Research Topics

Advances in multimodal learning: Pedagogies, technologies, and analytics

Edited by
Heng Luo, Wang Qiyun, Zhongling Pi and Ikseon Choi

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-3917-0
DOI 10.3389/978-2-8325-3917-0

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

Advances in multimodal learning: Pedagogies, technologies, and analytics

Topic editors
Heng Luo — Central China Normal University, China
Wang Qiyun — Nanyang Technological University, Singapore
Zhongling Pi — Shaanxi Normal University, China
Ikseon Choi — University of Georgia, United States

Citation
Luo, H., Qiyun, W., Pi, Z., Choi, I., eds. (2023). *Advances in multimodal learning: Pedagogies, technologies, and analytics*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3917-0

Table of contents

TYPE Editorial
PUBLISHED 31 October 2023
DOI 10.3389/fpsyg.2023.1286092

OPEN ACCESS

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

*CORRESPONDENCE
Heng Luo
✉ luoheng@mail.ccnu.edu.cn

RECEIVED 30 August 2023
ACCEPTED 20 October 2023
PUBLISHED 31 October 2023

CITATION
Luo H (2023) Editorial: Advances in multimodal learning: pedagogies, technologies, and analytics. *Front. Psychol.* 14:1286092.
doi: 10.3389/fpsyg.2023.1286092

Editorial: Advances in multimodal learning: pedagogies, technologies, and analytics

Heng Luo*

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

KEYWORDS

multimodal learning, multimodal stimuli, multimodal learning spaces, multimodal behaviors, multimodal learning analytics

Editorial on the Research Topic
Advances in multimodal learning: pedagogies, technologies, and analytics

The rapid development of digital technologies and data-driven techniques has led to advances in multimodal learning featured by multimodality in instructional stimuli, learning spaces, behavioral pattern, and data sources (Blikstein and Worsley, 2016; Di Mitri et al., 2018; Chango et al., 2021). The combination of multiple sensory stimuli (e.g., visual, audio, verbal, tactile, and olfactory) in instructional content is known to promote cognitive performance, sense of presence, and learning engagement (Moreno and Mayer, 1999; Marucci et al., 2021). The fusion of traditional and digital learning spaces enabled by online platforms and extended reality has created more personal, accessible, and risk-free learning conditions that may otherwise be unavailable (Garrison and Kanuka, 2004; Allcoat et al., 2021; Luo et al., 2021). Multimodality in learning also emphasizes the multiplicity of mode of behaviors such as communication, interaction, and regulation, leading to increased interests in their impacts on learning achievement and experience (Mangaroska et al., 2020; Cloude et al., 2022; Ninaus and Sailer, 2022). More importantly, Multimodal Learning Analytics (MMLA) allows traces of various cognitive, behavioral, and affective indicators to be extracted from multiple data sources (e.g., eye-tracking, wearable cameras, gesture recognition systems, infrared imaging, biosensors) to assist the measurement and understanding of complex learning processes (Blikstein and Worsley, 2016; Di Mitri et al., 2018; Chango et al., 2021).

Consequently, to capture the unique benefits of multimodal learning, research attention needs to be paid to four key aspects of multimodality, including learning content, learning space, learning process, and learning analytics. This Research Topic is an attempt to address such a research need by illustrating recent research development of multimodal learning in the four aspects. It comprises a total of 15 articles contributed by 63 authors globally, from academic and research institutions in Mainland China, Taiwan, United States, United Kingdom, Austria, Canada, Spain, and Colombia. Based on the four aspects of multimodality in learning, we classified the 15 articles into four themes: *design of multimodal stimuli*, *affordances of multimodal learning space*, *analysis of multimodal behaviors*, and *application of multimodal analytics*.

The first theme addresses the importance of using multimodal stimuli for creating effective instructional content. Mou et al. describe and evaluate a novel instructional media featuring two modes of stimuli: video and in-video text messages in the forms of bullet comments. While this multiplicity of media content increases parasocial interactions among learners, it also induces greater cognitive load and thus hinders learning performance, revealing a potential caveat of using multimodal stimulus in instruction. Xia et al. take a similar interest in instructional video and its impact on motor skill learning and find that the creation and sharing of video lead to increased intrinsic motivation and motor task perseverance, as compared to self-exercise. Han et al. further investigate two specific stimuli (i.e., text annotations and color changes) as visual signals in immersive virtual reality learning environments, and report empirical evidence supporting their effectiveness on learning for students with low prior knowledge levels. The findings of these three studies contribute to the existing literature of multimodal stimuli by exploring new a design feature, learning domain, and learning environment.

The second theme focuses on the fusion of multiple learning spaces and their unique learning affordances. Sbaffi and Zhao document the design and implementation of an online gamified learning module as a hybrid learning space to teach knowledge of academic integrity for university students. Likewise, Zhang et al. explore in-service teachers' professional development in online learning space by designing and validating an instrument to measure the quality of supporting platforms. Zhao and Xue take a special interest in the transition between online and offline learning spaces in the post-pandemic era, highlighting the transitioning challenges in administration, infrastructure, pedagogy, and finance. In addition to online learning space, two studies focus on extended reality (XR) spaces by synthesizing the existing research findings: Amores-Valencia et al. conduct a systematic review of augmented reality in secondary education, and Chen et al. perform a meta-analysis to determine the effect of XR on language learning. Both studies report positive findings regarding the learning effect of XR learning spaces. Lastly, Schmidthaler et al. introduce a game space (Poly-Universe) absent of digital technologies and prove its effectiveness in teaching computational thinking for children in an interdisciplinary fashion.

The third theme explores the patterns of multimodal behaviors in diverse learning contexts using advanced analysis methods. Tlili et al. use a lag sequential analysis approach to examine the impact of personalities on students' navigational patterns among 12 learning behaviors in an online course, and identify the traits of extraversion, conscientiousness, neuroticism, and openness as potential moderating factors. In a similar online learning context, Luo et al. focus on students' multitasking behaviors in online learning. Using structural equation modeling (SEM), the study reveals predictive path relationships among polychronicity, multitasking behavior, and perceived learning performance. Tao et al. shift their research attention to teachers and employ Partial Least Squares to explore the factors affecting teachers' precision teaching ability with a focus on the mediating effect of data consciousness.

The fourth theme highlights the role of MMLA in assisting people to understand and optimize the learning process and environment with data-driven decisions. Wang et al. conduct an MMLA model enabled by natural language processing and machine learning to predict college students' problem-based learning performance in a blended course, informed by both clickstream data and learner-generated text content. Xiao et al. further integrate physiological data such as brainwaves, eye movements, and facial expressions in their MMLA model to predict in-service teachers' engagement and performance in an online training program. Moreover, Li et al. innovatively utilize heart rate to profile synchronized physiological arousal during collaborative argumentation and explore the potential influencing factors such as types of challenge and social regulation focus. Consistent with the previous literature (Blikstein and Worsley, 2016; Emerson et al., 2020), the three contributions demonstrate the superiority of multimodal data over omni-modal data in predicting learning performance and experience.

In conclusion, the current Research Topic presents recent findings regarding four important aspects of multimodal learning. It is our hope that the contributions of this topic can extend our conceptual, practical, and methodological understanding of learning in the digital era, and lead to a greater breadth of research perspectives in this rapidly evolving field.

Author contributions

HL: Writing—original draft, Writing—review & editing.

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

References

Allcoat, D., Hatchard, T., Azmat, F., Stansfield, K., Watson, D. G., and von Mühlenen, A. (2021). Education in the digital age: learning experience in virtual and mixed realities. *J. Educ. Comp. Res.* 59, 795–816. doi: 10.1177/0735633120985120

Blikstein, P., and Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *J. Learn. Analyt.* 3, 220–238. doi: 10.18608/jla.2016.32.11

Chango, W., Cerezo, R., Sánchez-Santillán, M., Azevedo, R., and Romero, C. (2021). Improving prediction of students' performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. *J. Comp. Higher Educ.* 33, 614–634. doi: 10.1007/s12528-021-09298-8

Cloude, E. B., Azevedo, R., Winne, P. H., Biswas, G., and Jang, E. E. (2022). System design for using multimodal trace data in modeling self-regulated learning. *Front. Educ.* 7, 928632. doi: 10.3389/feduc.2022.928632

Di Mitri, D., Schneider, J., Spanncht, M. M., and Drachsler, H. (2018). From signals to knowledge: a conceptual model for multimodal learning analytics. *J. Comput. Assist. Learn.* 34, 338–349. doi: 10.1111/jcal.12288

Emerson, A., Cloude, E. B., Azevedo, R., and Lester, J. (2020). Multimodal learning analytics for game-based learning. *Br. J. Educ. Technol.* 51, 1505–1526. doi: 10.1111/bjet.12992

Garrison, D. R., and Kanuka, H. (2004). Blended learning: uncovering its transformative potential in higher education. *Inter. Higher Educ.* 7, 95–105. doi: 10.1016/j.iheduc.2004.02.001

Luo, H., Yang, T., Kwon, S., Li, G., Zuo, M., and Choi, I. (2021). Performing versus observing: investigating the effectiveness of group debriefing in a VR-based safety education program. *Comput. Educ.* 175:104316. doi: 10.1016/j.compedu.2021.104316

Mangaroska, K., Martínez-Maldonado, R., Vesin, B., and Gašević, D. (2020). Challenges and opportunities of multimodal data in human learning: the computer science students' perspective. *J. Comput. Assist. Learn.* 37, 1030–1047. doi: 10.1111/jcal.12542

Marucci, M., di Flumeri, G., Borghini, G., Sciaraffa, N., Scandola, M., Pavone, E. F., et al. (2021). The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. *Sci. Rep.* 11, 8. doi: 10.1038/s41598-021-84196-8

Moreno, R., and Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *J. Educ. Psychol.* 91, 358–368. doi: 10.1037/0022-0663.91.2.358

Ninaus, M., and Sailer, M. (2022). Closing the loop – The human role in artificial intelligence for education. *Front. Psychol.* 13, 956798. doi: 10.3389/fpsyg.2022.956798

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpsyg.2022.1016519

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University, China

REVIEWED BY
Zhang Jiakai,
Central China Normal University, China
Qinna Feng,
Central China Normal University, China
Xu Han,
Central China Normal University, China

*CORRESPONDENCE
Liujie Xu
xulj2004@126.com

†These authors have contributed equally to this work

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

RECEIVED 11 August 2022
ACCEPTED 31 August 2022
PUBLISHED 20 September 2022

CITATION
Chen JY, Dai J, Zhu KK and Xu LJ (2022) Effects of extended reality on language learning: A meta-analysis.
Front. Psychol. 13:1016519.
doi: 10.3389/fpsyg.2022.1016519

Effects of extended reality on language learning: A meta-analysis

Jingying Chen[1†], Jian Dai[2,3†], Keke Zhu[4†] and Liujie Xu[3*†]

[1]School of Marxism, Zhejiang University of Technology, Hangzhou, China, [2]School of Management, Zhejiang University of Technology, Hangzhou, China, [3]College of Educational Science and Technology, Zhejiang University of Technology, Hangzhou, China, [4]College of Foreign Languages, Zhejiang University of Technology, Hangzhou, China

In recent years, there has been increasing use of extended reality (XR) in language learning. Many scholars have conducted empirical research on the relationship between the two, but conclusions have been inconsistent, which calls for an organization and reanalysis of relevant literature. Articles published between 2000 and 2022 on the impact of XR on language learning were retrieved from the Web of Science and Scopus databases, and 17 of them (including 21 independent samples and 993 subjects) were included in this meta-analysis. The findings indicate that XR could promote language learning (effect size=0.825). The moderating effects of education level, target language, and technology type were also tested, and the results indicate that the target language type significantly moderated the effect of XR technology on language learning (Q=30.563, p<0.001). Moreover, based on the subgroup analysis, several research questions worthy of further exploration in this field are discussed. Some suggestions are provided, noting that these technologies should be personally designed for learners and learning objects when applied in order to improve the effects of language learning.

KEYWORDS

extended reality, virtual reality, augmented reality, language learning, meta-analysis

Introduction

In the context of economic globalization and cultural diversity, language learning is an inevitable subject for all learners; therefore, it has become a popular research topic among scholars. Today, foreign language learning is still considered extremely difficult by a great number of learners. In response, schools have been trying to apply various technologies to foreign language education. Extended reality (XR), which gradually develops and matures, has garnered increasing attention (Chen, 2016; Luo et al., 2021b). This is because XR provides learners with the authentic language environment that language learning requires, thus enabling users to have a real learning experience (Moeller and Catalano, 2015).

XR is an umbrella term encapsulating augmented reality (AR), virtual reality (VR), and everything in between. AR and VR are two related but distinct technologies. VR, as the name suggests, is an immersive technology that simulates reality and provides users with real experiences (Radianti et al., 2020; Luo et al., 2021a), while AR is a technology that subtly integrates virtual information with the real world; that is, through extensive use of various technical means, the virtual information generated by computer, such as text, image, three-dimensional model, music, video, and so on, is simulated and applied to the real world, and these two types of information complement each other to "augment" the real world (Han et al., 2020; Lee, 2020). In contrast to VR, AR technologies are more effective at providing contextualized and socialized learning experiences (Lin et al., 2019). In the field of educational technology research, these technologies have been examined for a long time. The 2017 Horizon Report listed VR as the technology that holds great potential for boosting the development of education (Freeman et al., 2017). Moreover, in the 2020 EDUCAUSE Horizon Report, it was emphasized that the number of XR users had grown in recent years due to the decreasing cost and innovations in technical capabilities and immersive experience; XR had become an important technology to be implemented in the field of educational technology (Brown et al., 2020; Feng et al., 2021). Scores of scholars have noticed the potential applications of XR-related technologies, and there has been growing recognition among them that technological advances could decrease the costs of equipment and contribute to the promotion of technology (Hwang and Hu, 2013; Vapenstad and Buzink, 2013; Parong and Mayer, 2018; Li et al., 2021). Lee (2020) pointed out that in the past two decades, language learning and XR-related technologies have been very closely linked. Therefore, scholars have also begun to pay attention to this emerging topic. Furthermore, studies have verified that XR technology intervention is beneficial to language learning (Holden and Sykes, 2011; Liu et al., 2016; Binhomran and Altalhab, 2021; Nicolaidou et al., 2021; Zhang et al., 2022). The above findings provide the prerequisite for the educational transformation in language learning enabled by technology. At the same time, education researchers are obliged to explain changes in the education mechanism as well as improvement in learning performance that occurred after the integration of nascent technologies into education. By searching authoritative databases, including the Web of Science core collection and Scopus, we found no systematic meta-analysis of XR technology in language learning. Therefore, from the perspective of technological transformation, we adopted a method of comprehensive research (meta-analysis) that is scientific and systematic to evaluate XR's impact on language learning. This paper reached convincing conclusions by synthesizing various empirical research studies. We mainly explored the following questions:

1. Does XR technology have a significant positive impact on language learning, and if so, what is the effect size?
2. Do education level, target language, technology type, and specific language skills moderate the effect of XR technology on language learning?
3. Based on systematic review and quantitative analysis, what are the potential research questions or research directions that need to be further explored in terms of the effects of XR technology on language learning?

Literature review

XR has been integrated into language learning for a long time. Solak and Erdem (2015) traced research of this kind back to 1995, when they were conducting a content analysis of VR-assisted foreign language learning. However, a careful analysis of some research reveals that what they thought of XR at that time was very different from how it is understood today. For example, the understanding of virtual devices' applications in learning was comparatively simple in the past, so the use of mobile headset virtual devices was considered an application mode of virtual devices (Tai and Chen, 2021). However, with updated technology, many devices have now become obsolete or are no longer classified as XR.

The integration of XR into language learning has another characteristic: empirical research is mainstream now, and many studies are experimental or quasi-experimental, which constitutes an important basis for the present paper. Experimental or quasi-experimental studies are often characterized by highly standardized procedures, strict control of irrelevant variables, high reliability and validity of measurement tools, and high repeatability of research conclusions. Huang et al. (2020) is an example of typical research on this topic. The team developed a Chinese writing learning system that was based on spherical video-based virtual reality (SVVR), and they carried out experiments in a senior high school writing class. They found that compared with the control group, whose learning was supported by conventional technology, the experimental group that used the SVVR system had better writing performance and self-efficacy. The team took a further step to measure the cognitive loads of the two groups, discovering that technology-supported learning can considerably reduce the cognitive loads of learners and boost their learning performance; Nicolaidou et al. (2021) also applied normative randomized control-group pretest-posttest design, indicating that VR applications can bring better immersion and engagement to learners through between-group comparisons (*t*-test). Moreover, it was the advantages of the application of technology in language learning that attracted the most attention of the researchers, and studies have revealed that the advantages include enhancing learners' learning motivation, changing their learning attitude, and improving their academic performance (Wehner et al., 2011; Qiu et al., 2021; Tai and Chen, 2021).

Another prominent characteristic of the application of XR in language learning is its rapid growth in recent decades, which can be understood through a comparison of the following two studies. When Solak and Erdem (2015) were reviewing this research topic, they discovered that from 1995 to 2015, only 40 papers were published on applying XR in foreign language learning and teaching. However, Qiu et al. (2021) found that there had been up to 150 relevant studies published between 2008 and 2019. A systematic review of the literature revealed that the application of XR was mainly conducted in higher education. This seems easy to explain because colleges and universities have easier access to these technologies, and researchers are often university faculty, which facilitates their study of the topic.

Other scholars have conducted meta-analyses in studies relating to this topic. For example, Wang et al. (2020) reviewed and examined relative research on Three-Dimensional Virtual Worlds (3DVWs, a 3D game) in language learning between 2008 and 2019, including 13 articles, and found that the application of 3DVWs dramatically improved students' attitudes and self-efficacy. Although 3DVWs are not exactly the same as the VR technology explored in this study (3DVWs are not considered VR technology without the use of a wearable VR device), Wang et al. (2020) was the first meta-analysis conducted on emerging technologies intervening in language learning and published in two authoritative indexes of the Web of Science core collection; therefore, it is of great significance for this study. Furthermore, it changes the situation where the only qualitative method was used to describe and summarize the study of XR applications in language learning, and it is another example of a successful meta-analysis conducted on a cutting-edge topic. Based on the above review, we find that there have been a large number of empirical studies on the use of XR in the field of language learning—enough to support a complete and systematic meta-analysis. Although some scholars have also conducted a meta-analysis (Wang et al., 2020), previous studies have not included both VR and AR in the investigation. Meanwhile, we are also concerned that a series of experimental studies closely related to the use of XR technology in language learning emerged from 2020 to 2022 (Chen and Hwang, 2020; Huang et al., 2020; Lee, 2020; Binhomran and Altalhab, 2021; Chen et al., 2021; Lai and Chen, 2021; Nicolaidou et al., 2021; Tai and Chen, 2021; Xie et al., 2021; Tai et al., 2022); therefore, it is necessary to implement a complete and systematic meta-analysis that includes this most recent work. The present study analyzed the research regarding the effectiveness of XR-assisted language learning in a quantitative manner, and confirmed the moderator variables that affect the validity of XR to better guide its future application in language learning.

Materials and methods

Definition and application of meta-analysis

The present study used a meta-analysis to analyze the effect of XR technology on language learning. Meta-analysis belongs to a branch of evidence-based research, and is capable of enhancing credibility by increasing sample size to resolve the inconsistencies of research results (Glass, 1976; Oswald and Plonsky, 2010). As a special method of systematic review, meta-analysis has the following three advantages (Lipsey and Wilson, 2001; Cooper, 2015):

1. Meta-analysis is a systematic review of the existing research, and is able to shed deep insights into the hypothesis, process, and conclusion of the included research.
2. Meta-analysis takes account of the strength of the effect in every case of empirical research, which can better satisfy the requirement for data, therefore offering a more convincing conclusion.
3. Meta-analysis employs a programmed method to deal with the information in a vast literature, enabling it to consider both the depth and the range of the studies' contents.

Meta-analysis was first applied in clinical medicine and social psychology (Smith and Glass, 1977; Rosenthal and Rubin, 1978), and was then introduced in the field of education (Glass and Smith, 1979). Today, meta-analysis is widely applied to language learning. For example, Lin (2014) used it to study the impact of computer-mediated communication (CMC) on language learning; Plonsky and Ziegler (2016) used second-order analysis to explore the relationship between computer-assisted language learning (CALL) and second language acquisition (SLA). These successful cases demonstrated the application value of meta-analysis in education as well as in language learning. Thus, in this study, we sought to further develop educational technology by employing a meta-analysis to identify the impact of XR on language learning outcomes.

The software we used to carry out the meta-analysis was Comprehensive Meta-Analysis (CMA, version 3.0, developed by Wilson et al. in the United States and United Kingdom). To ensure that the research is scientific and precise, the present study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), which include a literature search, literature screening, and data coding.

Search strategy

According to the objective of this study, our team retrieved articles on the impact of XR on language learning outcomes from the Web of Science core collection (to ensure the research quality, only SCIE and SSCI indexes were selected from the Web of Science database) and Scopus databases between 2000 and 2022. After reviewing several relevant studies (Luo et al., 2021a; Villena-Taranilla et al., 2022) and several attempts, the final retrieval strategy was determined as follows:

TS = (("VR" OR "virtual reality" OR "AR" OR "augmented reality" OR "MR" OR "mixed reality" OR "XR" OR "extended reality") AND ("language learning")).

As some articles were duplicated in the two databases, our research team deleted them when incorporating the data.

Ultimately, 470 articles were procured before proceeding to the next stage.

Inclusion criteria and assessment of quality

Meta-analysis should follow strict criteria when screening the included literature (Moher et al., 2009; Xu et al., 2022). First, the articles must involve quantitative empirical research, and literature that does not meet the requirements, such as review research, narrative studies, and qualitative research, should be excluded. Second, the research topic needs to be closely related to both XR technology and language learning. Third, the research method should be a rigorous and scientific experimental method. Finally, studies need to report the statistics necessary to calculate the overall effect size. The details of these research criteria can be found in Table 1.

When implementing the inclusion criteria, the present study referred to the execution process of Luo et al. (2021a), and combined it with the process of PRISMA to complete this step. First, three research team members examined the titles and abstracts of articles independently, and excluded 376 articles that were irrelevant to the research topic. When there was any disagreement during this stage of the process, the three researchers would discuss it and vote. Then, we carefully read the remaining articles and discarded 77 articles that did not fulfill the inclusion criteria in terms of the research methods used and data provided. This process was jointly completed by three research team members to ensure the reliability of the data screening process. In the above two rounds of screening, the consistency of the three team members was higher than 95%. After completing the above screening process, 17 articles were included in the meta-analysis (Figure 1).

