarraie et al., 2020). These in-class activities are designed to actively engage students and promote a deeper understanding of the subject (Subari, 2017; Talbert, 2017; Cevikbas and Kaiser, 2020; Gopalan et al., 2020; Mojtahedi et al., 2020; Namaziandest et al., 2020; Kapur et al., 2022).

The evidence supporting the efficacy of flipped teaching over traditional methods continues to grow (Pierce and Fox, 2012; Freeman et al., 2014; Ozudogru and Aksu, 2020; Jdaitawi, 2021). By transferring lecture content to independent learning spaces outside the classroom, students are expected to assume greater responsibility for mastering fundamental concepts. This shift provides flexibility and enhances self-efficacy (Herreid and Schiller, 2013; Tune et al., 2013). As a result of this relocation of lectures, class time can be capitalized toward fostering peer and instructor interactions and facilitating formative assessments. Consequently, students gain more opportunities to engage in in-depth discussions and are repeatedly exposed to the course material. The incorporation of individual and group assessments allows students to test and apply their knowledge, which in turn contributes to improved learning outcomes and test scores (Li et al., 2020; Wei et al., 2020; Halpin and Gopalan, 2021; Sivarajan et al., 2021).

One common hurdle in adopting flipped teaching is students' initial resistance to this new modality (Lo and Hew, 2019). Due to the COVID-19 pandemic, this resistance has decreased somewhat (Asad and Srivastava, 2020). Flipped teaching demands a consistent time commitment throughout the semester, encouraging students to avoid cramming for major exams. Additionally, student motivation can be challenging because flipped instruction requires more preparation for class and participation in formative assessments compared to the traditional lecture method (Gopalan et al., 2020).

The outbreak of COVID-19 precipitated an abrupt shift to online teaching for instructors worldwide, with some opting for the flipped

teaching approach (Morton, 2020; Roy et al., 2020; Yoo et al., 2020; McWatt, 2021). The adoption of flipped instruction during this transition was likely due to its ability to allow students to review lecture recordings at their own pace in an asynchronous format. This flexibility enabled students to schedule their learning around external commitments and minimize distractions (Dove, 2013; McWatt, 2021; Riedl et al., 2021).

Faculty encounter two primary concerns when implementing flipped teaching: the time investment and student evaluations. The initial stages of adopting flipped instruction demand a substantial time commitment, often necessitating a temporary shift from other responsibilities including scholarship and service. Furthermore, course evaluations may fluctuate since flipped teaching represents a novel approach for many students (Hornstein, 2017). Instructors may oppose this teaching method due to unfamiliarity with its implementation and the initial preparation required (Bove and Conklin, 2019). To address the challenge of transitioning an entire curriculum at once, instructors can begin by flipping a small portion of the content, known as a partial flip (PF). As they develop more flipped content, they can eventually apply this method to the entire course, referred to as a full flip (FF) (Kostka and Marshall, 2018). In contrast, the traditional lecture method will henceforth be referred to as 'unflipped' (UF).

Notably, there is a paucity of evidence regarding the effectiveness of flipped teaching at community colleges, particularly in the Science, Technology, Engineering, and Mathematics (STEM) fields. Community colleges are known for their diverse student populations, with individuals of various ages, ethnicities, and backgrounds as well as additional responsibilities such as employment and family commitments. This diversity can enrich the learning environment, but it also presents challenges in adapting teaching methods to cater to a wide array of needs and experiences (Karp et al., 2010; Dougherty, 2016). Community colleges often serve students with diverse levels of educational preparedness. Some students possess a strong academic foundation, while others may require remedial coursework (Jenkins et al., 2014). Many community college students face financial constraints and often juggle work, family, and college, which could result in an increased workload, affecting their ability to dedicate time to coursework (Twigg, 2009; Shapiro et al., 2015). Addressing these varying needs while maintaining academic rigor is a unique challenge. Given its flexibility, flipped teaching is expected to be well-suited for community college students (Dove, 2013; Riedl et al., 2021). Students will be able to access pre-class content at their available times and as frequently as it takes to become familiar with the content. Active learning strategies employed during class time to engage students would reinforce their understanding of the subject further. A comprehensive grasp of student learning preferences is indispensable for optimizing anatomy instruction (Cullinane and Barry, 2023). The investigation of these preferences remains an ongoing endeavor, as they can vary due to a spectrum of factors encompassing cultural backgrounds, individuality, age, and academic preparation (Holtbrügge and Mohr, 2010). Central to this concept is the modality preference, which underscores the significance of employing an individual's preferred learning approach to enhance educational outcomes (Lodge et al., 2016). Flipped teaching allows students to be exposed to the content early on and this priming is beneficial in learning the content at a deeper level. Active learning strategies employed during class time to engage students would reinforce their

understanding of the subject further. Teachers being able to spend more time with students in the flipped teaching modality will be able to assess students' abilities and will use appropriate tools to enhance student engagement. Moreover, students are exposed to content multiple times and go through several formative assessments in the flipped pedagogy which may help them determine their learning styles.

Research on learning styles based on gender has yielded inconclusive results. While Idrizi et al. (2023) suggested gender differences in learning styles, with females preferring read/write and males favoring kinesthetic styles, another study by Voyer and Voyer (2014) highlighted a female advantage in learning. These disparate findings challenge stereotypical beliefs that females excel in language and males in math and science (Halpern et al., 2011). This study explored whether flipped teaching had any effect particularly on male versus female students.

The existing evidence on the benefits of flipped teaching for students who are earning better grades versus those barely passing is inconclusive (Wozny et al., 2018; Gopalan et al., 2020; Herrero and Quiroga, 2020). Gopalan et al. observed improvement in both upper and lower-median groups, while Herrero & Quiroga's findings suggested that below-median students derived greater benefits from this instructional approach. Conversely, Wozny et al. reported positive impacts on above-median students (Wozny et al., 2018). The conflicting results highlight the ambiguity surrounding the effects of flipped teaching on distinct student groups within a course. Acknowledging this uncertainty, our study aimed to contribute to the understanding by investigating the impact of flipped teaching on students performing above and below the median levels in their classes. This examination encompassed three instructional modalities: UF, partially PF, and FF. Through this exploration, we aimed to provide additional insights into the varied effects of flipped teaching across different student performance brackets.

Students have many course options to enroll in, such as accelerated programs, expedited degree completion, and/or accommodation for schedule conflicts especially at the community college level. Eames et al. (2018) studied examination scores between an accelerated and traditional accountant program and found no significant difference between the two. Thus, a well-designed accelerated course can increase the quantity of well-qualified accountants (Eames et al., 2018). There are reports indicating increased stress levels when the course length is shortened (Green, 2020).

1.1 Study goals

The objectives of this study were to: (1) examine flipped teaching at a community college in an accelerated (8-week) anatomy course by comparing final examination scores between the UF, PF, and FF modalities; (2) compare the effectiveness of flipped teaching in the form of final examination scores between a regular (16-week) and an accelerated anatomy course at a community college setting; (3) evaluate the effect of different modalities on male and female students between a regular and an accelerated anatomy course; and (4) examine the three teaching modalities, UF, PF, and FF, among the above-median and the below-median students. It was hypothesized that students experiencing FF in the 16-week course would perform better than their peers in the 8-week course since there is a greater ratio of time to tests. Also, in the regular course, the content would

be distributed over a more extended period of time which could make students feel less stressed and achieve higher scores in the flipped teaching modality. Flipped teaching, as it exposes students to content multiple times through pre-class and in-class activities, was expected to increase academic performance in both the above- and below-median students. It was speculated that both male and female students would benefit from FF, particularly the regular 16-week course (Gross et al., 2015; Yan et al., 2018; Gopalan et al., 2020). Yan et al. reported that female students were more efficient than their male counterparts in a UF course. This study also showed that the motivation of the male students was enhanced by the FF method (Yan et al., 2018). Gross et al. compared student performance between the FF method and found that female students benefited more than male students (Gross et al., 2015).

2 Materials and methods

This study was conducted at a Community College immediately before the COVID pandemic. A total of 197 students participated in the study, and all were instructed by the same faculty member. Participants self-selected to enroll in an Introduction to Human Anatomy taught by the same professor. There was no prerequisites for this class. There was no exclusion of students. All students enrolled in the course were included in this study. There were no repeat students in any of the classes studied. A comparison of final examination scores between the traditional UF versus PF and FF sections of an 8-week *Introduction to Human Anatomy* course was made (Figure 1). Next, a comparison of FF between the accelerated 8-week and the regular 16-week course was carried out (Figure 2). This study was approved by the Southern Illinois University Edwardsville Institutional Review Board. An exempt status was received for the protocol (Protocol # 112).

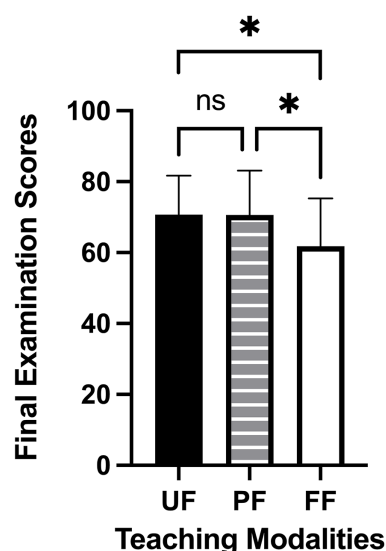
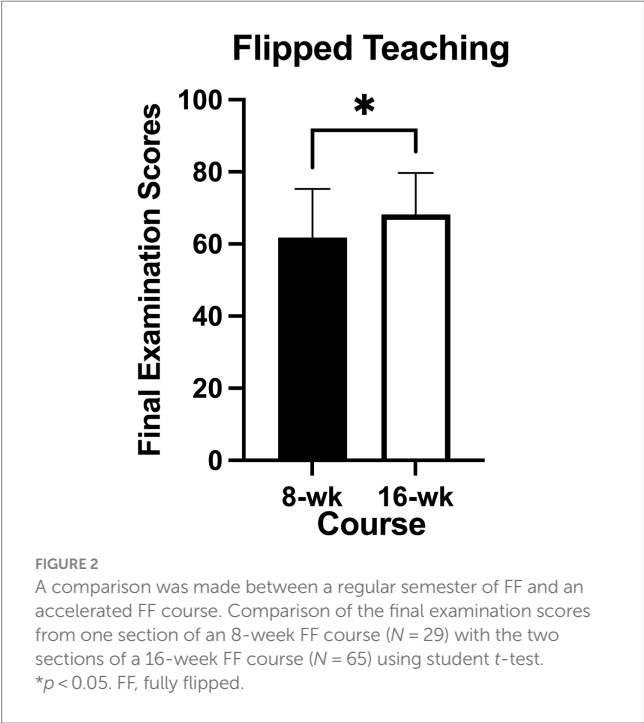


FIGURE 1

A comparison was made between UF, PF, and FF teaching modalities using the final examination scores of each of these courses. UF ($N = 41$), PF ($N = 62$), and FF course ($N = 29$) in an 8-week course using one-way analysis of variance. $*p < 0.05$. UF, unflipped; PF, partially flipped, and FF, fully flipped.

For UF classes, traditional lectures were delivered by the professor, slide by slide. Students were not required to do anything prior to coming to class. For PF and FF classes, students were required to read assigned pages in the textbook and study slides before coming to class.



Various activities, ranging from Q&A mini-lectures, small group problem-solving, and games, were implemented during class. The UF course was an 8-week accelerated course that was taught in the Spring of 2018 in two sections. The PF course was also an 8-week accelerated course that was taught in the Fall of 2018 as two separate sections. Both 8-week and 16-week FF courses were taught in Spring 2019 (Tables 1, 2).