After selecting the articles, the quality of the published research needed to be carefully evaluated. Since the 17 articles were published in academic journals or conference collections, they had undergone rigorous double-blind peer review, which ensured the quality of the research. To further ensure the quality of the study, we invited three professors who are experts in meta-analysis to review the 17 articles, mainly focusing on the rationality of the experimental design and the adequacy of data reporting. Finally, all three professors agreed on the quality of the studies, indicating that they met the requirements of a meta-analysis.

TABLE 1 Inclusion criteria for articles.

Category	Inclusion criteria
Literature type	Quantitative empirical research; literature that does not meet the requirements, such as review research, narrative studies, and qualitative research, should be excluded
Research topic	The effects of XR, such as VR, AR, MR, etc., on language learning
Research method	The experimental method is the only approach, and there have to be two types of experiments. One is a controlled experiment, in which XR is applied to language learning in the experimental group, but not in the control group. The other is setting a single group, which means that the same group will be given a pre-test and then a post-test after using XR for language learning
Data	The data provided in the article are enough to calculate the effect sizes, such as the sample size (N), Mean, standard deviation (SD), t-value, value of p, etc.
Access method	The article can be accessed *via* the internet

Data coding and descriptive statistics

To facilitate statistical analysis and the calculation of effect size, we needed to code the features of the literature. The present study takes the following features into account:

- Characteristics of the article (title, year of publication)
- Population characteristics (age, educational level)
- Experimental design (number of subjects, technology type, target language, language skill)
- Experimental results (mean, standard deviation, value of p).

The above features were independently extracted and coded by two authors. According to the coding standard proposed by Lipsey and Wilson (2001), the coding of a meta-analysis should take independent samples as the coding unit; that is, if there are multiple independent samples in a study, we code them separately and yield multiple independent effect sizes. Next, the 17 articles consisting of 21 sets of independent samples were studied. Any problem encountered was solved by discussion and voting by the three team members. The consistency of coding was higher than 98%, indicating that the coding results had good reliability and validity.

After the coding, the research team found that the final 17 articles, with a total of 993 subjects included in the meta-analysis, were mainly concentrated in terms of publication year: they were mostly published after 2010, especially in the most recent 5 years (2017–2021). As for the type of XR used, VR was applied in 11 studies, and AR was examined in six studies. Studies before 2016 all used VR, and since then, AR was gradually adopted in the research. The educational level of subjects covered a wide range,

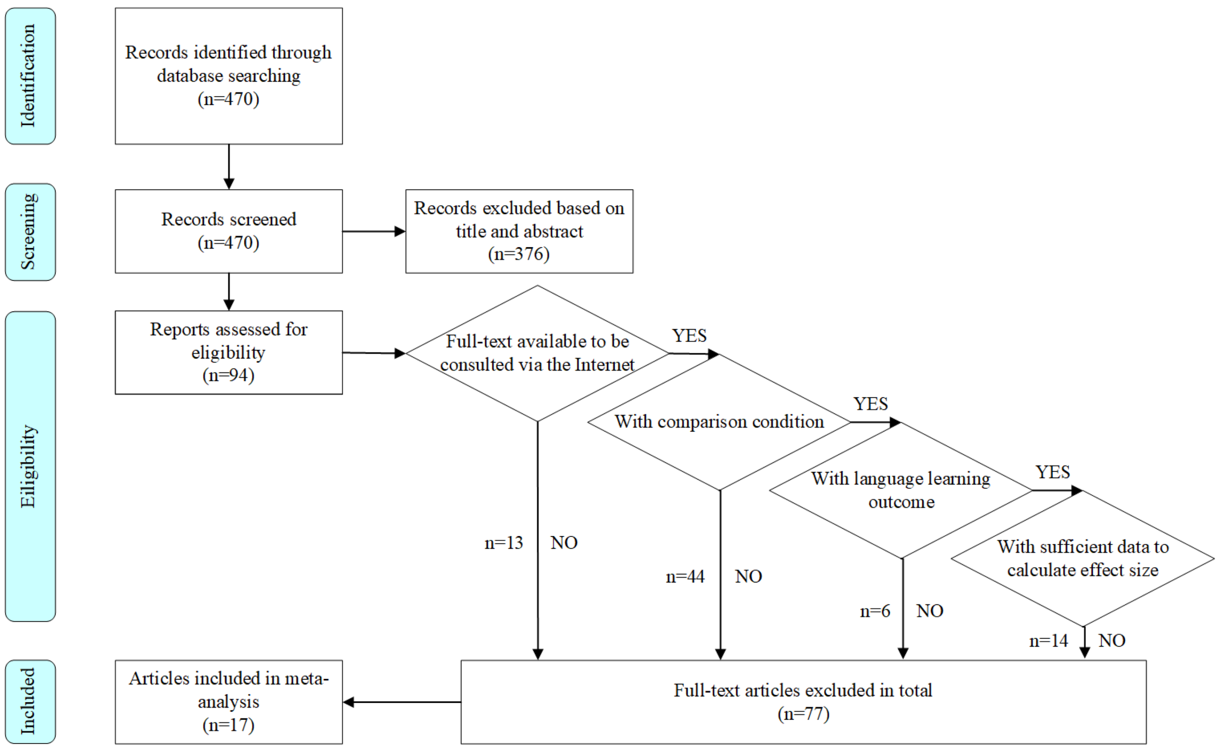


FIGURE 1
Flowchart for the article screening process.

from kindergarten to university. Among them, subjects in primary schools and universities were most frequently studied by researchers, appearing in seven and six studies, respectively. There were 13 studies that selected English as the target learning language, which may be explained by the fact that English is the most widely used language in the world.

Results and discussion

After the preliminary literature reorganization and data coding work, the CMA software was used to perform detailed analyses in sequence: the effect value calculation, heterogeneity test, sensitivity analysis, publication bias analysis, and subgroup analysis.

Effect size calculation

Effect size is an important indicator that estimates the magnitude of the effect or association between two or more variables (Snyder and Lawson, 1993). We used Hedges' *g* (1981) because, compared with the method of calculating effect size brought up by Cohen (1969) and Glass (1976), this method is more suitable for small sample sizes. Based on Cohen's *d* (1969), the effect sizes of 0.2, 0.5, and 0.8, respectively, corresponded to a small effect, medium effect, and large effect.

Heterogeneity test

As the experimental designs and characteristics of research subjects differed across the included articles, the findings on VR's impact on language learning would be affected, thus there was likely to be heterogeneity in the measured effect sizes. A heterogeneity test was conducted to determine whether to perform a moderator analysis to examine the source of heterogeneity, on the one hand, and to decide whether to adopt a fixed-effects model or a random-effects model, on the other hand. Heterogeneity was tested using *Q*-tests. The results showed that $Q = 177.356$ ($p < 0.001$) and $I^2 = 88.723\%$; the latter indicated that 88.723% of heterogeneity came from the differences in effect sizes. According to the standard set by Higgins et al. (2003), when I^2 is greater than 75%, this suggests a high degree of heterogeneity. Therefore, we adopted a random-effects model, and the summary effect size obtained was 0.825 ($p < 0.001$), indicating a large effect (Cohen, 1969). Figure 2 presents the forest plot of the study. The results suggest that the application of XR has a considerable effect on language learning. Figure 2 also shows that the 95% confidence interval is far higher than zero, which suggests that a large effect resulted from the message reflected in the research data itself, but not from random error.

Based on the above analyses, we can conclude that XR can significantly boost the learning effect, a finding that is consistent

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Lai, 2021	0.574	0.363	0.132	-0.137	1.286	1.582	0.114
Solak, 2016	1.642	0.293	0.086	1.067	2.217	5.599	0.000
Xie, 2019	0.308	0.397	0.157	-0.470	1.085	0.776	0.438
Nicolaidou, 2021	-0.324	0.312	0.097	-0.936	0.287	-1.039	0.299
Chen, 2019	1.132	0.216	0.047	0.708	1.555	5.234	0.000
Chen, 2021	0.865	0.318	0.101	0.241	1.489	2.717	0.007
Tai, 2021	0.495	0.237	0.056	0.031	0.959	2.091	0.037
Lee, 2020	-0.672	0.312	0.097	-1.284	-0.061	-2.156	0.031
Chen, 2020	2.236	0.265	0.070	1.716	2.756	8.433	0.000
Wen, 2018	2.694	0.408	0.166	1.894	3.493	6.602	0.000
Binhomran, 2021	1.080	0.248	0.062	0.593	1.567	4.346	0.000
Huang, 2020	0.455	0.212	0.045	0.040	0.870	2.151	0.031
Khatoony, 2019	5.472	0.723	0.522	4.056	6.888	7.573	0.000
Lin, 2019	0.857	0.256	0.066	0.355	1.360	3.344	0.001
Vazquez, 2018	-1.039	0.363	0.132	-1.750	-0.328	-2.863	0.004
Chung, 2012 Ⅰ	0.540	0.226	0.051	0.098	0.982	2.395	0.017
Chung, 2012 Ⅱ	0.640	0.227	0.052	0.195	1.085	2.817	0.005
Chung, 2012 Ⅲ	0.421	0.224	0.050	-0.018	0.860	1.880	0.060
Chung, 2011 Ⅰ	0.599	0.245	0.060	0.118	1.079	2.442	0.015
Chung, 2011 Ⅱ	0.715	0.247	0.061	0.230	1.200	2.890	0.004
Chung, 2011 Ⅲ	0.500	0.244	0.059	0.022	0.977	2.052	0.040
	0.825	0.179	0.032	0.474	1.176	4.609	0.000

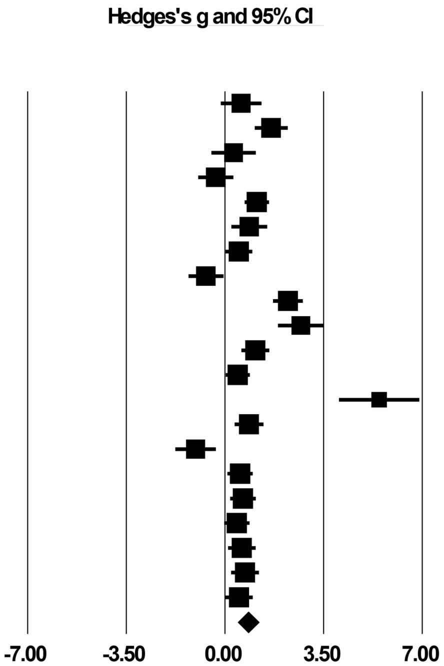


FIGURE 2
Forest plot for the impacts of XR on language learning.

with the conclusions of much comprehensive research (Solak and Erdem, 2015; Qiu et al., 2021). Many scholars gave reasonable explanations for this. For example, as XR assumes the characteristics of deep immersive interaction and high involvement in learning, it is able to break the limitations of the traditional media applied in education to provide language learners with a realistic simulated language learning environment, and to effectively support their language learning (Pinho et al., 2008; Nicolaidou et al., 2021).

Sensitivity analysis

Sensitivity analysis can help the meta-analysis practitioner assess the confidence of their results (Elvik, 2005). The sensitivity analysis was performed using the One Study Removed method in CMA 3.0. After removing each included study, one by one, the effect size was merged to observe whether it deviated. The results are shown in Figure 3. No matter which study was excluded, the offset of the effect size was relatively stable [in the interval of (0.686, 0.906)], which ensured the robustness and reliability of the meta-analysis.

Publication bias analysis

Amid a systematic review, the presence of publication bias means that the published articles cannot fully represent the overall study results in a field (Rothstein et al., 2006). To examine publication bias in this study, we used the funnel plot and the fail-safe *N*. As shown in Figure 4, the effect sizes are evenly distributed on the two sides of the aggregating effect size, preliminarily indicating that there is no serious publication bias in the selected studies. Because the funnel plot is only an initial and subjective test, the fail-safe *N* is needed to perform a further analysis. Rosenthal (1991) proposed that if $N > 5k + 10$, publication bias is not likely to affect the results of a meta-analysis. The *N* of this study is 882, which is much larger than 115 ($5 \times 21 + 10$), and is within the average range. Based on the analysis above, there is no significant publication bias in this study, which means that the results of the meta-analysis are robust and reliable.

Subgroup analysis

Due to the high degree of heterogeneity in the samples, there is likely to be significant moderator variables. Thus, a subgroup analysis was conducted to determine the source of the heterogeneity, and to study the moderation effects of the sample characteristics (i.e., educational level, target language, technology type, language skill) on effect size.

Educational level

As shown in Table 2, according to the *Q*-statistics, the educational level of learners and XR does not have a significant impact on language learning achievement ($Q = 4.096$, $p = 0.129$). Apart from that, it is clear that, supported by XR, primary school learners achieved strong learning effects ($g = 1.131$, $p < 0.001$), as

Statistics with study removed

Study name	Point	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Lai, 2021	0.839	0.186	0.035	0.474	1.203	4.510	0.000
Solak, 2016	0.783	0.182	0.033	0.426	1.140	4.296	0.000
Xie, 2019	0.850	0.185	0.034	0.487	1.212	4.595	0.000
Nicolaidou, 2021	0.881	0.181	0.033	0.526	1.235	4.868	0.000
Chen, 2019	0.812	0.189	0.036	0.442	1.182	4.298	0.000
Chen, 2021	0.825	0.187	0.035	0.459	1.191	4.416	0.000
Tai, 2021	0.846	0.189	0.036	0.475	1.217	4.472	0.000
Lee, 2020	0.895	0.176	0.031	0.550	1.239	5.085	0.000
Chen, 2020	0.743	0.170	0.029	0.410	1.075	4.372	0.000
Wen, 2018	0.736	0.173	0.030	0.397	1.074	4.259	0.000
Binhomran, 2021	0.814	0.188	0.035	0.446	1.183	4.331	0.000
Huang, 2020	0.849	0.190	0.036	0.476	1.222	4.464	0.000
Khatoony, 2019	0.686	0.159	0.025	0.374	0.998	4.311	0.000
Lin, 2019	0.826	0.189	0.036	0.457	1.196	4.380	0.000
Vazquez, 2018	0.906	0.173	0.030	0.567	1.245	5.234	0.000
Chung, 2012 Ⅰ	0.844	0.190	0.036	0.472	1.216	4.445	0.000
Chung, 2012 Ⅱ	0.839	0.190	0.036	0.466	1.211	4.412	0.000
Chung, 2012 Ⅲ	0.850	0.189	0.036	0.479	1.221	4.490	0.000
Chung, 2011 Ⅰ	0.840	0.189	0.036	0.470	1.211	4.443	0.000
Chung, 2011 Ⅱ	0.834	0.189	0.036	0.463	1.205	4.409	0.000
Chung, 2011 Ⅲ	0.845	0.189	0.036	0.475	1.216	4.476	0.000
	0.825	0.179	0.032	0.474	1.176	4.609	0.000

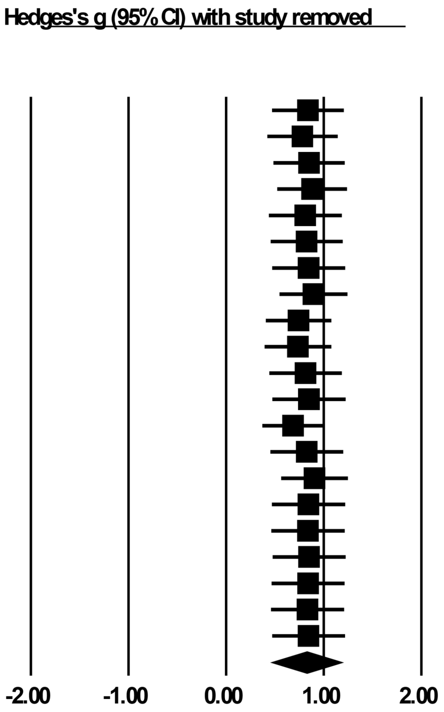


FIGURE 3
Results of the sensitivity analysis.

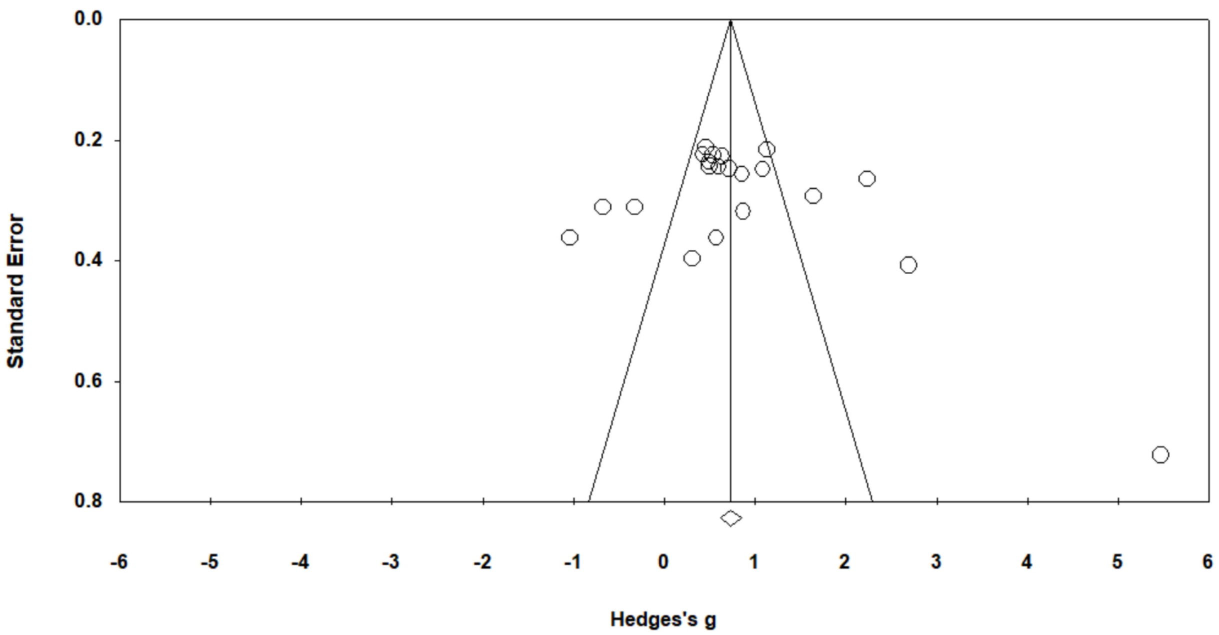


FIGURE 4
Funnel plot with effect sizes.

did middle school learners ($g = 0.646$, $p < 0.001$). When it comes to college learners, however, XR did not emerge as a significant booster of learning ($g = 0.238$, $p = 0.658$).

To figure out the reason why college learners were not significantly influenced, we scrutinized a number of studies of this population. The analysis found that many learners were quite

TABLE 2 Impacts of XR on the learning effects, by education level.

Educational level	*N*	*g*	*SE*	95% *CI*		*Z*	*p*	Between-group effects
				Lower	Upper			
Primary education	11	1.131	0.229	0.682	1.579	4.939	< 0.001	$Q = 4.096$, $p = 0.129$
Secondary education	3	0.646	0.157	0.338	0.953	4.115	< 0.001	
Higher education	6	0.238	0.537	−0.814	1.290	0.443	0.658	

N, number of samples; *g*, Hedges' g (effect sizes); *SE*, standard error; *CI*, confidence interval.

interested in XR, which received more attention than the learning material itself (Vázquez et al., 2018). In experimental studies, this phenomenon is similar to the Hawthorne effect. Vázquez et al. (2018) confirmed the above analysis: after a week spent familiarizing themselves with the XR device, the participants again conducted the experiment, with the experimental group using the VR; their learning effect began to grow significantly higher than that of the control group, who used the traditional learning method. Such a conclusion suggests that the cognitive cost, which resulted from learners' unfamiliarity with the VR interface, can possibly account for the poor learning effect at the start. Similar explanations are mentioned in another study (Cheng et al., 2017). Meanwhile, Vázquez et al. (2018) also found that college freshmen were more accustomed to the traditional mode of learning (because they are used to using traditional education approaches in K–12 schooling), so in the face of the new technology being used as a new teaching mode, these individuals may be unable to adapt to the new technology, resulting in a poor learning effect.

Lee (2020) provide a different explanation. Their team discovered that merely applying emerging technologies would not necessarily promote students' engagement and academic performance, and that innovative instructional design and new teaching principles were the right paths to foster language learning development. Their work offers a new perspective on technology-based study of language learning, reminds us to adopt a rational and objective attitude toward technology, and calls for a reexamination of language teaching that is integrated with technologies today (Luo et al., 2021a).

Xie et al. (2021) also analyzed and explained this problem. They argued that the plasticity of learners' language learning ability would be greatly reduced after they become adults, especially in terms of pronunciation and fluency, and it was difficult for adult college students to make significant progress in language learning in a short period of time with the intervention of XR technology. Solak and Cakır (2016) confirmed this, suggesting that the application of AR technology can significantly improve the satisfaction and enthusiasm of young learners. Another explanation mentioned by Xie et al. (2021) was that learners in primary school and middle school had higher self-consciousness compared with college learners; in some studies that lasted for a long time, learners in primary and middle school would continue to learn after class, which swayed the results to indicate that learners who did not use the XR technology were more likely to have significant differences than those who used the XR technology in primary and middle school.

In fact, as explained by Vázquez et al. (2018) and Lee (2020), we can improve the experimental design to obtain some more stable and practically instructive studies. To eliminate participants' inattention due to technical novelty, the learners can be exposed to the new technology environment for a period of time, helping them adapt to the equipment and the intervention learning mode (Gavgani et al., 2018). On this basis, we can eliminate some unnecessary interference factors, or we can use the Latin square design to exclude the above interference factors to reduce the influence of endogenous factors generated by the research design (Richardson, 2018). Undoubtedly, the research on this issue needs to be supplemented by more empirical studies.

Target language

It can be seen from the *Q*-statistics that there is a notable influence of the moderating effect of different target languages on the learning effect ($Q = 16.128$, $p < 0.001$; Table 3). When the target language is English, the application of XR could considerably promote the learning effect ($g = 0.904$, $p < 0.001$). There is no significant result for the learning of Chinese or for two other languages (Italian and Spanish) ($g = 1.499$, $p = 0.209$ and $g = -0.657$, $p = 0.065$, respectively).