All courses were model-based and used a systemic approach. The 8-week courses included weekly examinations and laboratory practice. Exams were bi-weekly for the 16-week class (Tables 1, 2). The examinations in all courses contained multiple-choice and short-answer questions. Except for the short answer questions, which were at the application, analysis, or synthesis level of Bloom's taxonomy, the multiple-choice questions were at the factual or comprehension levels. Students were not allowed to use any resources or collaborate with other students during the exam. The instructor proctored all the examinations throughout the entire period. The weekly examinations in the accelerated course and the bi-weekly examinations in the regular course were tested out of 50 points, whereas the final examination in both formats was worth 100 points. The same final examination was used between sections and was conducted during their scheduled class time. Students were not given copies of their final exam back to prevent leakage of questions but were encouraged to make appointments to review their performance. Moreover, every course had pre-lecture quizzes and pre-class assignments, except for the UF 8-week course. Furthermore, the instructor graded students for participation and extra credits in all classes. The total synchronous lecture and lab

TABLE 1 Course details between the regular 16-week and accelerated 8-week semesters.

8-week course curriculum				
Week	Content covered	Examinations	PF flipped content	
			Section 1	Section 2
1	Chemistry, Cell, Tissues	1	–	–
2	Heart, Respiratory System, Blood Vessels	2	–	–
3	Vascular, Digestive, and Lymphatic Systems	3	–	Flipped
4	Endocrine, Reproductive, and Urinary Systems.	4	–	Flipped
5	Skin and Nervous System I: Central Nervous System	5	Flipped	–
6	Nervous System II: Somatic and Autonomic Nervous Systems, Skeletal System I: Axial Skeleton	6	Flipped	–
7	Skeletal System II: Appendicular Skeleton, Articulations, Muscle Tissues, and Muscular System	7	–	–
8	Review	Comprehensive	–	–

16-week course curriculum		
Week	Content covered	Examinations
1–2	Chemistry, Cell, Tissues	1
3–4	Heart, Respiratory System, Blood Vessels	2
5–6	Vascular, Digestive, and Lymphatic Systems	3
7–8	Endocrine, Reproductive, and Urinary Systems.	4
9–10	Skin and Nervous System I: Central Nervous System	5
11–12	Nervous System II: Somatic and Autonomic Nervous Systems, Skeletal System I: Axial Skeleton	6
13–14	Skeletal System II: Appendicular Skeleton, Articulations, Muscle Tissues, and Muscular System	7
15–16	Review	Comprehensive

TABLE 2 Teaching Modalities, Course Characteristics, and Student Demographics.

Modality	Course details			Student details				
	Description	Examinations	Time between examinations	Section 1	Section 2	Male	Female	Total
8-week UF	A traditionally taught course in an accelerated format	7 Examinations and 1 Comprehensive Examinations	1 week	22	19	11	30	41
8-week PF	A course taught with components of flipped teaching for two exams	7 Examinations and 1 Comprehensive Examinations	1 week	30	32	14	48	62
8-week FF	A fully flipped course taught in an accelerated format	7 Examinations and 1 Comprehensive Examinations	1 week	29	-	6	23	29
16-week FF	A fully flipped course taught in a traditional 16-week format	7 Examinations and 1 Comprehensive Examinations	2 weeks	37	28	30	35	65

time per week for the UF and PF classes was 6 h and 20 min. The 8-week FF class had synchronous lectures and lab for 6 h and 20 min. The 16-week FF course had a synchronous lecture and lab time of 3 h and 10 min per week. Thus, the 8-week and 16-week course laboratory times were equated in the overall course configuration.

2.1 The 8-week accelerated course

The PF modality contained 25% flipped teaching, whereas the remaining 75% was taught in the UF style. To avoid the potential for the bias of having the flipped content be different in difficulty than the UF content (in the PF portion of the study), the topics flipped in one section were intentionally selected to be different from the second section. For example, the first section flipped the nervous system portion (tested in examinations five and six), whereas the second section flipped digestive, reproductive, and renal systems, which were covered in examinations three and four (Table 1). In addition, the scores of the PF portion of the semester were compared with the matching UF content.

All students in the flipped teaching design had access to PowerPoint slides (Windows 8, Redmond, WA) with voice annotation. Assigned readings and slides were posted on Canvas by week and were accessible to students the Friday before the week each system was to be discussed for asynchronous studying. A quiz that the students were expected to take ensured not only their participation but also to screen if students struggled with any particular content. Students were expected to independently review the content before class, participate in the discussion, and engage in learning activities. These learning activities include question-and-answer sessions, as well as informal group activities and quizzes during the synchronous teaching portion. Students in the flipped courses were informed that there would be no re-lecturing of the flipped content and they should ask questions by emailing the instructor before or at the beginning of the synchronous component, prior to engaging in learning activities. Students in the UF portion received slides with voice annotation on the day of lecture and a pop quiz during class time.

2.2 Statistical analysis

G* power (version 3.1) analysis was used to verify that the sample size was sufficiently large enough to provide at least 80% Power in the proposed statistical analysis. Specifically, the parameters used to conduct the power analysis for calculating the desired sample size included: Power = 80%, alpha = 5%, and a moderate to moderately large effect size of partial eta squared = 0.15. This analysis indicated that the desired sample size under these conditions, taking into account the time series experimental design as well as the four possible treatment combinations, was 116 subjects. The larger sample size used in this study (197 students) allowed for the detection of smaller effect sizes.

A one-way analysis of variance (ANOVA) was used to compare the student performance between UF, PF, and FF teaching formats in an 8-week course, where the student groups were the row factor and teaching modalities were the column factor. A post-hoc student *t*-test was used to compare (1) the final examination scores of the two sections of the PF; (2) flipped portions against UF sections, and (3) the test scores between the eight- and 16-week courses. A two-way ANOVA was used to compare sex differences (row factor) in the final examination scores between the UF, PF, and FF groups (column factor), as well as the comparison of the above- and below-median groups (row factor). The values were significantly different if the *p*-value was less than 0.05%. GraphPad Prism version 9.2 (GraphPad Software, San Diego, CA) was utilized in all of these analyses and the preparation of histograms (Figures 1–5). The values are presented as mean and standard deviation.

3 Results

3.1 Comparison of teaching modalities in the accelerated (8-week course)

A comparison was made between UF, PF, and FF using the final examination scores of each of these courses. Two sections each of UF (*N* = 41) and PF (*N* = 62) were compared with one section of the FF

Comparison of Teaching Modalities between Male and Female Students

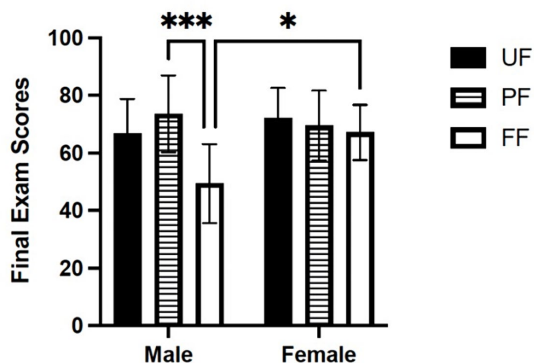


FIGURE 3

The effect of teaching modality on male versus female students was examined using an accelerated 8-week course. Comparison of final examination scores from the 8-week UF, PF, and FF courses between male (UF: $N = 11$, PF: $N = 14$, and FF: $N = 6$) and female students (UF: $N = 31$, PF: $N = 38$, and FF: $N = 14$) using two-way analysis of variance. * $p < 0.05$; *** $p < 0.001$. UF, unflipped; PF, partially flipped; and FF, fully flipped.

Flipped Teaching Duration and Gender

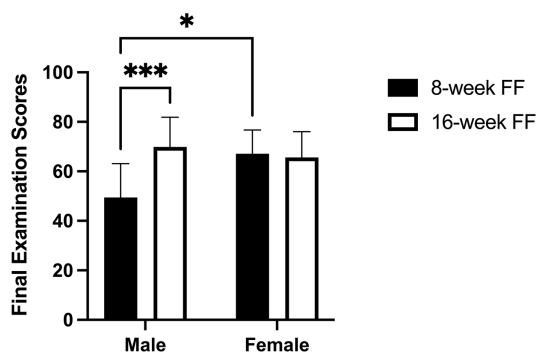


FIGURE 4

A comparison of the final examination scores between two FFs, 8-week and 16-week, revealed that the scores from the 16-week course were significantly greater among male students (male $N = 6$; female $N = 14$) versus 16-week FF (male $N = 30$ and female $N = 19$) using two-way analysis of variance. * $p < 0.05$; *** $p < 0.001$. FF, fully flipped.

course ($N = 29$) of an 8-week accelerated course. Post-hoc analyses revealed no significant difference in the final examination scores between the UF (70.71 ± 11.03) and PF (70.64 ± 12.52) groups ($p = 0.999$). Both UF ($p = 0.022$) and PF ($p = 0.018$) scores were significantly greater than in the FF group [61.80 ± 13.48 ; $F(2,111) = 4.387$; Figure 1].

The scores of the PF portion of the semester were studied to determine whether they would be similar or different from the matching UF content. The weekly examination scores from PF within section 1 (examinations three and four; 34.83 ± 8.412) were compared with the examination scores of the same content from the UF section (section 2; 36.28 ± 7.652). Similarly, the weekly examination scores from section 2 (examinations 5 and 6; 32.84 ± 7.212) were compared

with the same content from the UF (section 1; 34.38 ± 7.559) portion of the 8-week accelerated course and noted that they were not significantly different (section 1 $p = 0.3739$ and section 2 $p = 0.72$).

3.2 Accelerated (8-week) versus regular (16-week) FF

A comparison was made between a regular semester of FF and an accelerated FF course. When the final examination scores from one 8-week FF course (61.80 ± 13.48 , $N = 29$) were compared with two 16-week FF sections (68.22 ± 11.48 , $N = 65$), the scores in the 16-week course were significantly higher ($p = 0.0492$; Figure 2).

3.3 Effect of teaching modalities on male and female students

The effect of teaching modality on male versus female students was examined using an accelerated 8-week course. When the final examination scores were compared among the three teaching modalities, UF, PF, and FF, the scores from the FF (49.42 ± 13.72 , $N = 6$) were significantly lower compared to the PF group (73.61 ± 13.38 , $N = 14$) among male students [$F(2,108) = 5.316$; $p = 0.006$]. The FF scores in the male student group were not significantly different from the UF (70.71 ± 11.03 , $N = 11$). On the other hand, female students performed similarly in all three modalities (UF 72.07 ± 10.6 ; $N = 31$, PF 69.55 ± 12.19 ; $N = 38$, and FF 67.11 ± 9.6 ; $N = 14$), as there was no significant difference between the three teaching methods tested. Male students' final examination scores from the FF group were significantly lower compared to the female FF group ($p = 0.0194$; Figure 3).

A comparison of the final examination scores between two FFs, 8-week and 16-week, revealed that the scores from the 16-week course were significantly greater among male students (69.87 ± 12.01 , $N = 30$) than the accelerated 8-week FF [49.42 ± 13.72 ; $N = 6$; $F(1,65) = 11.67$; $p = 0.0011$]. Female students, on the other hand, performed similarly in both the 8-week (67.12 ± 9.60 , $N = 14$, $p = 0.0121$) and 16-week FF (65.63 ± 10.37 , $N = 19$, Figure 4).