XR can effectively promote English language learning, a conclusion that was supported by many studies (Khatoony, 2019; Chen and Hwang, 2020; Lee, 2020; Tai and Chen, 2021). The technical forms supporting English language learning were also very plentiful. In addition, the research on XR in English language learning started earlier and lasted a long time. In the long accumulation and practice, many teaching models and methods suitable for English teaching were formed, which undoubtedly improved the effect of English language learning after the technology intervention.

When reviewing studies of Chinese and Italian learners, we discovered that the research had a more detailed classification of language learning compared to work on English learners. For example, Xie et al. (2021) divided the language learning effect into five dimensions: content, fluency, vocabulary, pronunciation, and grammar. The results turned out to be vastly different. On content and vocabulary, the experimental group that used VR technology obtained notable greater learning performance than the control group, but there was no marked significance between the two groups in terms of fluency, pronunciation, or grammar.

TABLE 3 Impact of XR on the learning effects, by target language.

Target language	*N*	*g*	*SE*	95% *CI*		*Z*	*p*	Between-group effects
				Lower	Upper			
Chinese	2	1.499	1.193	−0.839	3.837	1.256	0.209	$Q = 16.128$, $p < 0.001$
English	17	0.904	0.172	0.567	1.241	5.262	< 0.001	
Other languages	2	−0.657	0.356	−1.356	0.041	−1.845	0.065	

N, number of samples; *g*, Hedges' g (effect sizes); *SE*, standard error; *CI*, confidence interval.

TABLE 4 Impacts of XR on the learning effects, by technology type.

Technology type	*N*	*g*	*SE*	95% *CI*		*Z*	*p*	Between-group effects
				Lower	Upper			
VR	15	0.709	0.205	0.308	1.111	3.462	0.001	$Q = 0.878$, $p = 0.349$
AR	6	1.101	0.387	0.567	1.814	3.024	< 0.01	

N, number of samples; *g*, Hedges' g (effect sizes); *SE*, standard error; *CI*, confidence interval.

After synthesizing various studies using a meta-analysis, we found that VR does not exert a great impact on learning in general. Still, we cannot arbitrarily assume that XR will not play a prominent role in the learning of Chinese or Italian in all respects. It is important that we carefully consider the question: On which language skills does XR have an impact? Further discussion and analysis of this is provided in section "language skill".

In fact, as all the included studies target the learning of a single language, it is hard to tell the difference in learning effects between languages and how different languages impact technology-integrated learning from the conclusions of these studies. However, we suppose that the features of diverse languages would, to some degree, moderate the impact of XR on language learning; in addition, the differences in the learning effects of different target languages may also come from the maturity of XR in the language learning field. More empirical research should analyze and explain the reason and mechanism behind this observation.

Technology type

The samples of this study contained VR and AR technologies, which both significantly contribute to language learning outcomes (Table 4). Yet, it is noteworthy that they differ in terms of the degree of their impact on learning. VR has a huge impact on language learning ($g = 1.101$, $p = 0.001$), while AR's impact is above average ($g = 0.709$, $p < 0.01$). Meanwhile, the different technologies do not show a significant difference in the moderating effect of language learning ($Q = 0.878$, $p = 0.349$).

Based on analyses of several typical studies (Solak and Cakır, 2016; Binhomran and Altalhab, 2021; Lai and Chen, 2021), we may assume that the difference in the impact on learning between VR and AR, to a large extent, originates from the characteristics of the technology itself. Compared with VR, AR emphasizes a real sense of presence and is therefore believed to help facilitate learning and visualization of abstract concepts in the mind (Gün, 2014). But it is a pity that only one XR technology was applied in all included studies, and there is a dearth of research on the difference between VR and AR, which makes it impossible for us to accurately determine the exact characteristic of the technology that leads to the different impacts on learning. The answer to this problem awaits further explanation from future empirical research.

Although the analysis in this present study cannot evaluate the differences between VR and AR when applied to language learning, some common factors of the two can be analyzed. Lee (2020) not only proved that XR would have better effects on language learning through experimental research, but also further analyzed the internal mechanism through interviews. In this setting, a student said that XR made him interested in exploratory learning, while in the traditional learning mode, information can only be learned and obtained by reading. In other words, the realistic situations provided by XR were the key for learners to achieve better learning outcomes. Of course, some studies found that the advantages of XR also consisted of improving students' understanding and enthusiasm for learning (Binhomran and Altalhab, 2021). In addition, previous studies have demonstrated that XR techniques can significantly improve memory retention time compared with traditional research methods (Pérez-López and Contero, 2013). As language learning was a learning activity requiring memory, this feature of XR matched the properties of language learning activities, which gave the application of XR in language learning extraordinary advantages and unique value.

Finally, it is worth noting that any technology intervention requires an adaptation period for the learners, especially for some complex technologies, and some even require corresponding training before use (Binhomran and Altalhab, 2021). To our relief, however, almost all articles included in the present research have considered this in the design of their experiments. In addition, teachers must pay attention to this point so that students' knowledge and skills can benefit from the new technology and some of the adverse effects caused by technical barriers can be reduced.

Language skill

Language skills were classified into five categories (vocabulary, speaking, grammar, reading, listening) according to the descriptions in the articles. The *Q*-statistics revealed no significant differences in the moderating effects of different language skills on language learning ($Q = 5.346$, $p = 0.375$; Table 5). Moreover, the learning of the five language skills is noticeably influenced by XR. Among them, speaking is most significantly promoted ($g = 2.100$, $p < 0.01$), and vocabulary shows a high degree of impact ($g = 0.762$, $p < 0.01$). The remaining three skills (grammar, listening, reading) are moderately impacted ($g = 0.590$, $p < 0.001$; $g = 0.495$, $p < 0.05$; $g = 0.457$, $p < 0.01$). The influence of XR on the five language skills occurred from the largest to the smallest as follows: speaking, vocabulary, grammar, listening, and reading.

Speaking has been noted by a number of studies (Khatoony, 2019; Chen et al., 2021; Xie et al., 2021) to have been aided by the integration of XR into language learning. This phenomenon is not difficult to understand given the fact that XR is famous for its authenticity and immersion, and that the realistic experience provides an excellent practice environment for speaking. Vocabulary is also greatly enhanced by XR technology, as expected; Pérez-López and Contero (2013) demonstrated that the intervention of XR can significantly improve learners' memory retention ability, so it is undoubtedly a great benefit to vocabulary memorization skills.

To learn more about the differences in the learning effects of diverse language skills, Xie et al. (2021) carried out systematic qualitative research based on experimental studies, and gave insightful explanations of their conclusions in the form of interviews. Quantitative research informs the relationship between research variables, while qualitative research provides insights into the causal logic behind relationships *via* case-by-case in-deep analysis. Through the interviews, Xie et al. (2021) found that, for some language skills, such as grammar and reading, it is difficult to make considerable improvement in a short time when a certain stage is reached; moreover, a large amount of input from a native speaker is needed to enhance these skills. Vocabulary, by contrast, is a kind of language skill that can be enhanced rapidly with the help of technology. This explains why the learning effects of grammar, listening, and reading were lower than that of speaking and vocabulary. The above conclusion is also supported by other studies (e.g., Bongaerts et al., 1997).

Examining the differences in language skills after technology intervention can provide learners and teachers with occasions to apply extended realistic technology and to avoid some adverse effects caused by the abuse of technology (such as blindly pursuing technology, but not paying attention to learning methods and teaching methods). Many scholars have pointed out that the greatest value of XR for language learning is to provide immersive realistic situations so that learners can conduct language learning as if they were in a foreign language community (Chung, 2011, 2012; Xie et al., 2021). XR is of great value in improving speaking, but this study also reminds us that we should not rely too much on XR in the practice of reading and other skills. We should also maintain a rational attitude toward emerging technologies, and actively explore alternative models that are conducive to improving learners' reading and other skills (Luo et al., 2021a).

Conclusion and prospects for future research

Conclusion and implications

The present study uses a systematic meta-analysis to review the application of XR in language learning over the past 20 years, reveals the learning effects of technology intervention in language learning, and answers the three questions raised in the introduction. First, through the comprehensive calculation of the effect size of the included articles, XR is confirmed to have a large positive effect on learning (effect size = 0.856, $p < 0.001$). Second, adjusting the analysis of the variables, such as education level, this study finds that the target language is an important adjustment variable affecting the results of the study, and different techniques have different benefits for language learning. This demonstrates that different languages and technologies have different features, and we need to select technologies according to the features of the language itself to make them match. Some researchers also point out that technology and its derivative products need to be designed according to language characteristics (Xie et al., 2021). This suggestion is undoubtedly of great significance. Finally, the present study also points out some meaningful research concerns, including:

TABLE 5 Impacts of XR on the learning effects, by language skill.

Language skill	*N*	*g*	*SE*	95% *CI*		*Z*	*p*	Between-group effects
				Lower	Upper			
Vocabulary	9	0.762	0.287	0.199	1.325	2.652	< 0.01	$Q = 5.346$, $p = 0.375$
Speaking	4	2.100	0.771	0.588	3.612	2.723	< 0.01	
Reading	2	0.457	0.165	0.134	0.780	2.773	< 0.01	
Listening	1	0.495	0.237	0.031	0.959	2.091	< 0.05	
Grammar	3	0.590	0.131	0.333	0.847	4.495	< 0.001	

N, number of samples; *g*, Hedges' g (effect sizes); *SE*, standard error; *CI*, confidence interval.

1. How does the difference between VR and AR influence the learning effect of language learning, which helps us to choose which technology to use? There is a lack of comparative study of VR and AR in the existing research on language learning.
2. The relationship between the language learning effects presented by different language types and the characteristics of the language itself needs to be further analyzed. Existing studies are all focused on a single language, which is valuable but insufficient. If necessary, the mother language can also be included as an influencing factor.
3. The relationship between the differences in language learning effects and the characteristics of education levels has not been explored. Existing studies only examine a single learning stage, and further empirical studies are needed.

The present study highlights some practical implications for the use of XR in language learning. First, based on the discussion of the results, blind application of these emerging technologies may not necessarily bring about the desired results (Xie et al., 2021). We need to understand how these technologies improve learners' performance after they are involved in language learning, and mastering these mechanisms can help these technologies to play a better auxiliary role. For example, XR has remarkable advantages in improving speaking, which is closely related to the sense of presence and immersion provided by XR (Radianti et al., 2020). Second, educators should evaluate the maturity of relevant technologies in the application field and take it as a condition for whether to apply relevant technologies. In the subgroup analysis of language types, experimental studies with English as the target language often have better learning effects, although there are factors of language characteristics, it is also closely related to whether the technology is mature in terms of product design in other language teaching fields. The bold application of immature technology in language learning may bring adverse effects. In addition, some scholars have emphasized some principles of the use of emerging technologies in empirical studies, pointing out that what really matters is not the use of these technologies, but the use of innovative teaching principles in language teaching (Lee, 2020), which is of great practical significance for front-line teaching staff. It is necessary for educators to examine the existing technology rationally and objectively and to introduce it into teaching activities scientifically and reasonably. Finally, XR technologies still face some user experience issues (Luo et al., 2021a), such as physical discomfort of learners, security risks, and low technical stability. These factors need to be considered when the application of XR is extended.

Limitations and prospects for future research

Although the effects of systematic error and random error on the conclusion have been accounted for as much as possible, there are still some design flaws and other factors that affect the reliability of this research. First, the present study only takes learning outcomes into consideration, without paying attention to the emotions, attitudes, and psychological state of learners. In addition, as the included studies did not explore the development of higher-order thinking in language acquisition, the present study did not conduct a detailed analysis of this topic.

Second, as analyzed in the present study, the design of the experiment may affect its conclusions, and the nature of the experimental design (i.e., that it is conducted over a short period of time) may influence the learning effects because students are not very familiar with the XR equipment. Among the included literature, some studies were long experimental studies, lasting up to 5 or 6 weeks (e.g., Chen and Hwang, 2020; Binhomran and Altalhab, 2021), and some were as short as 1–2 h (e.g., Huang et al., 2020). Future work might account for this difference.

Third, during the study, a series of reliability tests, which include the heterogeneity test, sensitivity analysis, and publication bias analysis, were undertaken, and the effect of publication bias has been proven non-significant on a technical level. Even so, since the study focused on published literature, if the primary researcher gave up publishing a paper due to the non-significance of the findings, or an editor thought that it was meaningless to publish an insignificant study, publication bias as such would noticeably impact the reliability of a meta-analysis's conclusions (Coursol and Wagner, 1986).

In spite of the above unavoidable problems, we have reason to believe that the integration of XR into language learning will continue; thus, related studies require constant evaluation and updates, which has also been recognized by many researchers (Chen and Hwang, 2020; Luo et al., 2021a,b). The limitations pointed out above need to be addressed in the future. For example, we can expand the data source, perhaps by including ERIC and other highly related databases, or we can include in the search databases that are highly relevant to pedagogy (e.g., ERIC). By doing so, empirical research on more languages could be accessed to make the subgroup analysis of language types more reliable. In terms of publication bias, over recent years this kind of cognitive bias has been reduced dramatically with the efforts of some scholars who have conducted meta-analyses (Halpern et al., 2005). Therefore, with the improvement of methodology and research design, scholars will be able to reach more accurate conclusions on this topic.

Data availability statement

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

Author contributions

All authors contributed to the conception of the idea, screening the literature, analyzing the data, and writing the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the General Scientific Research Projects of Zhejiang Provincial Department of Education (grant no. GZ20491080003), Education Science Planning Project of Zhejiang Province (grant no. 2021SCG236), Philosophy and Social Science Planning Project Fund Project of Zhejiang Province (grant no. 17GXSZ19YB), Teaching Reform Project of Zhejiang University of Technology (grant no. JG2022064), and National University Student Innovation and Entrepreneurship Training National Project of China (grant nos. 202210337027 and 202210337003).

Acknowledgments

The authors would like to express their sincere gratitude to the reviewers and editor for their valuable suggestions, to Chengliang Wang for his careful guidance, to Zhejiang University of Technology for providing places for learning and researching.

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

Binhomran, K., and Altalhab, S. (2021). The impact of implementing augmented reality to enhance the vocabulary of young EFL learners. *JALT CALL J.* 17, 23–44. doi: 10.29140/JALTCALL.V17N1.304

Bongaerts, T., Van Summeren, C., Planken, B., and Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Stud. Second Lang. Acq.* 19, 447–465. doi: 10.1017/S0272263197004026

Brown, M., McCormack, M., Reeves, J., Brooks, D. C., and Grajek, S. (2020). *2020 Educause Horizon Report Teaching and Learning Edition*. Louisville: EDUCAUSE.

Chen, Y. L. (2016). The effects of virtual reality learning environment on student cognitive and linguistic development. *Asia-Pac. Educ. Res.* 25, 637–646. doi: 10.1007/s40299-016-0293-2

Chen, C. Y., Chang, S. C., Hwang, G. J., and Zou, D. (2021). Facilitating EFL learners' active behaviors in speaking: a progressive question prompt-based peer-tutoring approach with VR contexts. *Interact. Learn. Envir.* 29, 1–20. doi: 10.1080/10494820.2021.1878232

Chen, M. R. A., and Hwang, G. J. (2020). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interact. Learn. Envir.* 28, 1–21. doi: 10.1080/10494820.2020.1734626

Cheng, A., Yang, L., and Andersen, E. (2017). "Teaching language and culture with a virtual reality game," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. eds. A. A. Maimani and A. Roudaut (Denver: ACM), 541–549.

Chung, L. Y. (2011). "Using avatars to enhance active learning: integration of virtual reality tools into college English curriculum," in *Proceedings of the 16th North-East Asia Symposium on Nano, Information Technology and Reliability*. eds. C. T. Hon, D. K. Kae and F. Simon (Macao: IEEE), 29–33.

Chung, L. Y. (2012). "Virtual reality in college English curriculum: case study of integrating second life in freshman English course," in *Proceedings of the 2012 26th International Conference on Advanced Information Networking and Applications Workshops*. eds. B. Leonard, E. Tomoya, X. Fatos and T. Makoto (Fukuoka), 250–253.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. London: Academic Press.

Cooper, H. (2015). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*. London; Thousand Oaks: Sage Publications.

Coursol, A., and Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: a note on meta-analysis bias. *Prof. Psychol. Res. Pract.* 17, 136–137. doi: 10.1037/0735-7028.17.2.136

Elvik, R. (2005). Can we trust the results of meta-analyses? A systematic approach to sensitivity analysis in meta-analyses. *Transport. Res. Rec.* 1908, 221–229. doi: 10.1177/0361198105190800127

Feng, Q., Luo, H., Li, W., Chen, Y., and Zhang, J. (2021). The moderating effect of debriefing on learning outcomes of IVR-based instruction: an experimental research. *Appl. Sci.* 11:10426. doi: 10.3390/app112110426

Freeman, A., Becker, S. A., and Cummins, M. (2017). *NMC/CoSN Horizon Report: 2017 K-12 Edition*. Austin: The New Media Consortium.

Gavgani, A. M., Walker, F. R., Hodgson, D. M., and Nalivaiko, E. (2018). A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different? *J. Appl. Physiol.* 125, 1670–1680. doi: 10.1152/japplphysiol.00338.2018

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educ. Res.* 5, 3–8. doi: 10.3102/0013189X005010003

Glass, G. V., and Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educ. Eval. Policy Anal.* 1, 2–16. doi: 10.3102/01623737001001002

Gün, E. (2014). *Effects of Augmented Reality Applications on Students' Spatial Abilities. [Master's Thesis].* Ankara: Gazi University.

Halpern, S. D., Berlin, J. A., Rothstein, H. R., Sutton, A. J., and Borenstein, M. (2005). "Beyond conventional publication bias: other determinants of data suppression," in *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. eds. H. R. Rothstein, A. J. Sutton and M. Borenstein (Hoboken: John Wiley & Sons), 303–317.

Han, X., Liu, Y., Li, H., Fan, Z., and Luo, H. (2020). "Augmenting the makerspace: designing collaborative inquiry through augmented reality" in *Proceeding of the International Conference on Blended Learning*. eds. S. K. S. Cheung, R. Li, K. Phusavat, N. Paoprasert and L. Kwok (Bangkok: Spinger), 148–159.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *J. Educ. Stat.* 6, 107–128. doi: 10.3102/10769986006002107

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Brit. Med. J.* 327, 557–560. doi: 10.1136/bmj.327.7414.557

Holden, C. L., and Sykes, J. M. (2011). Leveraging mobile games for place-based language learning. *Int. J. Game Based Lea.* 1, 1–18. doi: 10.4018/ijgbl.2011040101

Huang, X., He, J., and Wang, H. (2020). "A case study: students' perception of a collaborative game-based virtual learning environment," in *Proceedings of the 2020 6th International Conference of the Immersive Learning Research Network*. eds. E. Daphne, K. Alexander, D. Heather, P. Anasol, J. W. L. Mark, B. Dennis, et al (IEEE) 46–53.

Hwang, W. Y., and Hu, S. S. (2013). Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving. *Comput. Educ.* 62, 308–319. doi: 10.1016/j.compedu.2012.10.005

Khatoony, S. (2019). "An innovative teaching with serious games through virtual reality assisted language learning," in *Proceedings of the 2019 International Serious Games Symposium* (Tehran: IEEE), 100–108.

Lai, K. W. K., and Chen, H. J. H. (2021). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Comput. Assist. Lang. Learn.* 34, 1–34. doi: 10.1080/09588221.2021.1928226

Lee, J. (2020). Problem-based gaming via an augmented reality mobile game and a printed game in foreign language education. *Educ. Inf. Technol.* 27, 743–771. doi: 10.1007/s10639-020-10391-1

Li, G., Luo, H., Hou, S., and Zhu, M. (2021). "Comparison of direct and vicarious VR learning experience: a perspective from accessibility and equity," in *Proceeding of the 2021 7th International Conference of the Immersive Learning Research Network*. eds. E. Daphne, P. Anasol, D. Andreas, D. Heather, M. Markos, K. Alexander, et al. (Eureka: IEEE), 1–5.

Lin, H. (2014). Establishing an empirical link between computer-mediated communication (CMC) and SLA: a meta-analysis of the research. *Lang. Learn. Technol.* 18, 120–147.

Lin, S. Y., Chen, M. P., Wang, L. C., Kao, Y. T., Zou, D., and Xie, H. (2019). "Enhancing low achievers EFL learning with interactive digital technologies" in *Proceedings of the ICCE 2019 - 27th International Conference on Computers in Education*. eds. C. Maiga, S. Hyo-Jeong, W. Lung-Hsiang, Y. Fu-Yun, S. Ju-Ling, B. Ivica, et al. (Kenting: Asia-Pacific Society for Computers in Education), 617–623.

Lipsey, M. W., and Wilson, D. B. (2001). *Practical Meta-Analysis*. London; Thousand Oaks: Sage Publications

Liu, Y., Holden, D., and Zheng, D. (2016). Analyzing students' language learning experience in an augmented reality mobile game: an exploration of an emergent learning environment. *Procedia Soc. Behav. Sci.* 228, 369–374. doi: 10.1016/j.sbspro.2016.07.055

Luo, H., Li, G., Feng, Q., Yang, Y., and Zuo, M. (2021a). Virtual reality in K-12 and higher education: a systematic review of the literature from 2000 to 2019. *J. Comput. Assist. Learn.* 37, 887–901. doi: 10.1111/jcal.12538

Luo, H., Yang, T., Kwon, S., Li, G., Zuo, M., and Choi, I. (2021b). Performing versus observing: investigating the effectiveness of group debriefing in a VR-based safety education program. *Comput. Educ.* 175:104316. doi: 10.1016/j.compedu.2021.104316

Moeller, A. J., and Catalano, T. (2015). "Foreign language teaching and learning" in *International Encyclopedia of the Social & Behavioral Sciences*. ed. W. James (New York: Pergamon Press), 327–332.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G.PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151, 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135

Nicolaidou, I., Pissas, P., and Boglou, D. (2021). Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: a quasi-experiment. *Interact. Learn. Environ.* 29, 1–15. doi: 10.1080/10494820.2020.1870504

Oswald, F. L., and Plonsky, L. (2010). Meta-analysis in second language research: choices and challenges. *Annu. Rev. Appl. Lingui.* 30, 85–110. doi: 10.1017/S0267190510000115

Parong, J., and Mayer, R. E. (2018). Learning science in immersive virtual reality. *J. Educ. Psychol.* 110, 785–797. doi: 10.1037/edu0000241

Pérez-López, D., and Contero, M. (2013). Delivering educational multimedia contents through an augmented reality application: a case study on its impact on knowledge acquisition and retention. *Turk. Online J. Educ. Technol.* 12, 19–28.

Pinho, M. S., Bowman, D. A., and Freitas, C. M. (2008). Cooperative object manipulation in collaborative virtual environments. *J. Braz. Comput. Soc.* 14, 53–67. doi: 10.1007/BF03192559

Plonsky, L., and Ziegler, N. (2016). The CALL-SLA interface: insights from a second-order synthesis. *Lang. Learn. Technol.* 20, 17–37.

Qiu, X. Y., Chiu, C. K., Zhao, L. L., Sun, C. F., and Chen, S. J. (2021). Trends in VR/AR technology-supporting language learning from 2008 to 2019: a research perspective. *Interact. Learn. Envir.* 1–24, 1–24. doi: 10.1080/10494820.2021.1874999

Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput. Educ.* 147:103778. doi: 10.1016/j.compedu.2019.103778

Richardson, J. T. (2018). The use of Latin-square designs in educational and psychological research. *Educ. Res. Rev.* 24, 84–97. doi: 10.1016/j.edurev.2018.03.003

Rosenthal, R. (1991). Meta-analysis: a review. *Psychosom. Med.* 53, 247–271. doi: 10.1097/00006842-199105000-00001

Rosenthal, R., and Rubin, D. B. (1978). Interpersonal expectancy effects: the first 345 studies. *Behav. Brain Sci.* 1, 377–386. doi: 10.1017/S0140525X00075506

Rothstein, H. R., Sutton, A. J., and Borenstein, M. (2006). *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Hoboken: John Wiley & Sons.

Smith, M. L., and Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *Am. Psychol.* 32, 752–760. doi: 10.1037/0003-066X.32.9.752

Snyder, P., and Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. *J. Exp. Educ.* 61, 334–349. doi: 10.1080/00220973.1993.10806594

Solak, E., and Cakır, R. (2016). Investigating the role of augmented reality technology in the language classroom. *Croat. J. Educ.* 18, 1067–1085. doi: 10.15516/cje.v18i4.1729

Solak, E., and Erdem, G. (2015). A content analysis of virtual reality studies in foreign language education. *Particip. Educ. Res.* 15, 21–26. doi: 10.17275/per.15.spi.2.3

Tai, T. Y., and Chen, H. H. J. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *J. Educ. Comput. Res.* 59, 1272–1293. doi: 10.1177/0735633121994291

Tai, T. Y., Chen, H. H. J., and Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Comput. Assist. Lang. Learn.* 35, 892–917. doi: 10.1080/09588221.2020.1752735

Vapenstad, C., and Buzink, S. N. (2013). Procedural virtual reality simulation in minimally invasive surgery. *Surg. Endosc.* 27, 364–377. doi: 10.1007/s00464-012-2503-1

Vázquez, C., Xia, L., Aikawa, T., and Maes, P. (2018). "Words in motion: kinesthetic language learning in virtual reality," in *Proceedings of the 2018 IEEE 18th International Conference on Advanced Learning Technologies*. eds. C. Maiga, C. Nian-Shing, H. Ronghuai, Kinshuk M. Kannan, M. Sahana, et al. (Mumbai: IEEE), 272–276.

Villena-Taranilla, R., Tirado-Olivares, S., Cózar-Gutiérrez, R., and González-Calero, J. A. (2022). Effects of virtual reality on learning outcomes in K-6 education: a meta-analysis. *Educ. Res. Rev.* 35:100434. doi: 10.1016/j.edurev.2022.100434

Wang, C. P., Lan, Y. J., Tseng, W. T., Lin, Y. T. R., and Gupta, K. C. L. (2020). On the effects of 3D virtual worlds in language learning – a meta-analysis. *Comput. Assist. Lang. Learn.* 33, 891–915. doi: 10.1080/09588221.2019.1598444

Wehner, A. K., Gump, A. W., and Downey, S. (2011). The effects of second life on the motivation of undergraduate students learning a foreign language. *Comput. Assist. Lang. Learn.* 24, 277–289. doi: 10.1080/09588221.2010.551757

Xie, Y., Chen, Y., and Ryder, L. H. (2021). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Comput. Assist. Lang. Learn.* 34, 225–245. doi: 10.1080/09588221.2019.1604551

Xu, W., Dai, T., and Shen, Z. (2022). Effect sizes and research directions of technology application in museum learning: evidence obtained by integrating meta-analysis with co-citation network analysis. *J. Comput. Assist. Learn.* 38, 565–580. doi: 10.1111/jcal.12628

Zhang, J., Li, G., Huang, Q., Feng, Q., and Luo, H. (2022). Augmented reality in K-12 education: a systematic review and meta-analysis of the literature from 2000 to 2020. *Sustainability* 14:9725. doi: 10.3390/su14159725

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 October 2022
DOI 10.3389/fpsyg.2022.1009305

Evaluating a pedagogical approach to promoting academic integrity in higher education: An online induction program

Laura Sbaffi† and Xin Zhao*†

Information School, The University of Sheffield, Sheffield, United Kingdom

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University, China

REVIEWED BY
Yovav Eshet,
Zefat Academic College, Israel
Hui Gao,
Durham University, United Kingdom
Xinyi Huang,
Durham University, United Kingdom
Shiu Pong Kenneth Ng,
City University of Macau, Macao SAR, China

*CORRESPONDENCE
Xin Zhao
xin.zhao@sheffield.ac.uk

†These authors have contributed equally to this work

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

RECEIVED 01 August 2022
ACCEPTED 15 September 2022
PUBLISHED 05 October 2022

CITATION
Sbaffi L and Zhao X (2022) Evaluating a pedagogical approach to promoting academic integrity in higher education: An online induction program. *Front. Psychol.* 13:1009305. doi: 10.3389/fpsyg.2022.1009305

Academic integrity is at the heart of excellent education. However, resources explaining the concept tend to be definition-driven, while using complex language and sometimes even an austere tone designed to discourage students from breaches. This study aims to design and evaluate an online module at a UK University across 2 years, designed to improve students' understanding of concepts of academic integrity and practice. The module includes a range of interactive resources (e.g., gamified quizzes and e-booklets) and was made available to a large cohort of postgraduate students (448). The study adopts a mixed-methods approach composed of three sequential phases involving first collecting students' views on existing academic integrity resources (7 students participating in a focus group and 39 competing a questionnaire), then developing a range of new ones based on students' feedback to form the content of the module, and finally gathering students' evaluation on the newly created resources (sample size: 361 students). Results illustrate a clear improvement in relation to the accessibility, usefulness and understandability of new resources. Results also highlight a remarkable increase in student confidence levels regarding academic integrity. Students also considered the new module as more appealing and informative. This manuscript offers a good example of a pedagogical approach aimed at promoting academic integrity in an innovative and engaging fashion.

KEYWORDS

academic integrity, policy and practice, educational games, online module, online pedagogy

Introduction

Academic integrity is at the heart of excellent teaching and learning and can be defined as a commitment and demonstration of honest and moral behaviour in an academic setting and it is applicable to both students and academic staff. The lack of such commitment can lead to academic dishonesty, which is "related to the deterioration of educational goals, specifically ideas that impact learners' intellectual, civic, and psychosocial development" (Eshet and Margaliot, 2022, p. 2). Macfarlane et al. (2014) conducted a systematic literature review showing that research into academic integrity is often centred on unethical student behaviours, which may be either accidental or intentional (Walker and White, 2014). For example, Greenberger et al. (2016) argued that one of the most common breaches of academic integrity is plagiarism, caused by poor paraphrasing practices and incorrect referencing formats. At the other end of the spectrum is the deliberate attempts at cheating, which range from buying, selling, or trading essays, to arranging for someone else to take an exam (Bretag et al., 2019; Awdry et al., 2021). However, research suggests that the definition of academic integrity is not universally understood and is open to different interpretations; this may cause misunderstanding among staff and students (Waltzer and Dahl, 2022), and ultimately lead to unethical academic behaviours (Gullifer and Tyson, 2014).

Upholding academic integrity in higher education is vital for all stakeholders. Nevertheless, universities face numerous challenges related to breaches in student academic integrity, whether unintentional or deliberate (Mahmud and Ali, 2021). A key issue appears to be centred on aligning concepts of academic integrity, policies and processes, with teaching and learning practices. This study aims to design, implement and evaluate an online academic integrity module with resources tailored to students' needs. This manuscript presents results of qualitative and quantitative evaluation data, aiming to inform current teaching practice related to academic integrity in higher education.

The challenges of student academic integrity development in higher education

Existing literature highlights many barriers to the teaching of academic integrity within the higher education sector. One significant barrier appears to be the complex terminology and unclear processes associated with breaches in academic integrity (Wong et al., 2016; Ransome and Newton, 2018; Bretag et al., 2019). Consequently, students may be left vulnerable to penalties due to inadequate understanding of academic integrity (Palmer et al., 2019). Another barrier is often caused by competing views across cultures of what constitutes academic integrity and unethical academic conduct (Zhang et al., 2014; Chien, 2017; Palmer et al., 2017; Kam et al., 2018; Khanal and Gaulee, 2019; Błachnio et al., 2021), or even across academic institutions within a single national culture (Bretag et al., 2014; Walker and White, 2014). This may result in mixed messages to students who enter a new university with misaligned prior knowledge of these concepts in relation to institutional requirements (Bertram Gallant, 2017). Other barriers that have been frequently mentioned in the literature include language (e.g., unfamiliarity with academic writing style) (Newton, 2016) and cultural barriers (Bista, 2011), which are often experienced by students whose first language is not English (Fass-Holmes, 2017; Jian et al., 2019). For example, Mahmud et al. (2019) highlighted significant differences in student perceptions of academic integrity in relation to the UK and Eastern European countries, advocating that academic integrity policies should be considerate of national cultures. Clough et al. (2015) interviewed 30 Unfair Means officers, finding that unfair means, or breaches of academic integrity in essay-based assignments, are more common among non-native English speakers. Other research highlighted a danger in the over-simplistic view that international students cheat due to culturally diverse values (Bretag et al., 2019), potentially leading to staff being biased towards them (Zhao and Kung, 2021).

Promoting academic integrity in higher education

Recent research developments heightened the requirement for a holistic approach in promoting academic integrity, including the establishment of clearly defined academic integrity principles and terminologies. Löfström et al. (2015) observed that there are mismatches in the perspectives of teachers and students regarding responsibility for upholding academic integrity standards in universities. Bealle (2017) argued that students, as those most affected by academic integrity policies, tend to become passive recipients rather than active upholders. In the same vein, Bretag et al. (2014) called for universities to move beyond mere information provision on academic integrity to engage students by integrating education and support into their academic curriculum.

Existing research recognises two dominant approaches to preventing student academic integrity breaches: punitive and educative. Richards et al. (2016, p. 243) suggested that a punitive approach aims to "deter students from committing breaches through the threat of penalties," whereas an educative approach aims to "reduce the likelihood of students committing breaches by providing them with relevant skills and knowledge." Conversely, Walker and White (2014) proposed two plagiarism prevention models, the "ethical" model, emphasising students' active role in adhering to the academic integrity code of conduct, and the "pedagogical model," focused on equipping students with appropriate academic skills. Although a punitive approach can communicate to students that plagiarism has serious consequences, research indicates that this approach alone is not sufficient to reduce cheating (Miller et al., 2011; Sun and Hu, 2020). Zhao and Kung (2021) highlighted the importance for universities to adopt an educative (pedagogical) approach to provide consistent and continuous teaching for students related to academic integrity before applying a more severe punitive or judicial approach as more serious cases, such as contract cheating, are found. Similarly, Palmer et al. (2019) warned that overly punitive measures may not be effective at reducing academic integrity breaches at universities.

Following a pedagogical approach, a number of researchers have highlighted the value of an early intervention strategy that is positive, proactive, engaging and continuous to help students' academic integrity development when entering a new department (Belter and Du Pré, 2009; Bista, 2011; Bretag et al., 2014; Walker and White, 2014; Newton, 2016). For example, Bertram Gallant (2017, p. 89) suggested that universities should shift the focus from enforcing rules and policies to ensuring students learn about academic integrity, fostering "a learning-oriented environment, improving instruction, and enhancing institutional support for teaching and learning." Pàmies et al. (2020) argued that universities need to take more responsibility for educating their students about plagiarism and explaining how to properly cite sources.

Research suggests that there has been improvement in university commitment to addressing academic integrity issues (Burbidge and Hamer, 2020); however, the implementation process within teaching remains unsatisfactory (Bretag et al., 2014; Gottardello and Karabag, 2020). Christie et al. (2013) reported a disparity between academic integrity promotion and actual teaching practice which assimilates academic integrity in the classroom. Particularly, postgraduate students are less informed of academic integrity policies (Fatemi and Saito, 2020), and often left underprepared for research-based academic assignments (Mahmud and Bretag, 2013). Therefore, it is of vital importance to help students develop a clear understanding of academic integrity (Tatum and Schwartz, 2017).

A number of educative strategies have proven effective. For example, researchers have demonstrated that educational initiatives, such as online modules that focus on academic integrity, can positively impact students' attitudes, reducing potentially unethical behaviours (Belter and Du Pré, 2009; Ballard, 2013; Bealle, 2017; Palmer et al., 2019; Sefcik et al., 2019; Du, 2020). Stephens et al. (2021) argued, however, that online courses are only partially useful for students, but they can become much more effective within a comprehensive approach to promoting academic integrity. Boehm et al. (2009) suggested that providing clear definitions with specific examples of what constitutes unethical behaviours can effectively prevent academic integrity breaches. Bretag et al. (2014) argued that education resources on academic integrity should be engaging and creative, for example by using storytelling and narrations. Furthermore, they suggested that regular email reminders should be incorporated, providing ongoing support for students. Macfarlane et al. (2014) proposed that the provision of educative resources, and design of student-centred activities, should be informed by student feedback and tailored towards student needs. However, relatively less empirical studies directly addressed the topic of academic integrity through the design and evaluation of interventional strategies, such as academic integrity courses (Elander et al., 2010; Cronan et al., 2017; Stoesz and Yudintseva, 2018; Perkins et al., 2020; Sotiriadou et al., 2020).

Research aim and objectives

This research aims to design and evaluate a newly created module enhancing students' understanding of academic integrity concepts, policies and practices. The module is designed to include a suite of interactive activities and engaging materials which are based on student feedback and suggestions.

The objectives of this research are:

- To explore students' needs for academic integrity related activities and resources.
- To design and implement an online module with interactive activities and learning resources to support students' academic integrity development based on student feedback.
- To examine the effectiveness of the online module in enhancing students' awareness and knowledge of academic integrity.

Materials and methods

Research design

This study adopted a mixed-methods inductive approach and was conducted in a post-graduate school at a research-intensive university in the UK. Both quantitative and qualitative data were gathered across the course of a sequential study composed of three main phases (**Figure 1**).

In phase 1 (January-February 2020), a focus group interview and an online survey were conducted with postgraduate students within the Social Science faculty at a UK university to explore their views and understanding of academic integrity. In the focus group, in-depth discussions were carried out with a sample of seven students from different postgraduate programmes of study and genders. Students were asked questions around three main areas: their general understanding of the principles and concepts of academic integrity, the support received to date with respect to academic integrity, and their

FIGURE 1
Phases of the study.

views on the existing academic integrity resources. The focus group interview questions had been previously piloted with two independent students to assess the clarity and relevance of the questions. Analysis of focus group discussions informed the design of an online questionnaire covering aspects related to student understanding of academic integrity, approaches to searching for academic integrity-related resources, their perception and needs in terms of content and style of the academic integrity resources available, and suggestions to make them more relevant and accessible. Furthermore, it listed eight aspects of existing academic integrity resources (reliable, comprehensive, easy to use, useful, credible, convenient and accessible, easy to understand and trustworthy) which students were asked to rate on a Likert scale from 1 (completely disagree) to 5 (completely agree). Cronbach's Alpha (0.941) was used to measure the internal consistency of the eight aspects. The questionnaire was pre-tested for content validity by two expert colleagues in the field of Information Science. It was then piloted with three independent students for readability and coherence of questions and was distributed to postgraduate students within the Faculty of Social Science using the university volunteer student list.

In phase 2 (April-September 2020), the authors developed a range of resources based on student feedback from phase 1, aimed at enhancing accessibility, with more detailed content and support regarding academic integrity education within the department. These resources, collated under the term "Academic Integrity Activities," were hosted in an online Blackboard module accessible to students. The package included an interactive video recording with embedded questions on academic integrity and other key information for students; a gamified academic integrity quiz with 20 scenario questions; a FAQ document related to common academic integrity-related queries; an electronic booklet (e-booklet) containing detailed examples of commonly occurring academic integrity problems and their solutions (see **Figure 2** for screenshots examples of some of the resources). Students were asked to access and familiarise themselves with the resources and undertake the quiz as part of induction activities at the beginning of the academic year. Additionally, a series of four online, synchronous online sessions were offered to students, spread throughout the Autumn Semester; the sessions were student-led, covering topics suggested by students *via* online polls which informed the design and content of the next session.

In phase 3 (September-November 2020), a new online questionnaire was distributed to obtain students' feedback regarding the newly created academic integrity resources. The questionnaire included, similarly to that of phase 1, the rating of eight key aspects of academic integrity and questions regarding the timeliness of support in the academic year, and in what ways the new resources could be further improved and promoted to students. Furthermore, respondents were asked about their confidence in relation to academic integrity-related concepts before and after use of the new resources. Phases 1 and 3 sought and received ethical approval from the University Research Ethics Committee.

Participants

In phase 1, both focus group and online survey participants were postgraduate students from the Faculty of Social Science at a UK university. The reason for adopting this approach was twofold: first, to gain a deep understanding of the issues specifically faced by the students enrolled in the department under study (which is part of the Faculty of Social Science) and second, to acquire a wider appreciation of academic integrity perceptions at institutional level. The integration of the two perspectives would aid the formulation of truly comprehensive resources. Research was advertised *via* email; only volunteer students participated and no incentives were offered. The students participating in the focus group were four females and three males (mean age: 23 years), all holding international student status. The questionnaire from phase 1 was completed by 39 students. It was not possible to establish a response rate as the questionnaire was sent *via* the university volunteering mailing list. In this sample 2/3 of the students were female (66.7%) and almost half (46.2%), domestic students (**Table 1**)

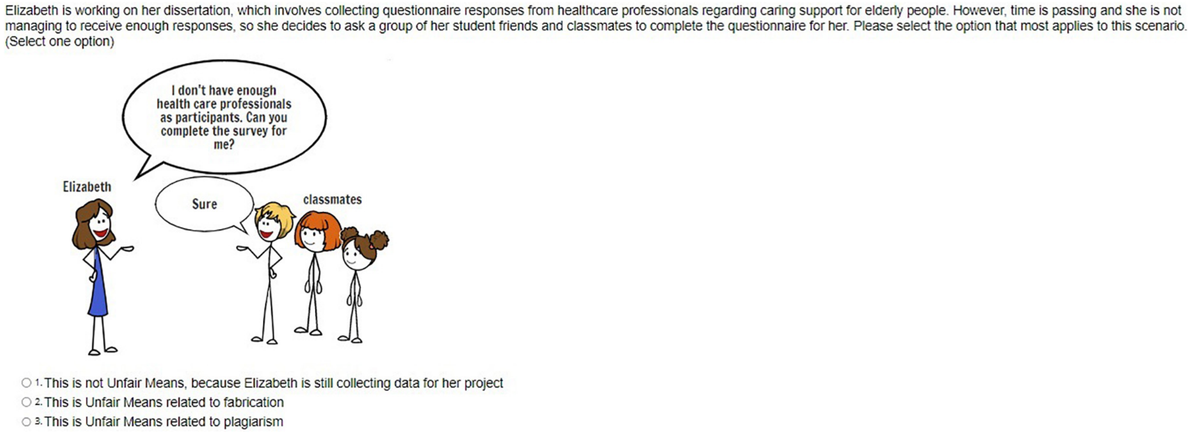


FIGURE 2
Pictured left is a representative example of a question included in the 20-question quiz; on the right is a screenshot of a page from the e-booklet.

TABLE 1 Demographic characteristics of the students participating in the research.

	Phase 1 questionnaire (*N* = 39)	Phase 1 focus group (*N* = 8)	Phase 3 questionnaire (*N* = 130)
Gender			
Male	33.3% (13)	50.0% (4)	36.2% (47)
Female	66.7% (26)	50.0% (4)	60.8% (79)
Other/Prefer not to say	0.0% (0)	0.0% (0)	3.0% (4)
Nationality			
Domestic	46.2% (18)	0.0% (0)	25.4% (33)
European	10.2% (3)	25.0% (2)	4.6% (6)
Overseas	43.6% (17)	75.0% (6)	69.2% (90)
Prefer not to say	0.0% (0)	0.0% (0)	0.8% (1)
Awareness of academic integrity concepts before joining the department			
Yes	66.7% (26)	87.5% (7)	76.2% (99)
No	17.9% (7)	0.0% (0)	15.4% (20)
Not sure	15.4% (6)	12.5% (1)	8.4% (11)

and had a mean age of 25 years. Participants in phase 3 were postgraduate students enrolled in varied programmes of study, but working within one department of the Faculty of Social Science; all accessed the induction module containing phase 2 induction activities. The students in these programmes learn in a hybrid environment offering a combination of online and face-to-face activities. Historically, however, academic integrity aspects are covered in an online format as they involve the whole student cohort and would be impractical to manage in other ways.

The questionnaire was included as the last task for completion in the induction activities of the online module. The rationale behind the inclusion of only postgraduate students from an academic department was based on the fact all resources created in phase 2 were designed specifically for this cohort (e.g., examples and quizzes were designed with the specific academic writing requirements of the department). 130 students out of 448 returned the phase 3 feedback questionnaire (response rate 29%, mean age: 26 years). Similar to phase 1, the majority of respondents were female (60.8%). However, this time a higher proportion of overseas students (69.2%) completed the survey compared to phase 1.

Research ethics

The project has received ethics approval (ID: 032172) from The University of Sheffield Ethics Committee. The approval letter has been uploaded as part of **Supplementary material**. An information sheet was provided to participants prior to data collection. Inform consent was collected prior to the survey and interviews. All data were anonymised by using a number system. Participants were reminded about the right to withdraw freely from the project.

Data analysis

Descriptive statistics, independent sample *t*-tests and comparison of means, were used to analyse quantitative data derived from the two questionnaires using IBM SPSS 26. Qualitative data from the focus group, and open-ended questions from the two questionnaires, were manually transcribed and analysed using the six-step approach to thematic analysis established by Braun and Clarke (2006). Thematic analysis was adopted to support a deeper understanding of participants' individual circumstances and experiences of academic integrity.