3.4 Student performance among the above- and below-median students

The final examination scores were compared between the above- and below-median students in the three teaching modalities of 8-week courses, UF, PF, and FF. Although the interaction between the above- and below-median groups and the teaching modalities was not significant [$F(1,65) = 0.04361$; $p = 0.8352$], the row factor [$F(1,65) = 82.86$; $p < 0.0001$] and the column factor [$F(1,65) = 9.2$; $p = 0.0034$] were significantly different. Turkey's multiple comparison test revealed that the above-median students' performance in the FF (71.35 ± 7.01 , $N = 10$) was significantly lower compared to PF (80.92 ± 5.30 , $N = 26$, $p < 0.01$) whereas the below-median students' performance in the FF (52.25 ± 11.48 , $N = 10$) was significantly lower compared to the UF and PF portion (UF 61.93 ± 6.45 ; $N = 21$; PF 60.36 ± 8.48 ; $N = 26$; $p < 0.05$). Overall, the student performance in FF was the

Teaching Modality among Upper and Lower Half of the Class

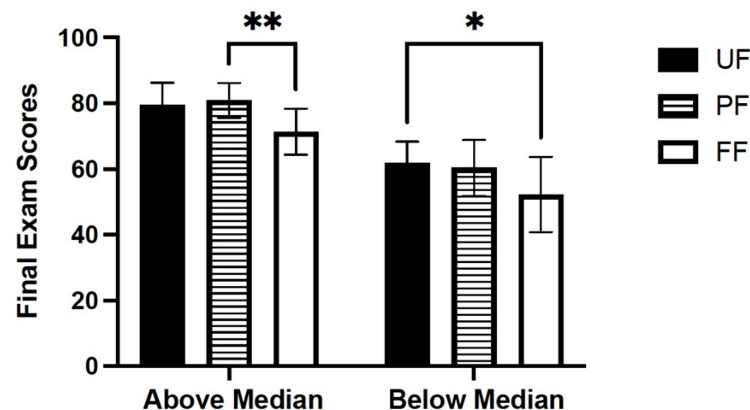


FIGURE 5

The final examination scores were compared between the above- and below-median students in the three teaching modalities of 8-week courses, UF, PF, and FF. Comparison of final examination scores of different teaching modalities of the accelerated 8-week courses between the above- and below-median students using two-way analysis of variance. * $p < 0.05$; ** $p < 0.01$. UF, unflipped; PF, partially flipped; and FF, fully flipped.

lowest among all the teaching methods tested in the accelerated course (Figure 5).

A comparison of student performance in their final examinations of an 8-week versus a 16-week FF course was made after separating the students into above- and below-median groups. Although there was a slightly higher in student scores in the 16-week FF (above median 77.16 ± 5.8 ; below-median 58.92 ± 7.93 ; $N = 25$) compared to the 8-week (above median 71.35 ± 7.01 ; and below-median 52.25 ± 11.48 ; $N = 10$), these numbers were not significantly different.

4 Discussion

Traditional approaches, such as didactic lectures and dissection-based laboratory sessions, have prevailed thus far in the educational approach of anatomical sciences, but continued testing of new techniques is essential to advancing student-centered teaching modalities in the future (Salinas-Alvarez et al., 2021). This study attempted to examine student performance between the UF, PF, and FF formats in an accelerated 8-week *Introduction to Human Anatomy* course at a community college. A comparison of above- and below-median groups, as well as male and female students was made. Additionally, the final examination scores in the accelerated 8-week semester were evaluated with those of the regular 16-week FF courses to analyze the course length's impact on student performance.

As shown in Figure 1, the scores in the accelerated 8-week FF were significantly lower than both UF and PF. Since it was an accelerated course, the students may not have had the time to complete expected pre-class activities in order to be prepared for the in-class engagement. Since this course was an accelerated course, the limited time for the students to interact with the flipped teaching modality appears to be a challenge. In a condensed version of a course, the study material is generally more overwhelming to handle in between scheduled classes.

More likely, any performance difference while comparing the effectiveness of flipped teaching in the form of final examination

scores between a regular (16-week) and an accelerated anatomy course was due to the fact that the students in the regular format had more contact hours with the instructor, which helped them learn the content better than in the accelerated course, where there was limited time between the instructor and the students (bi-weekly instead of weekly assessments). In contrast to the findings from this study, Barral et al. reported improved scores with flipped teaching in their accelerated *Introductory Biology* course (Barral et al., 2018). Reimer et al. also reported flipped teaching to promote student performance in the 8-week summer organic chemistry course (Reimer et al., 2021). It must be noted that both studies were conducted at four-year institutions, whereas the current study tested community college students. Whether the discrepancy in the findings is due to the difference in the student populations is unknown. Community colleges typically have a unique student population where students have outside responsibilities such as employment, family commitments, and a greater range of diversity in terms of age (Karp et al., 2010; Dougherty, 2016). The considerable amount of out-of-class time required in this teaching method could be one reason for the lower scores observed in the 8-week FF course. Any absence due to increased commitment could affect student performance since they are more likely to be unfamiliar with the concepts.

When comparing PF between two sections of the 8-week courses, it was intentionally designed to select content from different units of the course, such as the nervous system in one section versus the digestive, renal, and reproductive topics in the second section, to avoid assuming that the selected unit was the most difficult or the easiest compared to the rest of the content. Irrespective of the content flipped, PF was not effective. Partial flipping in this study design occurred either in the middle (3rd and 4th examinations) or at the end (5th and 6th examinations) of the semester. It is shown that the students adjust to their courses at the very beginning of the semester, and it is rather difficult to introduce new modes in the latter part of the semester (Syofyan and Siwi, 2018). Students are often preoccupied by due dates for papers, presentations, and tests,

which may lead some to not fully focus on readjusting to the course in the middle or the end of the semester (Gerrard et al., 2017). On a positive note, this study suggests that instructors can add elements of the flipped classroom without negatively impacting the students since PF was equally effective as UF. Educators must carefully plan the implementation of flipped components in their courses.