Results

Focus group

In phase 1, seven students participated in a focus group which highlighted three areas for improvements: academic integrity-related terminology, delivery formats, and tone of communication. Students found that the terminology around academic integrity tended to be complex and less engaging; they recommended that academic integrity information should include "real examples" with analysis, and "scenario questions" that are easy to follow and appear relatable:

> "The definitions are ok, but I would have liked more examples, practical scenarios so that it's easier to know what to avoid" (Focus group participant 1, China, Female)

They also recommended that the content of academic integrity should be delivered in a more engaging manner, suggesting interactive videos, humorous print/online brochures, and drop-in sessions for Q&A:

> "Maybe print brochures and include them in the welcome pack, or draw some comics/make short videos (better if humorous)" (Focus group participant 5, China, Female)

Additionally, students found that the tone of academic integrity related communications tended to be "scary." They expressed a preference towards a more neutral or supportive tone:

> "I hope there could be a better way to approach this issue, such as the officer should hold a neutral ground and try to guide and explain to students of their mistakes and not try to make them feel ashamed of what happened" (Focus group participant 3, Singapore, Male)

Pre and post-intervention quantitative data analysis

Two thirds of the students who completed the phase 1 questionnaire were aware of academic integrity concepts prior to their arrival at university. However, the remaining ones did not, or were unsure. In phase 3, 361 postgraduate students from one academic department completed academic integrity activities, representing a completion rate of 81%. Of those students, 130 completed the feedback questionnaire (**Table 1**). In terms of initial awareness of concepts of academic integrity, 76.2% reported a clear understanding prior to joining the department. Of those reporting no awareness of academic integrity, 74% were non-domestic students.

As described in the Section "Materials and methods," for comparative purposes, students participating in phases 1 and 3 were asked to rate the same eight aspects of academic integrity resources (see Section "Research design") on a Likert scale from 1 (completely disagree) to 5 (completely agree). Such aspects and their ratings before (phase 1) and after (phase 3) the design of the suite of new academic integrity resources are summarised in **Table 2** and **Figure 3**. Because of the time of data collection, participants in phases 1 and 3 were from two different cohorts of students, i.e., different academic years. This is because postgraduate programmes in the UK only last for 1 year. Nevertheless, a steady improvement in all aspects has been observed from the January 2020 (phase 1) to November 2020 questionnaire (phase 3), demonstrating an overall positive reception of the new academic integrity material developed in phase 2.

All eight aspects scored higher ratings in phase 3. Independent sample *t*-tests with Bonferroni correction performed on the eight aspects showed a statistically significant improvement of the perceived usefulness of the resources (from $M = 3.79$ in phase 1, to $M = 4.27$ in phase 3, $p = 0.028$; $t = 2.210$) and their understandability (from $M = 3.62$ in phase 1 to $M = 4.09$ in phase 3, $p = 0.031$; $t = 2.168$). Aspects of trustworthiness and credibility increased from a $M = 4.10$ to $M = 4.20$ for trustworthiness and $M = 4.10$ to $M = 4.14$ for credibility in phase 1 and phase 3 respectively. Although their increases in phase 3 were not statistically significant, they had been the two highest ranked aspects in phase 1, so the margin for improvement was relatively smaller.

The exploration of changes in perceived usefulness and understandability of the resources from phase 1 to phase 3 in terms of proportion of students rating them 4 (agree) and 5 (completely agree), also showed striking results; in phase 3, in fact, 17.1% more students agreed that the new resources were useful, and 28.5% more students agreed that they were easy to understand.

TABLE 2 Mean values of the eight aspects of academic integrity resources before (phase 1) and after (phase 3) implementation of the new academic integrity module.

Aspects of academic integrity resources	Phase 1 (N = 39)		Phase 3 (N = 130)		Comparison			
	M	SD*	M	SD*	M difference	p-value	SE*	t statistic
Reliable	4.00	1.08	4.11	1.28	0.11	0.627	0.226	0.487
Comprehensive	3.92	0.87	4.16	1.28	0.24	0.275	0.219	1.096
Credible	4.10	1.07	4.14	1.30	0.04	0.861	0.228	0.175
Convenient and accessible	3.82	1.12	4.05	1.29	0.23	0.316	0.229	1.005
Easy to use	3.72	0.92	4.08	1.20	0.36	0.086	0.209	1.726
Useful	**3.79**	**1.06**	**4.27**	**1.22**	**0.48**	**0.028**	**0.216**	**2.210**
Easy to understand	**3.62**	**0.96**	**4.09**	**1.24**	**0.47**	**0.031**	**0.216**	**2.168**
Trustworthy	4.10	1.21	4.20	1.30	0.10	0.669	0.234	0.428

In bold are the statistically significant differences. *M, mean; SD, standard deviation; SE, standard error.

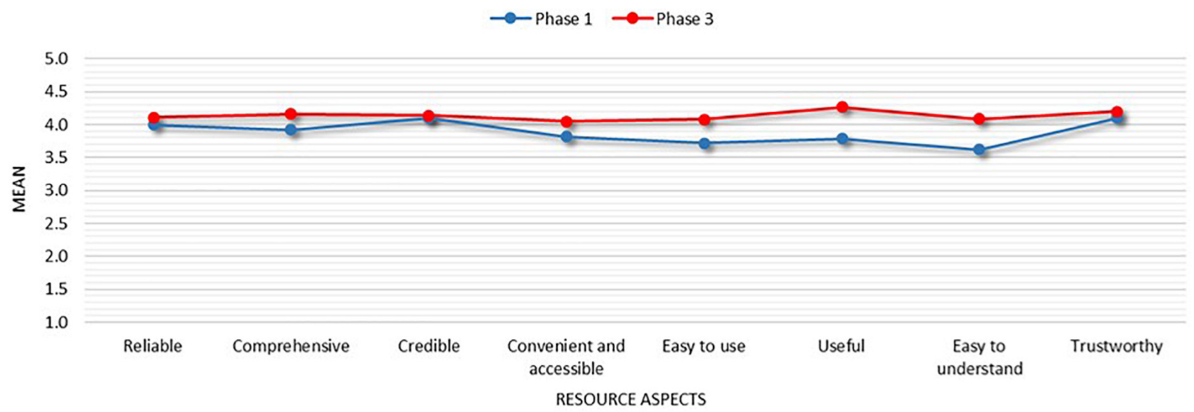


FIGURE 3
Mean values of key aspects of academic integrity resources before (phase 1) and after (phase 3) implementation of induction tasks.

The vast majority of students (85%) also found the new module adequate in supporting their learning; the remaining 12% of students answered "maybe" and a very small percentage of students (3%) considered the module inadequate. It is notable that this feedback was received from postgraduate students who attend 1-year programmes, and therefore do not have access to previous years' resources or modules. Furthermore, the study was conducted prior to students' first assessed submissions, hence, it was not possible to gauge their actual understanding of academic integrity principles.

In the phase 3 questionnaire, students were asked to rate their confidence regarding academic integrity concepts prior to and following the use of the new resources on a Likert scale from 1 (very worried) to 5 (very confident). Results show that only 26% of respondents were confident (rated 4 and 5) before using resources. However, this percentage increased to 60.6% after use. Nevertheless, despite the considerable increase in positive perceptions recorded in phase 3, around 28.3% of respondents still believed resources have the potential to be improved further.

In the phase 3 questionnaire, students were asked to indicate which of the new academic integrity resources they found most useful; they could select more than one option when answering this question (**Figure 4**).

All resources were well-received; however, the e-booklet scored particularly highly (41.1%), perhaps due to the attractiveness of its innovative format, one not normally used to deliver academic integrity content; this was followed by the scenario-based 20-question quiz (36.7%) and the FAQ document (35.2%). A small percentage of students (6.3%) reported none of the resources as useful. However, this could be due to the poor familiarity, at such an early stage of their studies, with the platform (Blackboard) in which they are embedded. Live online sessions were not included in the list, as the questionnaire was distributed before the sessions took place. Nevertheless, their usefulness was demonstrated by the consistently high number of students (about half of the full cohort in each one) attending them.

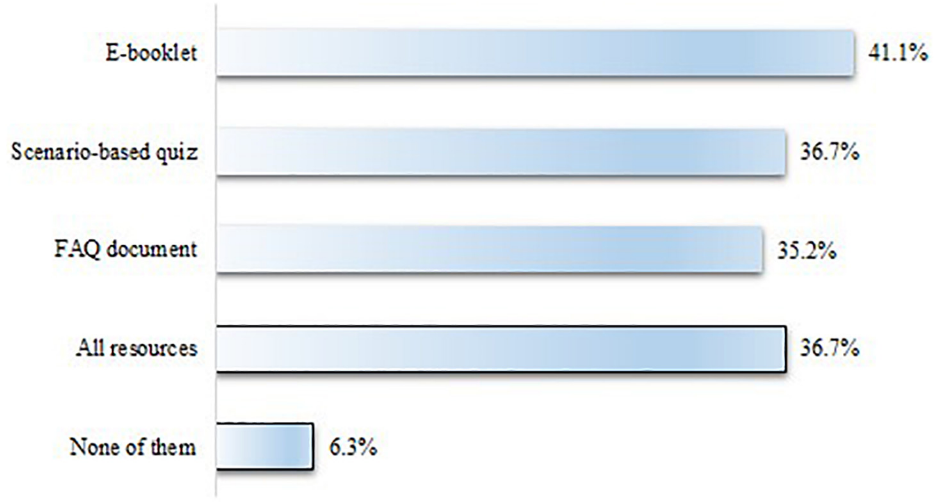


FIGURE 4
New academic integrity resources as ranked by perceived usefulness (according to student feedback).

Analysis of open-text data from phase 3 questionnaire

Open-text questions were included in the phase 3 questionnaire, allowing students to elaborate on their view of the new resources and provide suggestions for future improvement. Qualitative data suggests that students highly valued the resources provided to them, which helped them develop an enhanced understanding of academic integrity concepts and varied types of breaches in academic integrity:

> "I didn't expect to get such a detailed (and not boring) introduction on this subject." (Participant 28, Switzerland, Male)

> "I have attended lectures of [anonymised], their experience with students who used unfair means (knowingly or unknowingly) helped me a lot. Before that, I know just the theory behind unfair means but after examples, I realised that I do not know much about unfair means in practice." (Participant 114, India, Male)

Structured and accessible academic integrity materials

Two main themes were identified in qualitative data, corresponding to the Academic Integrity Model (Bretag et al., 2014): aspects of access and support of academic integrity resources. Students reported that having all resources in one virtual space was very helpful. For example, having a single online module covering academic integrity throughout their study allowed them to learn about these concepts at the beginning of their studies, as well as throughout the academic year and at critical times, such as before coursework submission periods:

> "Everything I need to establish what falls or doesn't under Academic Integrity can be found in the folders; which has been helpful." (Participant 31, Netherlands, Female)

> "I appreciated the introductory material, and know where to go for more information/guidance when I need it." (Participant 37, UK, Male)

However, students also expressed a preference towards a structure of the resources that avoids information repetition or cognitive overload. Most suggestions for improvement related to making resources organised in a clear fashion on Blackboard. Many students who had not used Blackboard prior to joining the university found it challenging to navigate and struggled to follow material, particularly at the beginning of the academic year. Student recommendations included walkthrough videos, explaining the overall structure of resources, brief videos with clearly labelled topics, and a navigation side panel in the online space:

> "The Blackboard software can be overwhelming as a new user. Clearer signposting to resources in general may improve access to learning materials." (Participant 8, UK, Male)

> "They are fine, though I think it needs to be streamlined a bit. There is a lot of content, and some of it is a bit repetitive. There should just be an explanatory video and then the quiz." (Participant 23, UK, Male)

> "I guess creating the whole information in form of some movie, while sticking only to the topic, would be more beneficial. And keeping the length of the videos small or making different videos for each topic would make it easy to search in hour of need. E.g., a student might suddenly get confused about something and search for an answer, but for that he'll have to go through the whole video." (Participant 29, India, Female)

These recommendations will be taken into consideration in future modifications of the resources.

Engaging contents and mixed resources formats

When considering the format of academic integrity resources and support, students welcomed the variety in means of delivery (videos, e-booklets, gamified quizzes with cartoons, FAQ, and online drop-ins) finding them particularly engaging. Students made very positive comments regarding the online quizzes, which contain cartoon scenarios stories based on real cases. According to students, real examples support comprehension of academic integrity concepts better, connecting them with practices and the creative use of cartoons enhanced student engagement:

> "Gave overview of what unfair means (academic integrity) are the quiz to test our understanding–made it more fun, and meant it wasn't just documents that we had to read." (Participant 27, UK, Female)

> "I particularly liked the quiz as it made me reread things in depth to get the right answer and it was real life examples rather than just pages of theory." (Participant 44, UK, Female)

> "I liked the idea of presenting doodles in each question. I never felt bored because of them and I think I can use this idea of creating doodles." (Participant 25, India, Male)

Physical copies and printable versions were also frequently mentioned by students in comments. Although students preferred the mixed format provided by online resources, they suggested that physical copies, or downloadable materials, should be made available for students as an additional option:

> "Email some of the documentation out because for the first couple of weeks I was still trying to get the hang of Blackboard and its multiple folders/files." (Participant 31, Netherlands, Female)

> "The booklet of plagiarism is only available online, it's better if we can download it." (Participant 28, China, Female)

This goal could be easily achieved by sending regular email reminders to students with links to online resources and PDF attachments.

Discussion

This manuscript adopted a pedagogical approach to promoting academic integrity to students in higher education (Walker and White, 2014; Richards et al., 2016). The study aimed to showcase the design process and evaluate the effectiveness of a newly created academic integrity module in relation to enhancing student understanding of academic integrity concepts, policies, and practices. Tailored to students' needs (Macfarlane et al., 2014), a sequential three-phase study was undertaken, collecting initial student feedback (phase 1), implementing an online academic integrity module (phase 2), and conducting a post-launch evaluation survey (phase 3). The project showed strong potential in supporting the development of students' academic integrity.

Consistent with existing literature, this research found that, compared with home students, international students are less likely to be informed about academic integrity policies before their arrival in academic departments (Bista, 2011; Fatemi and Saito, 2020). However, both home and international students participating in the study expressed a desire for academic departments to improve academic integrity resources, rather than using punishment and threats to prevent breaches. The results of this study provide supporting evidence for the arguments of previous studies, which argue that an educative approach to academic integrity is more effective than punitive measures (Miller et al., 2011; Palmer et al., 2019; Sun and Hu, 2020).

This study also highlights the importance of an early educative intervention to support the academic integrity of students, particularly with programmes with diverse student cohorts. The findings of this research revealed that a main barrier to student understanding of academic integrity is definition-driven terminologies that lack concrete examples. This finding is consistent with the literature, suggesting that universities should avoid complex language related to academic integrity (Bretag et al., 2019) and go beyond providing statements and definitions of academic misconduct (Risquez et al., 2013).

Another barrier reported by participants is a lack of creative resources and an absence of supportive tone from academic departments; this discourages student engagement and understanding of academic integrity; this broadly corresponds to findings of Bertram Gallant (2008), who stressed the importance of creating a supportive learning-oriented environment which fosters development in students' academic integrity.

According to the findings of this research, an online academic integrity module has proven to be an effective intervention strategy for increasing student awareness and understanding of academic integrity related concepts, procedures, and policies. This is consistent with the literature suggesting that academic integrity education programmes can positively influence student attitudes and reduce breaches of academic integrity (Greer et al., 2012; Obeid and Hill, 2017; Levine and Pazdernik, 2018; Sefcik et al., 2019). As previously reported, there are relatively few empirical studies directly addressing academic integrity through both the design and evaluation of intervention strategies. This manuscript fills this important gap in the literature, showing both the design and evaluation of the online module, such as how the design of the module was informed by student feedback and how the module promoted academic integrity among students (Stoesz and Yudintseva, 2018; Perkins et al., 2020). Student participants in the research reported a strong preference towards academic integrity resources and support that are structured and easy to access at different stages of their study. Comparative data analysis revealed the online academic integrity module significantly enhanced various aspects of academic integrity resources, particularly aspects of usefulness and comprehension of student perspectives. Results also revealed a remarkable increase in students' confidence regarding their knowledge of concepts of academic integrity. This evidences the importance of establishing early intervention, and a continuously accessible online course that promotes academic integrity in higher education.

Findings also highlight the importance of the use of a variety of media when delivering academic integrity resources to enhance student engagement and understanding. Students showed a preference towards online academic integrity booklets with examples of detailed analysis of good and poor practice; this allowed them to learn about different types of breaches. This concords with the findings of Boehm et al. (2009), who showed clear examples of what constitutes unethical academic behaviour are able to help prevent academic integrity breaches. Furthermore, students appreciated cartooned scenarios, entailing storytelling, and praised the entertaining and relatable nature of these materials. Results here reflect those of Macfarlane et al. (2014), who highlighted the need to use engaging techniques, such as storytelling and narration, to teach academic integrity related topics.

Limitations and future directions

The results of this study have important implications for higher education with respect to the design of academic integrity resources to support students' academic transition. As participants to this research were postgraduate students with essay-based assignments of social science subjects, generalisability of the results is limited. A second limitation regards the possibility of response bias in research and the associated implications. Participating students knew that the authors sought their feedback to evaluate a newly developed module, and they may have wished to please them and give them what they thought they expected (i.e., positive perceptions of the new materials). The authors attempted to minimise this issue by asking a research assistant to moderate the focus group. However, it would have been impossible to minimise this possibility. A further limitation regards the small sample size of the phase 1 study, which may also contribute to reducing the generalisability of the results. Future research may seek to focus on different subject disciplines, and compare results with the current study. Additionally, further studies may also explore the impact of the new module on students' long-term engagement with academic integrity practices.

Data availability statement

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

Ethics statement

The project has received ethics approval (ID: 032172) from The University of Sheffield Ethics Committee. An information sheet was provided to participants prior to data collection. Inform consent was collected prior to the survey and interviews.

Author contributions

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

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1009305/full#supplementary-material

References

Awdry, R., Dawson, P., and Sutherland-Smith, W. (2021). Contract cheating: To legislate or not to legislate-is that the question? *Assess. Eval. High. Educ.* 47, 712–726. doi: 10.1080/02602938.2021.1957773

Ballard, I. B. (2013). The impact of an academic integrity module and turnitin® on similarity index scores of undergraduate student papers. *Res. Sch.* 20, 1–13. doi: 10.1080/00405841.2017.1283573

Bealle, P. (2017). Community college academic integrity lessons that put research into practice. *Theory Pract.* 56, 144–155.

Belter, R. W., and Du Pré, A. (2009). A strategy to reduce plagiarism in an undergraduate course. *Teach. Psychol.* 36, 257–261. doi: 10.1080/00986280903173165

Błachnio, A., Cudo, A., Kot, P., Torój, M., Oppong Asante, K., Enea, V., et al. (2021). Cultural and psychological variables predicting academic dishonesty: A cross-sectional study in nine countries. *Ethics Behav.* 32, 44–49. doi: 10.1080/10508422.2021.1910826

Bertram Gallant, T. (2008). Academic integrity in the twenty-first century: A teaching and learning imperative. *ASHE High. Educ. Rep.* 33, 1–143.