Based on the results from this study, the ideal FF approach for community college students, especially male students, was the regular 16-week rather than the accelerated 8-week course (Figures 3, 4). The success of the 16-week format could be that the content was spaced out, made available throughout the semester, and delivered in more digestible chunks. Perhaps, students felt less constricted by time to process information and become familiar with the teaching method, which then encouraged more interaction in class. A study by Romaker (2021) using a regular semester reported that male students at the community college exhibited higher passing rates in the flipped teaching design when compared to traditional learning (Romaker, 2021). The opposite effect was seen in this study. Flipped teaching design requires structured independent study time to be successful and students in the accelerated course have limited time to prepare for the in-class work (Heijstra and Sigurðardóttir, 2018). The number of male students in the current study was small in the 8-week ($N=6$) compared to the 16-week FF course ($N=30$). The small number in the 8-week FF course might have contributed to the lower scores, especially if one or two students were not as strong as the rest. Female students performed similarly in the 8-week and 16-week FF, suggesting that they were able to adapt to both the accelerated 8-week and the regular 16-week courses (Mašić et al., 2021). Whether the higher number of female students compared to male students in every section of the 8-week accelerated course may have influenced these findings is unknown. The 16-week FF course had a similar distribution of male and female students, and hence the results were comparable in terms of their performance (Table 2).

When the class was separated into upper and lower halves and analyzed the modality that was effective for these two groups, it was apparent that FF in an 8-week course was not well received by both halves of the class. Flipped teaching implementation occurred for the first time in this course. The comfort level of the instructor may help improve the situation in the future with necessary adjustments. Students in an accelerated 8-week class require substantially more organization to embrace active and innovative learning in the classroom (Barral et al., 2018).

Several factors can affect student performance without a selection process or prerequisite requirements, including varied student abilities, different backgrounds, and motivation. Some students may need more foundational knowledge to succeed, making it challenging for the teacher to meet the needs of all students. Specific teaching strategies may provide additional support to ensure that all students have the opportunity to succeed. Flipped teaching could be one of those, but more work is needed, especially at the community college level.

4.1 Limitations of the study

The limitations of this study were: (1) the number of sections for the different learning modalities. The 8-week FF only had one

section, while the 8-week UF and PF, as well as the 16-week FF, had two sections; (2) the absence of 16-week courses for UF and PF; (3) the male sample size which was noticeably smaller in the 8-week course than the female students; (4) Although the final exams were identical, the scores between sections might reflect differences in the student body due to many factors that are beyond our capability to analyze. Having a more uniform sample size would strengthen the findings of this study and yield more reliable conclusions.

5 Conclusion

The findings from this study suggest that FF in a regular 16-week course increased examination scores compared to the accelerated 8-week course. Male students struggled with the accelerated 8-week FF course compared to the 16-week semester. Thus, the length of the course appears to be a key factor for students to absorb course content where the regular 16-week course is more effective than the accelerated course for flipped teaching. Both above- and below-median groups of students were negatively affected by FF in the accelerated course. Female students scored similarly in all modalities tested, suggesting that they could handle different modalities equally well. PF did not negatively impact student performance, which suggests that educators can add elements of the flipped classroom wherever it is appropriate for their courses. There are many contexts in which the flipped classroom could be easier and more practical to implement.

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 humans were approved by the Southern Illinois University Edwardsville Institutional Review Board. An exempt status was received for the protocol (Protocol # 112). The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because an exempt status was received for the protocol.

Author contributions

CG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. EB: Writing – original draft, Writing – review & editing. SD: Writing – original draft, Writing – review & editing. K-LN: Investigation, Methodology, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was supported by the NSF-Research Coordination Network award (Grant Abstract # 1624200) and by the American Physiological Society Teaching Career Enhancement Award.

Acknowledgments

Thanks to Julia Armbruster, Jacen King, and Paris Kinsley for their assistance in the revision of this manuscript.