Bertram Gallant, T. (2017). Academic integrity as a teaching & learning issue: From theory to practice. *Theory Pract.* 56, 88–94. doi: 10.1080/00405841.2017.1308173

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., et al. (2019). Contract cheating: A survey of Australian university students. *Stud. High. Educ.* 44, 1837–1856. doi: 10.1080/03075079.2018.1462788

Bretag, T., Mahmud, S., Wallace, M., Walker, R., McGowan, U., East, J., et al. (2014). Teach us how to do it properly!'An Australian academic integrity student survey. *Stud. High. Educ.*39, 1150–1169. doi: 10.1080/03075079.2013.777406

Bista, K. K. (2011). 12: Academic dishonesty among international students in higher education. *Improve Acad.* 30, 159–172. doi: 10.1002/j.2334-4822.2011.tb00655.x

Boehm, P. J., Justice, M., and Weeks, S. (2009). Promoting academic integrity in higher education. *Commun. Coll. Enterp.* 15, 45–61.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706QP063OA

Burbidge, T., and Hamer, R. (2020). Academic honesty in the international baccalaureate diploma programme: Student, teacher, and school perspectives. *J. Int. Stud.* 10, 265–285. doi: 10.32674/jis.v10i2.667

Chien, S. C. (2017). Taiwanese college students' perceptions of plagiarism: Cultural and educational considerations. *Ethics Behav.* 27, 118–139. doi: 10.1080/10508422.2015.1136219

Christie, B. A., Miller, K. K., Cooke, R., and White, J. G. (2013). Environmental sustainability in higher education: How do academics teach? *Environ. Educ. Res.* 19, 385–414. doi: 10.1080/13504622.2012.698598

Clough, P., Willett, P., and Lim, J. (2015). "Unfair means: Use cases beyond plagiarism," in *Proceedings of the international conference of the cross-language evaluation forum for european languages* (Cham: Springer), 229–234. doi: 10.1007/978-3-319-24027-5_21

Cronan, T. P., McHaney, R., Douglas, D. E., and Mullins, J. K. (2017). Changing the academic integrity climate on campus using a technology-based intervention. *Ethics Behav.* 27, 89–105. doi: 10.1080/10508422.2016.1161514

Du, Y. (2020). Evaluation of intervention on Chinese graduate students' understanding of textual plagiarism and skills at source referencing. *Assess. Eval. High. Educ.* 45, 14–29. doi: 10.1080/02602938.2019.1601680

Elander, J., Pittam, G., Lusher, J., Fox, P., and Payne, N. (2010). Evaluation of an intervention to help students avoid unintentional plagiarism by improving their authorial identity. *Assess. Eval. High. Educ.* 35, 157–171. doi: 10.1080/02602930802687745

Eshet, Y., and Margaliot, A. (2022). Does creative thinking contribute to the academic integrity of education students? *Front. Psychol.* 13:925195. doi: 10.3389/fpsyg.2022.925195

Fass-Holmes, B. (2017). International students reported for academic integrity violations: Demographics, retention, and graduation. *J. Int. Stud.* 7, 664–669. doi: 10.5281/zenodo.570026

Fatemi, G., and Saito, E. (2020). Unintentional plagiarism and academic integrity: The challenges and needs of postgraduate international students in Australia. *J. Further High. Educ.* 44, 1305–1319. doi: 10.1080/0309877X.2019.1683521

Greenberger, S., Holbeck, R., Steele, J., and Dyer, T. (2016). Plagiarism due to misunderstanding: Online instructor perceptions. *J. Scholarsh. Teach. Learn.* 16, 72–84. doi: 10.14434/josotl.v16i6.20062

Greer, K., Swanberg, S., Hristova, M., Switzer, A. T., Daniel, D., and Perdue, S. W. (2012). Beyond the web tutorial: Development and implementation of an online, self-directed academic integrity course at Oakland University. *J Acad. Libr.* 38, 251–258. doi: 10.1016/j.acalib.2012.06.010

Gottardello, D., and Karabag, S. F. (2020). Ideal and actual roles of university professors in academic integrity management: A comparative study. *Stud. High. Educ.* 47, 1–19. doi: 10.1080/03075079.2020.1767051

Gullifer, J. M., and Tyson, G. A. (2014). Who has read the policy on plagiarism? Unpacking students' understanding of plagiarism. *Stud. High. Educ.* 39, 1202–1218. doi: 10.1080/03075079.2013.777412

Jian, H., Marion, R., and Wang, W. (2019). Academic integrity from China to the United States: The acculturation process for chinese graduate students in the United States. *Ethics Behav.* 29, 51–70. doi: 10.1080/10508422.2018.1468760

Kam, C. C. S., Hue, M. T., and Cheung, H. Y. (2018). Plagiarism of Chinese secondary school students in Hong Kong. *Ethics Behav.* 28, 316–335. doi: 10.1080/10508422.2017.1333909

Khanal, J., and Gaulee, U. (2019). Challenges of international students from pre-departure to post-study: A literature review. *J. Int. Stud.* 9, 560–581. doi: 10.32674/jis.v9i2.673

Levine, J., and Pazdernik, V. (2018). Evaluation of a four-prong anti-plagiarism program and the incidence of plagiarism: A five-year retrospective study. *Assess. Eval. High. Educ.* 43, 1094–1105. doi: 10.1080/02602938.2018.1434127

Löfström, E., Trotman, T., Furnari, M., and Shephard, K. (2015). Who teaches academic integrity and how do they teach it? *High. Educ.* 69, 435–448. doi: 10.1007/s10734-014-9784-3

Macfarlane, B., Zhang, J., and Pun, A. (2014). Academic integrity: A review of the literature. *Stud. High. Educ*. 39, 339–358. doi: 10.1080/03075079.2012.709495

Mahmud, S., and Ali, I. (2021). Evolution of research on honesty and dishonesty in academic work: A bibliometric analysis of two decades. *Ethics Behav*. 1–15.doi: 10.1080/10508422.2021.2015598

Mahmud, S., and Bretag, T. (2013). Postgraduate research students and academic integrity:'It's about good research training'. *J. High. Educ. Policy Manage*. 35, 432–443. doi: 10.1080/1360080X.2013.812178

Mahmud, S., Bretag, T., and Foltýnek, T. (2019). Students' perceptions of plagiarism policy in higher education: A comparison of the United Kingdom, Czechia, Poland and Romania. *J. Acad. Ethics* 17, 271–289. doi: 10.1007/s10805-018-9319-0

Miller, A., Shoptaugh, C., and Wooldridge, J. (2011). Reasons not to cheat, academic-integrity responsibility, and frequency of cheating. *J. Exp. Educ.* 79, 169–184. doi: 10.1080/00220970903567830

Newton, P. (2016). Academic integrity: A quantitative study of confidence and understanding in students at the start of their higher education. *Assess. Eval. High. Educ.* 41, 482–497. doi: 10.1080/02602938.2015.1024199

Obeid, R., and Hill, D. B. (2017). An intervention designed to reduce plagiarism in a research methods classroom. *Teach. Psychol.* 44, 155–159. doi: 10.1177/0098628317692620

Pàmies, M. D. M., Valverde, M., and Cross, C. (2020). Organising research on university student plagiarism: A process approach. *Assess. Eval. High. Educ.* 45, 401–418. doi: 10.1080/02602938.2019.1658714

Palmer, A., Oakley, G., and Pegrum, M. (2017). A culture of sharing: Transnational higher education students' views on plagiarism in the digital era. *Int. J. Manag. Educ.* 11, 381–404. doi: 10.1504/IJMIE.2017.086893

Palmer, A., Pegrum, M., and Oakley, G. (2019). A wake-up call? Issues with plagiarism in transnational higher education. *Ethics Behav.* 29, 23–50. doi: 10.1080/10508422.2018.1466301

Perkins, M., Gezgin, U. B., and Roe, J. (2020). Reducing plagiarism through academic misconduct education. *Int. J. Educ. Integr.* 16, 1–15. doi: 10.1007/s40979-020-00052-8

Ransome, J., and Newton, P. M. (2018). Are we educating educators about academic integrity? A study of UK higher education textbooks. *Assess. Eval. High. Educ.* 43, 126–137. doi: 10.1080/02602938.2017.1300636

Richards, D., Saddiqui, S., White, F., McGuigan, N., and Homewood, J. (2016). A theory of change for student-led academic integrity. *Q. High. Educ.* 22, 242–259. doi: 10.1080/13538322.2016.1265849

Risquez, A., O'Dwyer, M., and Ledwith, A. (2013). 'Thou shalt not plagiarise': From self-reported views to recognition and avoidance of plagiarism. *Assess. Eval. High. Educ.* 38, 34–43. doi: 10.1080/02602938.2011.596926

Sefcik, L., Striepe, M., and Yorke, J. (2019). Mapping the landscape of academic integrity education programs: What approaches are effective? *Assess. Eval. High. Educ*. 5, 30–43. doi: 10.1080/02602938.2019.1604942

Sotiriadou, P., Logan, D., Daly, A., and Guest, R. (2020). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Stud. High. Educ.* 45, 2132–2148. doi: 10.1080/03075079.2019.1582015

Stoesz, B. M., and Yudintseva, A. (2018). Effectiveness of tutorials for promoting educational integrity: A synthesis paper. *Int. J. Educ. Integr.* 14, 1–22. doi: 10.1007/s40979-018-0030-0

Stephens, J. M., Watson, P. W. S. J., Alansari, M., Lee, G., and Turnbull, S. M. (2021). Can online academic integrity instruction affect university students' perceptions of and engagement in academic dishonesty? Results from a natural experiment in New Zealand. . *Front. Psychol.* 12:569133. doi: 10.3389/fpsyg.2021.569133

Sun, X., and Hu, G. (2020). What do academics know and do about plagiarism? An interview study with Chinese university teachers of English. *Ethics Behav.* 30, 459–479. doi: 10.1080/10508422.2019.1633922

Tatum, H., and Schwartz, B. M. (2017). Honor codes: Evidence based strategies for improving academic integrity. *Theory Pract.* 56, 129–135. doi: 10.1080/00405841.2017.1308175

Walker, C., and White, M. (2014). Police, design, plan and manage: Developing a framework for integrating staff roles and institutional policies into a plagiarism prevention strategy. *J. High. Educ. Policy Manag.* 36, 674–687. doi: 10.1080/1360080X.2014.957895

Waltzer, T., and Dahl, A. (2022). Why do students cheat? Perceptions, evaluations, and motivations. *Ethics Behav*. 1–21. doi: 10.1080/10508422.2022.2026775

Wong, S. S. H., Lim, S. W. H., and Quinlan, K. M. (2016). Integrity in and beyond contemporary higher education: What does it mean to university students? *Front. Psychol.* 7:1094. doi: 10.3389/fpsyg.2016.01094

Zhang, D., Joy, M., Cosma, G., Boyatt, R., Sinclair, J., and Yau, J. (2014). Source-code plagiarism in universities: A comparative study of student perspectives in China and the UK. *Assess. Eval. High. Educ.* 39, 743–758. doi: 10.1080/02602938.2013.870122

Zhao and Kung (2021). Academic integrity of international students: Investigation in the era of 'The Presumption of Innocence'. *Educ. Dev.* Available online at: https://www.seda.ac.uk/seda-publishing/educational-developments/past-issues-2000-onwards/educational-developments-issue-22-2-2021/

TYPE Original Research
PUBLISHED 07 October 2022
DOI 10.3389/fpsyg.2022.998196

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University, China

REVIEWED BY
Chenchen Liu,
Wenzhou University, China
Rui S. U.,
Guangxi Normal University, China

*CORRESPONDENCE
Bing Wang
sxwangbing@mail.ccnu.edu.cn
Harrison Hao Yang
Harrison.yang@oswego.edu

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

RECEIVED 19 July 2022
ACCEPTED 15 September 2022
PUBLISHED 07 October 2022

CITATION
Zhang J, Wang B, Yang HH, Chen Z, Gao W and Liu Z (2022) Assessing quality of online learning platforms for in-service teachers' professional development: The development and application of an instrument. *Front. Psychol.* 13:998196. doi: 10.3389/fpsyg.2022.998196

Assessing quality of online learning platforms for in-service teachers' professional development: The development and application of an instrument

Jing Zhang[1], Bing Wang[2]*, Harrison Hao Yang[3]*, Zengzhao Chen[2], Wei Gao[4] and Zhi Liu[2]

[1]School of Education, Jianghan University, Wuhan, China, [2]National Engineering Laboratory for Educational Big Data, Central China Normal University, Wuhan, China, [3]School of Education, State University of New York at Oswego, Oswego, NY, United States, [4]School of Education, Central China Normal University, Wuhan, China

To help optimize online learning platforms for in-service teachers' professional development, this study aims to develop an instrument to assess the quality of this type of platforms on teacher satisfaction. After reliability and validity tests and expert empowerment, the 27-item instrument was formed. Based on the information systems (IS) success model, this instrument was designed to measure teacher perceptions of the quality of online learning platforms from three dimensions, namely, content quality, technical quality, and service quality. Moreover, the developed instrument was used to analyze the effects of the National Teacher Training Platform amid the COVID-19 outbreak in China. The findings revealed that the improvement of the platform's style, tool function, operating efficiency, and teaching methods could enhance teachers' experience of online training.

KEYWORDS
teacher education, online learning, quality evaluation, user experience, training platform

Introduction

Over the years, education reform and teacher training initiatives have worked hard to build and promote scalable, sustainable online communities for education professionals (Schlager and Fusco, 2003). One of these types of communities, online learning platforms for teacher professional development (TPD), has garnered widespread attention and has been growing rapidly for its capacity and flexibility to support teachers to continuously reflect, learn, and act to augment their practice throughout their teaching careers. Online learning platforms make teacher training and self-development feasible and more convenient in terms of time and space. A quality online teacher learning platform is an essential guarantee for positive teacher learning effects (Pengxi and Guili, 2018). Online technology can help supply high-quality TPD when it is suitably integrated into the learning platform. However, using online technology as a "quick fix" or integrating it into a platform without a clear purpose will not result in the desired changes in teaching and learning outcomes (Cobo Romani et al., 2022). Schlager and Fusco (2003) contended that focusing solely on online technology as a means of delivering training and/or creating online networks puts the cart before the horse by ignoring the Internet's even greater potential to support and strengthen local communities of practice in which teachers work. The design and delivery of high-quality TPD warrant an understanding of the applicable technology, resources required, and teachers' needs. Previous research suggested that to exert the greatest impact, professional development must be designed, implemented, and evaluated to satisfy the needs of particular teachers in particular settings (Guskey, 1994). Thus, the teachers' perceptions of online learning platforms for TPD have become an issue of great importance, especially with regard to the quality of these platforms. Obtaining and analyzing teachers' perceptions can help trainers, administrators, platform designers, and technologists better use and improve online learning platforms for TPD, thereby helping teachers acquire knowledge and skills more effectively. As Greenberg (2009, p. 2) pointed, "without a programmatic understanding of best practices and methods of anticipating potential roadblocks, far too many initiatives may falter or fail."

So far, despite the availability of some instruments to measure users' perception of online platforms or websites in general, there are few instruments specifically for teachers' perception of the quality of online training platforms. Without appropriate measurement, some online learning platforms might not be used effectively to support TPD. To address this significant issue and research gap, this study aims to: (i) develop a teacher perception scale of the quality of online learning platforms for TPD; and (ii) apply the developed scale to evaluate the National Online Teacher Training Platform in Central China.

Study design

Overall, this study contained two parts, instrument development and application study. As shown in **Figure 1**, the first part included four steps of the instrument development: conceptual framework and related works, initial scale, preliminary test, and expert evaluation empowerment. Then, the developed instrument was used to validate and evaluate the optimization effect of the online learning platform for teachers. The second part of this study included the following steps: pre-analysis of data, platform quality comparison, and benefit analysis of platform optimization.

Instrument development

Conceptual framework and related works

One of the leading models for measuring information systems (IS) is the IS success model, which aims to provide a thorough understanding of IS success by elucidating the relationships among the critical success factors frequently considered when assessing IS. The IS success model was initially developed and later revised by DeLone and McLean, 1992, 2002, 2003 in response to input from other researchers. As shown in **Figure 2**, in the IS success model, the three dimensions of quality (information or content quality, system quality, and service quality) directly affect usage intentions and user satisfaction, and, consequently, net system benefits. *Information quality*, also known as content quality, denotes the quality of the information or content that a system can store, deliver, or produce in terms of completeness, relevance, and consistency. *System quality*, also known as technical quality, denotes the quality of the system in terms of functionality, usability, efficiency, and portability. *Service quality* usually denotes the quality of support provided to the users, including reliability, responsiveness, assurance, and empathy. So far, the IS success model has been adopted widely by existing studies to measure different types of information systems, including online learning systems and websites (Lin, 2007). **Table 1** presents some main relevant instruments based on one or all of the quality dimensions. Based on the IS success model and previous related research, this study develops an instrument specifically on online platforms for teachers' professional development. Specifically, this study attempts to identify key indicators from technical quality, content quality, and service quality to measure the effectiveness of online learning platforms for teachers' professional development on teachers' experience.

Content quality

Content is the core of online learning platforms, including information and resources for learning and practice. Yang and Chan (2008) focused on the content quality in the evaluation of English learning websites and prepared a differentiated evaluation of the general content, professional learning content, and exclusive training content of each website, highlighting the authority and practicality of content. In addition, Devi and Verma (2018) investigated libraries in general, as well as service information and different types of professional books, stating that attention should also be paid to the design of learning methods and strategies besides the quality of the content itself. Desjardins and Bullock (2019) claimed that only when teachers experience the problem-based learning mode in online training, their learning stays at the theoretical level with no conflicts and contradictions. Fuentes and Martínez (2018) analyzed different

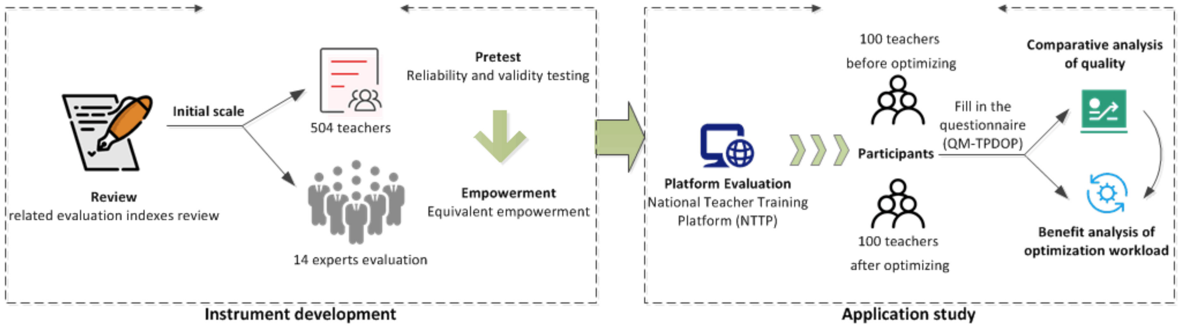


FIGURE 1
Study design.

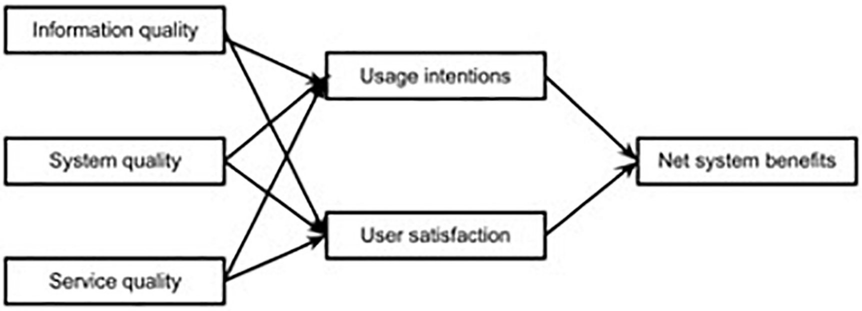


FIGURE 2
Information systems (IS) success model.

evaluation frameworks for English learning platforms and proposed an assessment list, including multimedia, interactive and educational content, and communication items, covering the assessment of teaching content, teaching methods, and strategies.

With the progress of teachers' online activity, many online learning platforms are promoting both teacher tuition and practice, providing teachers with curriculum research, materials, and resources for their teaching practice. Thus, the content quality of this study denotes the quality of online learning courses and teaching-related professional resources. Regarding specific sub-dimension design, we integrated previous research on evaluating the teaching content and method strategies with the teaching effect suggested by experts. Then, content quality indicators were constructed from three aspects: resource, method, and effectiveness. Of note, resource evaluation highlights the authority and practicality of the content. Method evaluation focuses on diversity and individual motivation. Teaching effect evaluation emphasizes the impact of learning on Teachers' teaching practices.

Technical quality

Regarding evaluating the platform's technical quality, most studies evaluated the two general aspects of the functional effectiveness and technical aesthetics of online platforms or websites and then proposed enhancement in operation technology and interface design. Reportedly, improving online platforms helps to increase user satisfaction and user application continuity (Ng, 2014; Laperuta et al., 2017; Liu et al., 2017; Manzoor et al., 2018; Mousavilou and Oskouei, 2018). Santos et al. (2016, 2019) followed the five principles, namely, multimedia quality, content, navigation, access speed, and interaction in graphic design, to measure user satisfaction in online learning websites. They revealed that usability and navigation were the two evaluation subindicators most preferred by users.

Besides the basic usability aspect, the technical quality of the online platform is the operation's efficiency and the users' stickiness in the application process. For example, ÖZkan et al. (2020) aimed at the rapid upsurge of resource retrieval, assessing the retrieval quality of academic online learning platforms from the standpoint of performance, design content, meta-tags, backlinks, and other indicators. Pant (2015) examined the quality of the central science library's websites primarily from the aspects of efficiency, satisfaction, and accessibility. The evaluation items included the smooth and fast operation of web pages, the ease of use of functions, and the reliance

TABLE 1 Main instruments of online learning platforms and websites.

Author	Object	Sample/target	Measurement	Indicators	Category		
					Content quality	Technical quality	Service quality
Hassanzadeh et al., 2012	Measurement of e-learning systems success model (MELSS)	33 experts, and 369 instructors, students, and alumni from five universities	5-point scale, quantitative	Technical system quality, educational system quality, content and information quality, service quality	√	√	√
ÖZkan et al., 2020	Evaluation criteria of search engine optimization (SEO)	70 Turkish industrial engineering departments' websites	Quantitative form, qualitative/quantitative	Performance, design, content, meta tags, backlink, technical		√	
Velasquez and Evans, 2018	A website's assessment spreadsheet protocol	46 postgraduate students	A spreadsheet protocol, quantitative	Accessibility of websites, available online resources, library staff's responses	√	√	√
Liu et al., 2011	Evaluation criteria for Web usability of English learning websites	160 university students and seven learning technology and English teaching experts	The derived criteria combined with a checklist, qualitative/quantitative	Web usability, learning materials, functionality of assisting language learning, technology integration, and learner preferences	√	√	√
Yang and Chan, 2008	Set of evaluation criteria for English learning websites	Eight students and eight English teachers selected from junior high schools, 17 experts	4-point scale, qualitative/quantitative	General information, integrated English learning, listening, speaking, reading, and writing	√		

and trust of website services after application. Hassanzadeh et al. (2012) investigated the success factors of digital learning systems and proposed that user loyalty is a factor affecting the system's success. Loyalty signifies users' dependence on the platform and the willingness to recommend the platform. This study considered that teachers are different in sensitivity to technical efficiency, interface design, and overall design. Thus, in terms of assessing the rationality and effectiveness of online technology, we combined users' perception of the development speed of the platform and users' willingness to recommend the platform. Overall, the evaluation of online teacher learning platforms is conducted from the aspects of effectiveness, style, and development in terms of technical quality.

Service quality

Service quality is one of the key factors influencing learner satisfaction. Examining the quality of digital learning services, user satisfaction, and loyalty, reported that organizational management and learning support, as well as course quality, could affect learners' satisfaction. Online learning services included similar human services, resources, and tool services. Velasquez and Evans (2018) focused on the public library website staff's response time to users in terms of service quality and revealed that user satisfaction with the website could be increased by refining the staff's service. Devi and Verma (2018) included webpage emergency response and tools as one of the quality evaluation items when assessing library websites. Fuentes and Martínez (2018) involved learning support tools and learning resources in assessing website quality when evaluating English learning websites. Moreover, the evaluation of learning services should reflect life-oriented characteristics. Using a design-based research method to propose an evaluation of English learning websites, Liu et al. (2011) analyzed those websites from the aspects of network availability and learning materials, as well as functions that assist language learning, technology integration, and learner preferences. The evaluation indicator system not only accentuates the platform that provides the systematic learning of course content for teachers but also focuses on providing extensive resources and functional applications. Regarding the evaluation of resources and functions, the indicator focuses on the platform's richness and diversity, while it focuses on personalization in terms of learning services.

With the diversification of online learning methods for teachers and personalized online learning trends in data analysis, the formulation of service quality indicators comprehensively considers manual services, such as answering questions and providing guidance. In addition, developing data analysis and diagnosis services is included, as well as supporting teachers' teaching, function modules, and apps for behaviors like research, discussion, and reflection. Regarding the evaluation of service items, this study focused on the effectiveness of individualized assistance, as well as the relevance and availability of related tool functions and services.

Initial scale

We constructed the teacher perception scale of the quality of online learning platforms for TPD (TPS-Online-TPD), including 27 question items (Table 2). Moreover, the scheme provided the meaning described by each indicator.

Technical quality measured the platform performance and the rationality of the functional interface design. The technical quality evaluation included the evaluation of users' dependence on the platform and the development of perception, including the following: efficacy (3 items), style (4 items), and development (4 items). An example is "I think the overall layout of the platform is reasonable."

Content quality measured the quality of all content, including online courses and teaching resources. The evaluation involved not only the authority, rationality, and reliability of the resource itself but also the learning process and results, including resources (4 items), effectiveness (2 items), and method (3 items). An example is that "The teaching content and resources provided on the platform are authoritative."

Service quality measured the rationality and effectiveness of platform support services. It focused on the service quality that a platform is able to deliver, including: help (2 items), functional tools (3 items), and guidance (2 items). An example: "The analysis of learning data provided by the platform is very good for my learning."

Preliminary test

A preliminary test TPS-Online-TPD containing 27 items was developed and distributed for online participation. The target population was primary and secondary school in-service teachers who had just completed a period of online training. They were required to complete evaluations of the learning platform they had used, based on their real experience. An exploratory factor analysis was performed on the recovered questionnaire with SPSS22, and the reliability and validity of the results were verified.

Participants

Participants were in-service teachers who completed their professional development training on different online learning platforms hosted by the National Teacher Training Center at Central China Normal University in 2018. A total of 567 questionnaires that contained items of TPS-Online-TPD were collected online. After excluding the high rate of identical answers and incomplete or blank answers, 504 valid questionnaires were obtained, with the efficiency of the questionnaire at 88.89%. Of 504 participants, 77.4% were teachers in primary schools, and 22.6% were teachers in middle schools. In addition, 41.2% of teachers held senior titles in total. Regarding gender, 71.2% are women, while 28.8% are men. For the length of their teaching careers, 3.9% of respondents had been teaching for >3 years, 16.3% from 3 to 10 years, and 73.8% for >10 years. The age information range of participants was as follows: 14.3% aged <30 years, 43.3% aged 30–39 years, 34.9% aged 40–49 years, and 7.5% aged >50 years. Finally, 11.3% of the responding teachers had no experience in online training.

TABLE 2 The dimensions and indicators of TPS-Online-TPD.

Dimensions	Indicators	Description
Technical quality	Efficacy	Speed and efficiency of platform operation, page jump, data, and resource transfer.
	Style	Reasonable and aesthetic level of platform function, interface, font, color.
	Development	The growth rate of users and resources of the platform; trust and recommendation intention of users with regard to the platform.
Content quality	Resources	The authority, reliability, and rationality of the courses and teaching resources.
	Effectiveness	The influence of network learning on the teacher's teaching theory and practice.
	Method	The reasonableness, diversity, and stimulating nature of the teaching methods and strategies.
Service quality	Help	Access to help and response time when users encounter problems.
	Functional tools	The utility and richness of the tools and functions that support teacher learning and long-term professional development.
	Guidance	The degree of user satisfaction and personalization of the learning guidance, answering of questions, and other support services provided by the platform.

Validity and reliability

First, the questionnaires were classified to some extent using exploratory factor analysis. We used the principal factor analysis, and Varimax performed factor rotation in the factor analysis process. The cumulative variance explanation rate was the proportion of the variance due to all factors to the total variance, suggesting the total influence of all factors on the dependent variable. Typically, the cumulative variance explanation rate should at least be >50%; while >70% is better, >85% is excellent. According to the cumulative variance interpretation rate, three factors with eigenvalues >1 were extracted. Table 3 shows that three common factors have eigenvalues >1; the total variance interpretation rate for the three factors is 71.523%. Hence, most of the factors were considered to have been covered. Thus, the original three-dimensionality is retained as a common factor.

The load factor was the load of a variable on a common factor, which, in turn, reflected the relative importance of the variable on the common factor. Typically, the load factor after rotation must be >0.71 to be excellent, >0.63 to be very good, and >0.55 to be good. As shown in Table 3, the load factor of only one question in this questionnaire was <0.63, while that of all questions was >0.55, indicating that the questionnaire had a high degree of correspondence with the dimensions, and all 27 questions could be reserved. Supplementary Appendix 1 lists all questions.

Then, we tested the reliability of the questionnaire using the Cronbach coefficient for the 27 selected items. Usually, a Cronbach coefficient of 0.70 is credible, while 0.70–0.98 indicates high reliability. Table 4 demonstrates that the reliability coefficients of the three dimensions are all >0.70, with some even >0.95; the overall reliability coefficient of the questionnaire reached 0.98, indicating that the questionnaire had excellent internal consistency. Next, the KMO test statistic of the screened questionnaire was analyzed. Of note, the KMO test statistic is an index used to compare the simple correlation coefficient and the partial correlation coefficient between variables. The closer the KMO value is to 1, the stronger is the correlation between variables, and the more suitable the original variables are for factor analysis. Of note, a KMO value of 0.90 indicates excellent (marvelous), while 0.8 indicates good (meritorious), and 0.70 indicates middling. The KMO value of this questionnaire was 0.98, indicating that the questionnaire was highly suitable for factor analysis. Finally, we performed a factor analysis on the screened questionnaire. The factor loading coefficients of each question are in the table, and the corresponding factors were 0.61–0.80, indicating that the questionnaire had excellent structural validity.

Expert evaluation empowerment

As indicators are of different importance, we used the expert assessment method to empower the indicators at all levels. The data were collected through an online questionnaire and calculated with an equal weight evaluation method. First, we obtained the weights assigned to the indicators by experts through a questionnaire survey. Thus, the average overall score of each indicator item was calculated per the ranking of all the fill-in options, reflecting the overall ranking of the indicator. The calculation method is:

$$S_{\mathrm{i}} = \frac{\sum \left(F_{\mathrm{ij}} \times W_{\mathrm{ij}}\right)}{F} \quad (1)$$

where S_i is an average overall score of option i; F_{ij} denotes the times of option i in position j; W_{ij} denotes the weight of option i in position j; F denotes the times fill in this question.

The option average overall score was further processed to obtain the platform indicator weight value. The calculation formula is as follows:

$$k_{\mathrm{i}} = \frac{S_i}{S_n} \times 100\% \quad (2)$$

where k_i denotes the weight of option i; S_i denotes the average overall score of option i; S_n denotes the sum of the average overall scores of all options in the dimension.

The ranking questionnaires were distributed to 14 experts, including teachers and subject supervisors from primary and

TABLE 3 Eigenvalue and accumulative rate of factor analysis.

Principal component	Initial eigenvalues			Rotation sums of squared loadings		
	Eigenvalue	Explained variance (%)	Accumulative rate of contribution (%)	Eigenvalue	Explained variance (%)	Accumulative rate of contribution (%)
1	26.42	64.43	64.43	10.59	25.84	25.84
2	1.76	4.28	68.71	9.88	24.09	49.92
3	1.15	2.81	71.52	8.86	21.60	71.52
4	0.77	1.87	73.40			
5	0.68	1.66	75.06			

secondary schools and eight instructors and advisors from universities, all of whom specialized in teacher mentorship, teacher preparation, and online learning programs. They have about an average of 23 years working experience. Supplementary Appendix 2 provides the experts' primary background information. According to the formula above, the weights of the indicators at all levels of the platform were calculated, as shown in Table 5.

TABLE 4 Analysis of the final questionnaire's validity and reliability.

Measure	Item	Factor loading	α	KMO
Technical quality	T1. The resources on the platform are growing fast.	0.74	0.96	
	T2. I would like to recommend this platform to friends and colleagues.	0.76		
	T3. I am delighted with the way the tools/sections on the webpage open, run, and jump.	0.80		
	T4. The webpage on the platform runs smoothly.	0.68		
	T5. The running speed of webpages on the platform and the uploading and downloading of resources are fast.	0.68		
	T6. It is very efficient to communicate and share resources through the platform.	0.71		
	T7. I think the overall page layout of the platform is reasonable.	0.71		
	T8. The colors and fonts of the platform page are well designed.	0.69		
	T9. The functional navigation of the webpages on the platform is clear.	0.72		
	T10. The colors and fonts of the platform pages are well designed.	0.67		
	T11. The functions and resources of the platform can support my long-term use of the platform.	0.67		
Content quality	C1. On the platform, I can retrieve many resources that I need for teaching.	0.71	0.95	
	C2. I trust the learning content and resources provided by the platform.	0.74		
	C3. The teaching content and Q and A provided on the platform are authoritative.	0.71		
	C4. The content on the platform is in line with the needs of our teachers' learning and development.	0.75		
	C5. The content of this learning platform is closely related to my teaching practice.	0.78		
	C6. After stages of learning, my teaching concepts and ideas have changed.	0.78		
	C7. The teaching content on the platform and the teacher's teaching are exciting and help me maintain my continuous enthusiasm for learning.	0.68		
	C8. The learning method of the course is suitable for my professional development needs.	0.68		
	C9. I am satisfied with the variety of learning methods available on the platform.	0.61		
Service quality	S1. There are clear channels on the platform to help with problems.	0.67	0.96	
	S2. The platform has fast support services.	0.69		
	S3. The functional tools provided on the platform are very useful.	0.72		
	S4. The functional tools available on the platform are plentiful.	0.75		
	S5. The analysis of learning data provided by the platform is very beneficial to my learning.	0.67		
	S6. The platform provides personalized services, such as learning strategies and learning guidance.	0.71		
	S7. The quality of support services provided by the platform, such as answering questions and guidance, is high.	0.68		
Total			0.98	0.98

TABLE 5 Indicator weights.

Dimensions	Weight	Indicators	Weight
Technical quality	0.29	Efficacy	0.42
		Style	0.25
		Development	0.33
Content quality	0.49	Resources	0.42
		Effectiveness	0.30
		Method	0.28
Service quality	0.23	Help	0.36
		Functional tools	0.29
		Guidance	0.36

Application study

Data analysis

Study process

To validate the practicability of TPS-Online-TPD, we applied it to the annual evaluation of the National Online Teacher Training Platform in Central China in November 2019 and July 2020, respectively, which was before and after the lockdown of Wuhan. The Pearson correlation test and independent-sample *t*-test were conducted to ensure the data from two different groups of teachers were available for comparison before conducting a comparative analysis. Of note, the difference between the two evaluations could represent platform improvement. To further investigate the efficiency of each part of the improvement work, we further calculated the ratio between quality improvement and workload, to comprehend which improvement work could attain the fastest improvement of user satisfaction with the least workload.

Participants

We commissioned the development agency of the training platform; they invited 200 teachers (100 in November 2019 and 100 in July 2020) to fill in the questionnaires carefully in the form of a formal invitation letter through the administrative personnel of the school. All invited in-teachers were from primary and secondary schools in urban areas who had online learning experience before. We checked the collected questionnaires and found that all questionnaires were available in terms of time spent answering questions and repetition rates of the same answers.

Data preprocessing

We performed the following analyses to ensure that the data collected from the two evaluation surveys were effective for comparative analysis. First, the Pearson correlation test was conducted on the correlation between scores measured in 2019 and 2020 and the background factors of samples. The significance coefficient *P*-value obtained was >0.05, as shown in Table 6, suggesting that the correlation was not significant; thus, the personnel background factors would not affect the scoring results, and the data were usable. Then, we conducted an independent-sample *t*-test. In addition, Levene's test for equality of variances was not significant, and the corresponding *P*-value was <0.05, as shown in Table 7, indicating a significant difference between the two measurement results and that the results could be compared.

Results

Table 8 presents the results of the two weighted measurements. The scores of all indicator items were higher in 2020 than those in 2019, showing that optimizing online teacher learning platforms amid COVID-19 improved user experience in various aspects. According to the data collected during the COVID-19 pandemic, online learning was teachers' only professional development method. Thus, the influence of other forms of professional learning for teachers can be excluded from the study based on the data collected during this period.

The more significant the difference measured, the higher the improvement. We observed that the improvement in technical and service quality were the most significant, followed by the improvement in content quality. Overall, the improvement in webpage operation efficiency was the most significant in technical quality. The improvement in the resource quality in terms of content quality was relatively high, and so was the improvement in assistance and guidance in service quality, which were more significant than the improvement in functional tools.

After comparing the weighted average scores of the two measurements, we further combined the National Online Teacher Training Platform Operation Company's annual improvement efforts in each module to elucidate the benefits and effects of platform improvement. The calculation formula is shown in Eq. 3, and Table 7 presents the calculation results. The bigger the input–output ratio, the higher the benefit of the improvement.

$$P_{\mathrm{n}} = \frac{A_n - A_{nm}}{L_n} \times C \tag{3}$$

where P_n denotes efficiency improvement in dimension *n*; A_n denotes the weighted average score of dimension *n* this year; A_{nm} denotes the weighted average score of dimension *n* last year;

TABLE 6 Correlation between the total platform score and variables.

Version of platform	Coefficient	Variables				
		Seniority	Age	Title	Online training experience	Gender
2020	Correlation	−0.01	−0.47	0.03	−0.05	−0.07
	Significance	0.91	0.64	0.80	0.66	0.52
2019	Correlation	−0.09	−0.06	−0.07	−0.12	0.13
	Significance	0.36	0.58	0.48	0.23	0.20

TABLE 7 Independent-sample *t*-test of the two versions of the platform.

Version of platform	Mean	SD	Difference of mean	SE	*P*	95% Confidence intervals		Levene's test for equality of variances	
						Upper limit	Lower limit	*F*	Significance
2020	4.20	0.42	0.14*	0.07	0.04	0.005	0.27	3.12	0.079
2019	4.06	0.53							

*$P < 0.05$.

TABLE 8 Comparison of workload, improvement, and improved efficiency.

Dimensions/key indicators	2019	2020	Workload	Improvement	Improved efficiency
Technical quality	4.15	3.97	65	0.18	28.15
Content quality	4.25	4.17	58	0.09	14.66
Service quality	4.18	4.02	88	0.16	18.18
Efficacy	4.13	3.87	37	0.26	70.27
Style	4.19	4.05	10	0.15	146.00
Development	4.16	4.04	18	0.12	65.00
Resources	4.25	4.15	41	0.11	25.61
Effectiveness	4.28	4.21	–	0.07	–
Method	4.22	4.15	14	0.07	50.00
Help	4.18	4.01	25	0.17	66.00
Functional tools	4.22	4.07	18	0.15	81.67
Guidance	4.17	4.01	45	0.17	36.67
Total	4.21	4.08	211	0.13	6.16

L_n is the workload of dimension n; C denotes the fixed rounding factor with a value of 10,000.

As shown in Table 8, it is apparent that the degrees of improvement and improved efficiency in technical quality were higher than those of the quality of service and content. Among the key indicators, the improvement in the interface style was the most efficient. Regarding the optimization in content quality, improvement in the teaching method could lead to high efficiency. The improvement of each dimension of service quality was relatively balanced, and the improvement of the platform function tools was the most efficient.

Discussion and conclusion

Based on the IS success model and previous relevant research, this study developed TPS-Online-TPD, which provides a set of key indicators for assessing the quality of online learning platforms for teacher professional development from three aspects: technical quality, content quality, and service quality. TPS-Online TPD is further used to analyze the optimization of the National Teacher Training Platform amid the COVID-19 outbreak in China. By comparing the quality evaluation of the platform before and after the lockdown of Wuhan. The findings revealed two main points. First, the optimization of the platform's technical quality and service quality can better improve teachers' learning experience, compared with the optimization of the platform's content. Second, more significant improvements in the teacher learning experience can be generated when the design style, tool functions, operation efficiency, and teaching methods of the platforms are optimized. These findings corroborate the previous studies that examined key factors on continuance intention in technology-assisted learning (Yang et al., 2022) and technological barriers to learning outcomes (He and Yang, 2021).

The findings of this study have crucial practical implications for developing online learning platforms for teacher professional development; however, these have some limitations in terms of generalizability. Notably, TPS-Online TPD is designed and developed primarily based on and focused on the circumstance of primary and secondary school teachers in China, which does not include other cultural backgrounds and college instructors. In addition, technology integration in teacher professional development is a long and dynamic process (Cai et al., 2019; Zhou et al., 2022). This study only measured the current teacher perceptions of the quality of online learning platforms. Thus, we suggest that future research should consider other cultures and college instructors, and periodically assess the quality of online learning platforms for teacher professional development to obtain additional accurate information and improve the use of the platform.

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 authors.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JZ and HY: conceptualization and writing–review and editing. JZ and ZC: methodology. JZ and WG: formal analysis. WG and ZL: investigation. BW, ZC, and HY: resources. JZ: writing–original draft preparation. BW: project administration. BW, ZC, and ZL: funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (Project Nos. 62077022 and 62077017).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.998196/full#supplementary-material

References

Cai, J., Yang, H. H., Gong, D., MacLeod, J., and Zhu, S. (2019). Understanding the continued use of flipped classroom instruction: A personal beliefs model in Chinese higher education. *J. Comput. High. Educ.* 31, 137–155. doi: 10.1007/s12528-018-9196-y

Cobo Romani, J. C., Mata-Sandoval, J. C., Quota, M. B. N., Patil, A. S., and Bhatia, J. (2022). *Technology for teacher professional development navigation guide: A summary of methods*. Washington, DC: World Bank Group.

DeLone, W. H., and McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Inform. Syst. Res.* 3, 60–95.

DeLone, W. H., and McLean, E. R. (2002). "Information systems success revisited," in *Proceedings of the 35th annual hawaii international conference on system sciences*, (Big Island, HI: IEEE), 2966–2976. doi: 10.1109/HICSS.2002.994345

DeLone, W. H., and McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *J. Manage Inform. Syst.* 19, 9–30. doi: 10.1080/07421222.2003.11045748

Desjardins, F., and Bullock, S. (2019). Professional development learning environments (PDLEs) embedded in a collaborative online learning environment (COLE): Moving towards a new conception of online professional learning. *Educ. Inf. Technol.* 24, 1863–1900. doi: 10.1007/s10639-018-9686-6

Devi, K. K., and Verma, M. K. (2018). Web content and design trends of Indian Institute of Technology (IITs) libraries' website: An evaluation. *Collnet. J. Scientomet.* 12, 165–181. doi: 10.1080/09737766.2018.1433100

Fuentes, E. M., and Martínez, J. J. R. (2018). Design of a checklist for evaluating language learning websites. *Porta Linguarum* 30, 23–41.

Greenberg, A. D. (2009). *Critical success factors for deploying distance education technologies*. Duxbury, MA: Wainhouse Research.

Guskey, T. R. (1994). Results-oriented professional development: In search of an optimal mix of effective practices. *J. Staff Dev.* 15, 42–50.

Hassanzadeh, A., Kanaani, F., and Elahi, S. (2012). A model for measuring e-learning systems success in universities. *Expert Syst. Appl.* 39, 10959–10966. doi: 10.1016/j.eswa.2012.03.028

He, X., and Yang, H. H. (2021). "Technological barriers and learning outcomes in online courses during the Covid-19 pandemic," in *Proceedings of the international conference on blended learning*, (Cham: Springer), 92–102. doi: 10.1007/978-3-030-80504-3_8

Laperuta, D. G. P., Pessa, S. L. R., Santos, I. R. D., Sieminkoski, T., and Luz, R. P. D. (2017). Ergonomic adjustments on a website from the usability of functions: Can deficits impair functionalities? *Acta Sci.* 39, 477–485. doi: 10.4025/actascitechnol.v39i4.30511

Lin, H. F. (2007). Measuring online learning systems success: Applying the updated DeLone and McLean model. *Cyberpsychol. Behav.* 10, 817–820. doi: 10.1089/cpb.2007.9948

Liu, G. Z., Liu, Z. H., and Hwang, G. J. (2011). Developing multi-dimensional evaluation criteria for English learning websites with university students and professors. *Comput. Educ.* 56, 65–79. doi: 10.1016/j.compedu.2010.08.019

Liu, Y. Q., Bielefield, A., and Mckay, P. (2017). Are urban public libraries websites accessible to Americans with disabilities? *Universal Access Inf.* 18, 191–206. doi: 10.1007/s10209-017-0571-7

Manzoor, M., Hussain, W., Sohaib, O., Hussain, F. K., and Alkhalaf, S. (2018). Methodological investigation for enhancing the usability of university websites. *J. Amb. Intel. Hum. Comp.* 10, 531–549. doi: 10.1007/s12652-018-0686-6

Mousavilou, Z., and Oskouei, R. J. (2018). An optimal method for URL design of webpage journals. *Soc. Netw. Anal. Min.* 8:53. doi: 10.1007/s13278-018-0532-z

Ng, W. S. (2014). Critical design factors of developing a high-quality educational website: Perspectives of pre-service teachers. *Issues Informing Sci. Inf. Technol* 11, 101–113. doi: 10.28945/1983

ÖZkan, B., Zceylan, E., Kabak, M., and Dadeviren, M. (2020). Evaluating the websites of academic departments through SEO criteria: A hesitant fuzzy linguistic MCDM approach. *Artif. Intell. Rev.* 53, 875–905. doi: 10.1007/s10462-019-09681-z

Pant, A. (2015). Usability evaluation of an academic library website: Experience with the Central Science Library, University of Delhi. *Electron. Library* 33, 896–915. doi: 10.1108/EL-04-2014-0067

Pengxi, L., and Guili, Z. (2018). "Design of training platform of young teachers in engineering colleges based on TPCK theory," in . *Proceedings of the 2018 4th international conference on information management (ICIM)*, (Oxford: IEEE), 290–295. doi: 10.1109/INFOMAN.2018.8392852

Santos, A. M., Cordón-García, J. A., and Díaz, R. G. (2016). "Websites of learning support in primary and high school in Portugal: A performance and usability study," in *Proceedings of the international conference on technological ecosystems for enhancing multiculturality*, (New York, NY: ACM), 1121–1125. doi: 10.1145/3012430.3012657

Santos, A. M., Cordón-García, J. A., and Díaz, R. G. (2019). Evaluation of high school websites based on users: A perspective of usability and performance study. *J. Inf. Technol. Res.* 12, 72–90. doi: 10.4018/JITR.2019040105

Schlager, M. S., and Fusco, J. (2003). Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse? *Inform. Soc.* 19, 203–220. doi: 10.1080/01972240309464

Velasquez, D., and Evans, N. (2018). Public library websites as electronic branches: A multi-country quantitative evaluation. *Inf. Res.* 23:25.

Yang, H., Cai, J., Yang, H. H., and Wang, X. (2022). Examining key factors of beginner's continuance intention in blended learning in higher education. *J. Comput. High. Educ.* 1–18. doi: 10.1007/s12528-022-09322-5 [Epub ahead of print].