References

- Al-Samarraie, H., Shamsuddin, A., and Alzahrani, A. I. (2020). A flipped classroom model in higher education: a review of the evidence across disciplines. *Educ. Tech. Res. Dev.* 68, 1017–1051. doi: 10.1007/s11423-019-09718-8
- Asad, A., and Srivastava, S., and Verma MK. (2020). Evolution of COVID-19 pandemic in India. *INAE* 5, 711–718. doi: 10.1007/s41403-020-00166-y
- Baig, M. I., and Yadegaridehkordi, E. (2023). Flipped classroom in higher education: a systematic literature review and research challenges. *Int. J. Educ. Technol. High. Educ.* 20:61. doi: 10.1186/s41239-023-00430-5
- Barral, A. M., Ardi-Pastores, V. C., and Simmons, R. E. (2018). Student learning in an accelerated introductory biology course is significantly enhanced by a flipped-learning environment. *CBE—life. Sci. Educ.* 17:ar38. doi: 10.1187/cbe.17-07-0129
- Bove, L. A., and Conklin, S. (2019). Using the technology adoption model to assess faculty comfort with the learning management system. *OJDLA* 22. Available at: https://www.westga.edu/~distance/ojdla/fall223/bove_conklin223.html (Accessed August 16, 2021)
- Cevikbas, M., and Kaiser, G. (2020). Flipped classroom as a reform-oriented approach to teaching mathematics. *ZDM* 52, 1291–1305. doi: 10.1007/s11858-020-01191-5
- Cullinane, D. P., and Barry, D. S. (2023). Dyad pedagogy in practical anatomy: a description of the implementation and student perceptions of an adaptive approach to cadaveric teaching. *Anat. Sci. Educ.* 16, 99–115. doi: 10.1002/ase.2184
- Dehghanzadeh, S., and Jafaraghaee, F. (2018). Comparing the effects of traditional lecture and flipped classroom on nursing students' critical thinking disposition: a quasi-experimental study. *Nurse Educ. Today* 71, 151–156. doi: 10.1016/j.nedt.2018.09.027
- Doğan, Y., Batdı, V., and Yaşar, M. D. (2023). Effectiveness of flipped classroom practices in teaching of science: a mixed research synthesis. *Res. Sci. Technol.* 41, 393–421. doi: 10.1080/02635143.2021.1909553
- Dougherty, K. J. (2016). The contradictory college: the conflicting origins, impacts, and futures of the community college. Albany, New York: SUNY Press.
- Dove, A. (2013). Students' perceptions of learning in a flipped statistics class. *SITE Proceed.*, 393–398.
- Eames, M., Luttman, S., and Parker, S. (2018). Accelerated vs. traditional accounting education and Ccpa exam performance. *JAED* 44, 1–13. doi: 10.1016/j.jaccedu.2018.04.004
- Faucette, A. N. (2023). Incision precision: engaging students during dissection labs. *J. Microbiol. Biol. Educ.* 24, e00193–e00122. doi: 10.1128/jmbe.00193-22
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS* 111, 8410–8415. doi: 10.1073/pnas.1319030111
- Gerrard, D., Newfield, K., Asli, N. B., and Variawa, C., (2017). Are students overworked? Understanding the workload expectations and realities of first-year engineering [online]. ASEE Annual Conference & Exposition. Available at: <https://peer.asee.org/are-students-overworked-understanding-the-workload-expectations-and-realities-of-first-year-engineering> (Accessed September 17, 2021)
- Gopalan, C., Fentem, A., and Rever, A. L. (2020). The refinement of flipped teaching implementation to include retrieval practice. *Adv. Physiol. Educ.* 44, 131–137. doi: 10.1152/advan.00143.2019
- Green, C. (2020). Teaching accelerated nursing students' self-care: a pilot project. *Nurs. Open* 7, 225–234. doi: 10.1002/nop2.384
- Gross, D., Pietri, E. S., Anderson, G., Moyano-Camihort, K., and Graham, M. J. (2015). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *LSE* 14:ar36. doi: 10.1187/cbe.15-02-0040
- Halpern, D. F., Straight, C. A., and Stephenson, C. L. (2011). Beliefs about cognitive gender differences: accurate for direction, underestimated for size. *Sex Roles* 64, 336–347. doi: 10.1007/s11199-010-9891-2
- Halpin, P. A., and Gopalan, C. (2021). Using dramatizations to teach cell signaling enhances learning and improves students' confidence in the concept. *Adv. Physiol. Educ.* 45, 89–94. doi: 10.1152/advan.00177.2020
- Heijstra, T. M., and Sigurðardóttir, M. S. (2018). The flipped classroom: does viewing the recordings matter? *Act. Learn. High. Educ.* 19, 211–223. doi: 10.1177/1469787417723217
- Herreid, C. F., and Schiller, N. A. (2013). Case studies and the flipped classroom. *J. Coll. Sci. Teach.* 42, 62–66.
- Herrero, J. I., and Quiroga, J. (2020). Flipped classroom improves results in pathophysiology learning: results of a nonrandomized controlled study. *Adv. Physiol. Educ.* 44, 370–375. doi: 10.1152/advan.00153.2019
- Holtbrügge, D., and Mohr, A. T. (2010). Cultural determinants of learning style preferences. *AMLE* 9, 622–637. doi: 10.5465/amle.9.4.zqr622
- Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Educ.* 4:1304016. doi: 10.1080/2331186X.2017.1304016
- Idrizi, E., Filiposka, S., and Trajkovikj, V. (2023). Gender impact on stem online learning: a correlational study of gender, personality traits and learning styles in relation to different online teaching modalities. *Multimed Tools Appl.* 82, 30201–30219. doi: 10.1007/s11042-023-14908-x
- Jdaitawi, M. (2021). Does flipped learning promote positive emotions in science education? A comparison between traditional and flipped classroom approaches. *ECEL* 18, 516–524. doi: 10.34190/JEL.18.6.004
- Jenkins, D., Jaggars, S. S., and Roksa, J. (2014). Completing college: a national view of student attainment rates – Fall 2008 cohort. New York, NY: Community College Research Center.
- Kapur, M., Hattie, J., Grossman, I., and Sinha, T. (2022). Fail, flip, fix, and feed – rethinking flipped learning: a review of meta-analyses and a subsequent meta-analysis. *Front. Educ.* 7:956416. doi: 10.3389/feduc.2022.956416
- Karp, M. M., Hughes, K. L., and O'Gara, L. (2010). An exploration of Tinto's integration framework for community college students. *J. Coll. Stud. Retent.: Res. Theory Pract* 12, 69–86. doi: 10.2190/CS.12.1.e
- Kostka, I., and Marshall, H. W. (2018). “Flipped learning in TESOL: past, present, and future” in Applications of CALL theory in ESL and EFL environments (Hershey, PA: IGI Global), 223–243.
- Li, B.-Z., Cao, N.-W., Ren, C.-X., Chu, X.-J., Zhou, H.-Y., and Guo, B. (2020). Flipped classroom improves nursing students' theoretical learning in China: a meta-analysis. *PLoS One* 15:e0237926. doi: 10.1371/journal.pone.0237926
- Lo, C. K., and Hew, K. F. (2019). The impact of flipped classrooms on student achievement in engineering education: a meta-analysis of 10 years of research. *JEE* 108, 523–546. doi: 10.1002/jee.20293
- Lodge, J. M., Hansen, L., and Cottrell, D. (2016). Modality preference and learning style theories: rethinking the role of sensory modality in learning. *Learn.: Res. Pract.* 2, 4–17. doi: 10.1080/23735082.2015.1083115

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.

- Loyens, S. M. M., van Meerten, J. E., Schaap, L., and Wijnia, L. (2023). Situating higher-order, critical, and critical-analytic thinking in problem- and project-based learning environments: a systematic review. *Educ. Psychol. Rev.* 35:39. doi: 10.1007/s10648-023-09757-x
- Mašić, A., Polz, E., and Bećirović, S. (2021). The relationship between learning styles, GPA, school level and gender. Available at: <https://papers.ssrn.com/abstract=3583905>
- McWatt, S. C. (2021). Responding to COVID-19: a thematic analysis of students' perspectives on modified learning activities during an emergency transition to remote human anatomy education. *Anat. Sci. Educ.* 14, 721–738. doi: 10.1002/ase.2136
- Mojtahedi, M., Kamardeen, I., Rahmat, H., and Ryan, C. (2020). Flipped classroom model for enhancing student learning in construction education. *J. Civ. Eng.* 146:05019001. doi: 10.1061/(ASCE)EJ.2643-9115.0000004
- Morton, D. A. (2020). "The flipped classroom: starting with the end in mind" in Teaching anatomy: a practical guide. eds. L. K. Chan and W. Pawlina (Cham, Switzerland: Springer International Publishing), 97–105.
- Namaziandost, E., Rezaei, Z., Etemadfar, P., and Alekasir, S. (2020). Implementing a flipped model of instruction in the EFL listening classroom: impact on comprehension. *JEFL* 10, 385–401. doi: 10.23971/jeff.v10i2.2065
- Özbay, O., and Çınar, S. (2021). Effectiveness of flipped classroom teaching models in nursing education: a systematic review. *Nurse Educ. Today* 102:104922. doi: 10.1016/j.nedt.2021.104922
- Ozudogru, M., and Aksu, M. (2020). Pre-service teachers' achievement and perceptions of the classroom environment in flipped learning and traditional instruction classes. *Australas. J. Educ. Technol.* 36, 27–43. doi: 10.14742/ajet.5115
- Pierce, R., and Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *Am. J. Pharm. Educ.* 76:196. doi: 10.5688/ajpe7610196
- Reimer, L. C., Denaro, K., He, W., and Link, R. D. (2021). Getting students back on track: persistent effects of flipping accelerated organic chemistry on student achievement, study strategies, and perceptions of instruction. *J. Chem. Educ.* 98, 1088–1098. doi: 10.1021/acs.jchemed.0c00092
- Riedl, A., Yeung, F., and Burke, T. (2021). Implementation of a flipped active-learning approach in a community college general biology course improves student performance in subsequent biology courses and increases graduation rate. *Life Sci. Educ.* 20:ar30. doi: 10.1187/cbe.20-07-0156
- Romaker, D. (2021). Benefits of flipped learning for developmental mathematics. *CCJRP* 47, 9–21. doi: 10.1080/10668926.2021.1919241
- Roy, H., Ray, K., Saha, S., and Ghosal, A. K. (2020). A study on students' perceptions for online zoom-app based flipped class sessions on anatomy organised during the lockdown period of COVID-19 epoch. *JCDR* 14, 1–4. doi: 10.7860/JCDR/2020/44869.13797
- Salinas-Alvarez, Y., Quiroga-Garza, A., Martinez-Garza, J. H., Jacobo-Baca, G., Zarate-Garza, P. P., Rodríguez-Alanís, K. V., et al. (2021). Mexican educators survey on anatomical sciences education and a review of world tendencies. *Anat. Sci. Educ.* 14, 471–481. doi: 10.1002/ase.2017
- Shapiro, D., Dundar, A., Wakhungu, P. K., Yuan, X., and Nathan, A. (2015). Tracking transfer: New measures of institutional and state effectiveness in helping community college students attain bachelor's degrees. Herndon, VA: National Student Clearinghouse.
- Sivarajan, S., Soh, E. X., Zakaria, N. N., Kamarudin, Y., Lau, M. N., Bahar, A. D., et al. (2021). The effect of live demonstration and flipped classroom with continuous formative assessment on dental students' orthodontic wire-bending performance. *BMC Med. Educ.* 21:326. doi: 10.1186/s12909-021-02717-5
- Strelan, P., Osborn, A., and Palmer, E. (2020). The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels. *Educ. Res. Rev.* 30:100314. doi: 10.1016/j.edurev.2020.100314
- Subari, K. (2017). "Improving understanding and reducing matriculation students' misconceptions in immunity using the flipped classroom approach" in Overcoming students' misconceptions in science. eds. M. Karpudewan, A. M. Zain and A. Chandrasegaran (Singapore: Springer), 209–226.
- Syofyan, R., and Siwi, M. K. (2018). The impact of visual, auditory, and kinesthetic learning styles on economics education teaching. In 1st International Conference on Economics Education, Economics, Business and Management, Accounting and Entrepreneurship (PICEBA 2018)* 642–649.
- Talbert, R. (2017). Flipped learning: A guide for higher education faculty: Stylus Publishing.
- Tune, J. D., Sturek, M., and Basile, D. P. (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. *Adv. Physiol. Educ.* 37, 316–320. doi: 10.1152/advan.00091.2013
- Twigg, C. A. (2009). Using asynchronous learning in redesign: reaching and retaining the at-risk student. *JALN* 13, 147–155.
- Voyer, D., and Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. doi: 10.1037/a0036620
- Wei, X., Cheng, I.-L., Chen, N.-S., Yang, X., Liu, Y., Dong, Y., et al. (2020). Effect of the flipped classroom on the mathematics performance of middle school students. *Educ. Technol. Res. Dev.* 68, 1461–1484. doi: 10.1007/s11423-020-09752-x
- Wozny, N., Balser, C., and Ives, D. (2018). Evaluating the flipped classroom: a randomized controlled trial. *J. Econ. Educ.* 49, 115–129. doi: 10.1080/00220485.2018.1438860
- Yan, J., Li, L., Yan, J., and Niu, Y. (2018). A comparison of flipped and traditional classroom learning: a case study in mechanical engineering. *Int. J. Eng. Educ.* 34, 1876–1887.
- Yoo, H., Kim, D., Lee, Y.-M., and Rhyu, I. J. (2020). Adaptations in anatomy education during COVID-19. *J. Korean Med. Sci.* 36:e13. doi: 10.3346/jkms.2021.36.e13

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