Yang, Y. T. C., and Chan, C. Y. (2008). Comprehensive evaluation criteria for English learning websites using expert validity surveys. *Comput. Educ.* 51, 403–422. doi: 10.1016/j.compedu.2007.05.011

Zhou, C., Wu, D., Li, Y., Yang, H. H., Man, S., and Chen, M. (2022). The role of student engagement in promoting teachers' continuous learning of TPACK: Based on a stimulus-organism-response framework and an integrative model of behavior prediction. *Educ. Inf. Technol.* 1–21. doi: 10.1007/s10639-022-11237-8 [Epub ahead of print].

frontiers | Frontiers in Psychology

TYPE Systematic Review
PUBLISHED 11 October 2022
DOI 10.3389/fpsyg.2022.1011409

Check for updates

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University, China

REVIEWED BY
Julio Cabero Almenara,
Sevilla University, Spain
Cesar Collazos,
University of Cauca, Colombia

*CORRESPONDENCE
Daniel Burgos
daniel.burgos@unir.net

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

RECEIVED 04 August 2022
ACCEPTED 22 September 2022
PUBLISHED 11 October 2022

CITATION
Amores-Valencia A, Burgos D and Branch-Bedoya JW (2022) Influence of motivation and academic performance in the use of Augmented Reality in education. A systematic review. *Front. Psychol.* 13:1011409. doi: 10.3389/fpsyg.2022.1011409

Influence of motivation and academic performance in the use of Augmented Reality in education. A systematic review

Antonio Amores-Valencia[1], Daniel Burgos[1,2]* and John W. Branch-Bedoya[2]

[1]Research Institute for Innovation and Technology in Education (UNIR iTED), Universidad Internacional de La Rioja, Logroño, Spain, [2]Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia

The recent technologies rise today as a tool of significant importance today, especially in the educational context. In this sense, Augmented Reality (AR) is a technology that is achieving a greater presence in educational centers in the last decade. However, Augmented Reality has not been explored in depth at the Secondary Education stage. Due to this, it is essential to analyze and concentrate the scientific research developed around this educational technology at that stage. Therefore, the aim of this research is to describe the influence that Augmented Reality shows on the motivation and academic performance of students in the Secondary Education stage. In relation to the methodology, a systematic review of the literature has been conducted using the Kitchenham protocol, where several factors have been analyzed, such as subjects, activities, and electronic implementation devices, together with the effects on motivation and student's academic performance. The Scopus and Web of Science (WoS) databases have been used to search for scientific papers, with a total of 344 investigations being analyzed between 2012 and 2022. The methodological stages considered were the formulation of research questions, the choice of data sources, search strategies, inclusion and exclusion criteria and quality assessment, and finally, data extraction and synthesis. The results obtained have shown that the use of AR in the classroom provides higher levels of motivation, reflected by factors such as attention, relevance, confidence, and satisfaction, and reflects better results in the tests carried out on the experimental groups compared to the control groups, which means an improvement in the academic performance of students. These results supply a fundamental theoretical basis, where the different teachers should be supported for the incorporation of AR in the classroom, since how this educational technology has been shown offers great opportunities. Likewise, the development of research in areas not so addressed can further clarify the generality of AR based on its influence on learning. In addition, the fields of natural sciences and logical-mathematical have been the most addressed, managing to implement their contents through object modeling. In short, this research highlights the importance of incorporating Augmented Reality into all areas and educational stages, since it is a significant improvement in the teaching and learning process.

KEYWORDS

Augmented Reality, motivation, academic performance, high school, systematic review

Introduction

Today, the use of technological devices is present in most of the activities that people conduct daily. This is largely due to the introduction of technology in the development and implementation process in many fields (Huang and Liao, 2017; Juan, 2019; McLean and Wilson, 2019; Schaffernak et al., 2020; Rezaee et al., 2021). In this sense, education cannot be relegated to the background, much less leave aside the new devices and existing technological tools (Macías-González and Manresa-Yee, 2013). For this reason, the teaching and learning process must adapt to today's society, to work in parallel with the demands of an increasingly changing market.

Under this premise, Information and Communication Technologies (ICT) are the tools that have brought about a great transformation of daily life, whatever the aspect that may be glimpsed. This fact, together with the creation of new jobs, requires the obligation of continuous training, which produces people fully prepared for the future changes that lie ahead (Cabero-Almenara, 2015). In addition, the introduction of educational technology supposes an enhancement of student motivation, which translates into better academic performance (Area-Moreira et al., 2018).

On the other hand, the little or no use of technological tools in the classroom shows very worrying levels of motivation, especially in Secondary Education, so teachers must reconsider their position in search of a more appropriate environment for the reality of their needs students (Amores-Valencia and De-Casas-Moreno, 2019). According to Hernandez (2017) the lack of motivation of the students is not only an obstacle for the learning of concepts, but also a problem in the daily teaching work of the teachers, because the students are inattentive, and it often generates disorder. In this negative environment, educational technologies play a significant role because they can be seen as a powerful motivational tool (Vedadi et al., 2017). In addition, it is essential to train students in critical thinking, emphasizing the reflective part, as well as giving greater efforts in the face of motivational and emotional difficulties such as low expectations, disinterest and high pressure from students (Barroso-Osuna and Cabero-Almenara, 2013).

One of the most attractive, dynamic and interactive proposals is the integration of Augmented Reality in the classroom, as shown by numerous studies in recent years. In reference to the origin of the concept, it was Azuma (1997) who defined it as a technology that enhances the sensory uptake of people since it can combine real and virtual elements in an interactive scenario in real time. In this sense, the implementation of Augmented Reality requires working with technological devices such as tablets, smartphones, and computers that generate interaction between users, producing empathic experiences (Cabero-Almenara and Barroso-Osuna, 2016b). In addition, with the introduction of new active methodologies that will affect students in quite diverse ways in face-to-face, online and blended learning and teaching contexts (Buchner, 2021).

This work is organized as follows. The first section shows the research related to this work, specifically those that bring together Augmented Reality with motivation and academic performance in Secondary Education. The second section presents the method developed where each of the phases is broken down. The third section shows the results obtained through an exhaustive analysis of the information. In the fourth section, the research questions are discussed and answered. Finally, the relevant conclusions are developed based on the contrast of the aims set, and the results obtained.

Theoretical framework

In this sense, Augmented Reality (AR) is the technology with the greatest effect on education in recent years, enabling the coordination of the real and the virtual at the same time (Cabero-Almenara and Barroso-Osuna, 2016a). In addition, it helps to create more motivating and attractive teaching and learning scenarios that, in turn, would be impossible to conduct in the real world (Duh and Klopher, 2013; Wojciechowski and Cellary, 2013; Huang et al., 2016). Likewise, the use of this educational technology can contribute positively to the interest, motivation and performance of students (Di Serio et al., 2013 Redondo-Domínguez et al., 2014; Sommerauer and Müller, 2014; Reinoso-Peinado, 2016).

In this way, many investigations highlight the effect and influence that Augmented Reality has and will have on teaching and learning processes (Yuen et al., 2013; Bower et al., 2014; Ibáñez et al., 2014; Jerábek et al., 2014; Coimbra et al., 2015; Solak and Cakir, 2015; Fonseca-Escudero et al., 2016; Marín-Díaz, 2016; Sánchez-Bolado, 2017; Cabero-Almenara and Marín-Díaz, 2018). In addition, some scientific works are related to games, applications and illustrated books, which use AR to show new ways of learning, incorporating animations of images and videos to the illustration of the texts (Lin et al., 2017; Álvarez-Marín et al., 2020).

On the other hand, most research has considered university students, where extra motivation is predisposed as it is a non-compulsory educational stage (Fernández-Robles, 2017; Barroso-Osuna et al., 2018; Marín-Díaz et al., 2018; Gómez-García et al., 2020). Another profile of the student body studied has been primary school students where performance and motivation tend to have high values (Toledo-Morales and Sánchez-García, 2017; Wang, 2017; Kirikkaya and Başgül, 2019; Lai et al., 2019; López-Belmonte et al., 2019).

Other fundamental aspects related to motivation are consciousness and emotions. According to Collazos et al. (2021) teachers must know when and how to participate, in such a way that they adapt to the situation of each group of students, and in the opposite case, the students must be aware of what is happening in their environment with the purpose of achieve proper interaction between all members. Along the same lines, Mestre-Navas et al. (2017) shows that emotional intelligence is closely related to motivation, since it supplies the necessary skills to control one's own and others' emotions, which produces better learning. For this, it is essential that emotions are perceived, found, valued and expressed adequately and precisely. In this sense, assimilating emotions is a vital step to achieve a good understanding of knowledge. Likewise, the organization of emotions brings with it a cognitive effort, which, guided by the teacher, produces an adequate management of emotions in real-life situations.

Finally, the integration of Augmented Reality in the educational field is a fact confirmed in the multiple investigations previously developed, however, it is essential to know the influence of the gender of the students in the results obtained when this educational technology is used (Abad-Segura et al., 2020; Álvarez-Marín et al., 2020). In this sense, there is a multitude of investigations that try to find out the repercussion that gender shows according to motivation and interest, the acceptance of the use of Augmented Reality as an educational tool and the acquisition of knowledge or performance (Hsu, 2019; López-García et al., 2019). For this reason, this first section will specify the results and conclusions that have been described by the different investigations on the influence that the gender of the students shows in motivation, in the acceptance and use of learning objects and in the acquisition of knowledge, and before the use of Augmented Reality in the different educational stages.

To know what other authors have investigated about how Augmented Reality influences the motivation and performance of students in the various learning and teaching contexts mentioned, a systematic review of the literature has been carried out, based on studies that include different techniques to show this technology. Specifically, it has focused on Secondary Education, since it shows the worst rates of motivation and academic performance of students (Amores-Valencia and De-Casas-Moreno, 2019). The studies included in this systematic review have been obtained from the Scopus and Web of Science (WoS) databases, where a series of keywords have been introduced, to later discard repeated documents and keep those that met the criteria of inclusion. Inclusion and exclusion detailed in the method.

Augmented Reality in education

Regarding the contributions made by Augmented Reality to the educational field, the considerable number of applications developed to work with this educational technology in the classroom in the different subjects stand out. However, the use of this educational technology is still very restricted due to the lack of teacher training, since teachers who wish to integrate Augmented Reality in their classrooms must get the knowledge in a self-taught way, during non-school hours and on many occasions without adequate resources (González-Segredo and Hernández-Cabrera, 2022).

On the other hand, said technology presents a series of fundamental characteristics for the new educational paradigm, since it provides interaction with resources, visualization of information or creation of scenarios that enhance the understanding of concepts, which has repercussions on the learning process (Aguilar-Acevedo et al., 2022). Another characteristic that Augmented Reality grants is the choice of applying active methodologies, where the teacher plays the role of guide and the students take an active role, managing their own learning. In addition, the student imposes his own rhythm, which eases the process of understanding and getting knowledge. All of this provides autonomous learning, where the student usually shows more interest than with traditional methodologies (Amores-Valencia, 2020).

According to Cabero-Almenara and Barroso-Osuna (2016b), the relevance of Augmented Reality in the future will be especially important since it combines reality with virtuality at the same time and place. In this sense, Castro-Marcos (2022) points out that this educational technology gives the possibility of creating learning scenarios that would be impossible to develop in the real world. According to this author, these more attractive scenarios enhance the interest and motivation of the student, an aspect that undoubtedly affects their academic performance.

In this way, many investigations highlight the impact that Augmented Reality will have on teaching and learning processes (Fonseca-Escudero et al., 2016; Marín-Díaz, 2016; Sánchez-Bolado, 2017). In addition, some scientific works are related to games, applications and illustrated books, which use Augmented Reality to show new forms of learning, incorporating animations of images and videos to the illustration of the texts (De-Paiva-Guimarães and Farinazzo-Martins, 2014; Lin et al., 2017; Álvarez-Marín et al., 2020).

Other applications of Augmented Reality are focused on the professional field since virtual scenarios can be created and implemented that do not pose a danger or physical harm to people. Thus, a safe and at the same time adequate training can be conducted, where the student's skills are experienced in virtual situations (Akçayir et al., 2016). This feature gives companies a fantastic opportunity to train their employees, guaranteeing their health (Carlton, 2017). Likewise, with the rise of e-learning training, Augmented Reality has made it possible to work on practical content, enabling the acquisition of knowledge that would only have been possible in face-to-face training (Reinoso-Peinado, 2016).

In reference to the acceptance and use of Augmented Reality learning objects, the authors Wang et al. (2017) and Bursztyn et al. (2020) state that no significant difference is observed between male and female students. These investigations show that factors such as perceived enjoyment, perceived usefulness and perceived ease of use show practically identical values in men and women. Although it is true that one cannot speak of the absolute non-existence of gender disparity, since some research has highlighted that said difference exists, emphasizing the real and palpable difference in the perception that students present in the use of Augmented Reality based on gender (Dirin et al., 2019; Park et al., 2019). This fact indicates that the digital divide between genders continues to exist, although fortunately over the years it is being reduced (Cabero-Almenara et al., 2019b). Thanks to the effort of educational centers to equate all students in an identical way, their digital competence is not related to gender (Hohlfeld et al., 2013).

In conclusion, Cabero-Almenara and Marín-Díaz (2018) show a series of educational possibilities of Augmented Reality:

- It improves the real contents easing its understanding.
- Develop multimedia training environments.
- Promotes online learning.
- Cut non-relevant information that hinders the acquisition of knowledge.
- Create safe learning scenarios.
- Helps active and productive learning of Augmented Reality resources.
- Incorporate extra information in the form of illustrations, videos, or audios.
- Design more attractive simulators for learning.
- It eases the visualization of contents from various perspectives.
- It promotes the use of active strategies or methodologies.

Motivation and academic performance

For many years, cognitive variables have been the most analyzed in the learning process, but in the 1990's there was an increase in the number of studies seeking the influence of the motivational aspect in the design of congruent models that explain academic performance (Pressley et al., 1992; Pintrich et al., 1993; Borkowski, 1994; García and Pintrich, 1994; Pintrich, 1994; Schunk and Zimmennan, 1994; García, 1995; Boekaerts, 1996). However, for these authors, there must be a relationship between the cognitive and the motivational to obtain an improvement in academic performance.

Motivation tries to activate student behavior aimed at achievements and goals if there is an effort behind it during the process. For this reason, motivation encompasses many variables such as expectations of achievement, relative attributions, self-value, self-esteem and self-concept (González-Pienda, 2003). In this sense, Weiner (1986) states that motivated behavior is obtained based on the possibilities of achieving goals and their value. These two components manage the success or failure of a student and are set up by the relative attributions that the student has of himself. Therefore, Weiner (1985, 1986) defines attributions as the greatest determinant of motivation in terms of the results and academic performance of the student.

In reference to self-concept, the research conducted focused on the student's academic behavior, due to the importance of knowing how academic goals are obtained and in what context they are achieved (González-Pienda et al., 1997). These studies confirmed the relationship between self-concept and the academic performance of students, however, there are still doubts between the processes that make this relationship possible and its directionality. In such a way that there is research that corroborates the reciprocity between self-concept and academic performance (Marsh and Yeung, 1997) and others that expose the unidirectionality of performance on the self-concept of students (Helmke and van Aken, 1995).

According to Núñez (2009), the skills and competencies of a student are not enough to improve their academic performance, but it is necessary to consider their motivation. This statement comes to extol the importance of motivation in a student's performance, since it does not depend exclusively on the knowledge and skills that he has. Similarly, Garrido-Macías et al. (2013) states that the effect of motivation on a student's academic performance grows the higher their self-esteem and their assessment of tasks.

For all that has happened, the motivational variables that are subject to the types of motivation are directly related to academic performance and are decisive in the educational process (Mascarenhas et al., 2005; Alonso-Tapia and Ruiz-Díaz, 2007; Martín et al., 2008; Miñano-Pérez and Castejón-Costa, 2011; Barca-Lozano et al., 2012). On this occasion, it can be affirmed that intrinsic and extrinsic motivation can be addressed jointly or separately, since they are not opposed based on the academic performance presented by students (Usán-Supervía and Salavera-Bordás, 2018). Therefore, it is vital to continue analyzing the relationship between motivation and academic performance, to learn more about the interrelated factors and thus improve the teaching and learning process (Jerez-Carrillo, 2021).

Finally, the motivation that students present in relation to their gender is reflected in the different investigations carried out by Hanafi et al. (2017) and Buchner (2021), where it is stated that the behavior that students present is different from that the students present and, therefore, a significant difference in motivation is appreciated before the use of Augmented Reality. However, other authors indicate that the difference shown is negligible to be considered (Bursztyn et al., 2017; López-Belmonte et al., 2019). Regarding academic performance, Del-Rio-Guerra et al. (2019) and Gómez-Tone et al. (2020) state that students and students show the same results when Augmented Reality is applied in classrooms. This means that the academic performance of students is invariant based on their gender. This information is valuable since it gives teachers the possibility of implementing this educational technology without the need to take this feature into account. However, one must always be aware of the disparity that exists in each of the classrooms, since sometimes a differentiation can be found between male and female students in the acquisition of knowledge (Chen et al., 2021).

The bibliographical reviews on the use of AR in Secondary Education reported up to now have almost never brought together two of the great educational factors, such as motivation and academic performance (Martín-Gutiérrez and Meneses-Fernández, 2014; Liu et al., 2019). For this reason, this research aims to analyze the influence that Augmented Reality shows on the motivation and academic performance of students in the Secondary Education stage.

Methods

General guidelines

This research has been developed through the systematic review of the literature (RSL) process, based on the proposal of Kitchenham and Charters (2007). According to the authors, this protocol requires an exhaustive, objective, and reliable general description, which is governed by defined and strict steps. Specifically, the steps followed for the development of the systematic review were the following:

1. Planning the review

- Identification of the need for a review
- Specifying the research questions

2. Conducting the review

- Identification of data sources
- Selection of search strategies
- Inclusion and exclusion criteria
- Study quality assessment
- Data extraction
- Data synthesis

3. Results report

- Included and Excluded Studies
- Interpretation of results
- Formatting the report

In this way, the analysis of the literature has been developed under the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA; Urrútia and Bonfill, 2010; Moher et al., 2015).

Planning the review

Given the research studies analyzed, where the factors of motivation and academic performance have been worked on individually, it is necessary to develop a work where these factors are grouped and detailed, particularizing in the Secondary Education stage, since it is about a stage where the lowest levels are shown (Picó-Lozano, 2014; González-Valenzuela and Martín Ruiz, 2019; Jerez-Carrillo, 2021). For this reason, the investigations that combine these two dimensions have been brought together, in such a way that the following research question can be answered:

What is the status of the use of Augmented Reality through markers in terms of population, interventions, comparators, results and study designs, considering studies between 2012 and 2022 included in two interdisciplinary databases: Scopus and Web of Science, in order to know the impact on motivation and academic performance in students of the Secondary Education stage?

According to this main research question, a series of research sub-questions were defined:

RQ1: What subjects and groups are the recipients of educational activities with AR?

RQ2: What technological devices have been used to generate and/or run AR applications?

RQ3: How are educational activities implemented with AR in the classroom?

RQ4: What motivational impact do students have based on the use of AR?

RQ5: How does the use of AR influence the academic performance of students?

Conducting the review

Identification of data sources

The search for jobs was conducted using the Scopus and Web of Science (WoS) databases, relevant scientific content platforms, since they bring together a multitude of scientific publications from various areas of knowledge. Specifically, they host a multitude of works related to Augmented Reality in secondary education.

On the other hand, these two databases allow searching in advanced structures thanks to the use of logical operators, which fit perfectly to the particularities of the systematic review proposed in this research. In addition, the use of filtering tools and bibliometric analysis provide excellent information to the work presented.

Selection of search strategies

The search strategies are one of the high points of the research, since the information available in the databases must be filtered, in such a way that the selected works allow answering the research questions posed and, consequently, fulfilling the marked target. According to Kitchenham et al. (2009) search strategies make it possible to assess the integrity of the information search.

In reference to this premise, the search strings were defined in such a way that the defined keywords could be reached, and in turn answer the research questions raised.

The structured search used to search for jobs was conducted on July 14, 2022, and followed the following format according to each database:

The search string adapted to the syntax required by the Scientific Information Institute-Scopus database was as follows: (TITLE-ABS-KEY ("Augmented Reality" OR "augmenting reality" OR "AR") AND TITLE-ABS -KEY (("motivation" OR "performance)) AND TITLE-ABS-KEY ("education")) AND DOCTYPE (ar OR cp) AND PUBYEAR > 2011 AND PUBYEAR < 2022 AND (LIMIT-TO (LANGUAGE, "English")).

On the other hand, the search string adapted to the syntax required by the Institute for Scientific Information-Web of Science database was as follows: TOPIC: (("Augmented Reality" OR "Augmenting reality" OR "AR")) AND TOPIC ((motivation OR "academic performance")) AND TOPIC: (("education")) AND YEAR PUBLISHED (2012–2021). Refined by: LANGUAGES: (ENGLISH) AND TYPES OF DOCUMENTS: (ARTICLE).

Inclusion and exclusion criteria

In the selection of studies, those works that meet the conditions to be considered in the RSL are chosen, under the premise of the inclusion and exclusion criteria. In the study selection process, works were included in which the areas of application, target groups, technological tools used, motivation and academic performance could be identified. In addition, the duplication of references and their subsequent elimination was conducted using Microsoft Excel software.

According to Kitchenham and Charters (2007), studies can be selected by title and abstract, obtaining a complete copy of them. Based on these suggestions, the study selection criteria were detailed, which in turn included the keywords and search strings defined from the research questions:

- In the title and abstract, the sequence of words "Augmented Reality" or "augmenting reality" or "AR" should be included.
- The abstract must contain the sequence of words "high school."
- And in the summary the term "motivation" or "academic performance" should appear, or both words at the same time.

The eligibility criterion taken to include and exclude studies was if the word appeared, it was marked with the number 1, otherwise, it was indicated with the number 0. However, in cases where the title and abstract were not enough to determine its inclusion or exclusion, the authors evaluated all the content of the work.

To clarify the selection criteria, the following function was detailed in the Microsoft Excel software:

IF(AND(TITLE=1;ABSTRACT=1; COUNT.IF(ABSTRACT:ABSTRACT;1)≥1);"candidate article;" "no").

Study quality assessment

One of the most relevant sections of the systematic review is the evaluation of the quality of the study, since it involves determining those works that allow an adequate response to the research questions and, therefore, fulfill the stated objective. For this reason, it is necessary to analyse the results without any type of interference and mistake, counting on the appropriate studies for the proposal (Carrizo and Moller, 2018).

TABLE 1 Quality assessment checklist.

Level	Description	Score
Si	Information is explicitly defined/evaluated	1
Partially	Information is implicit/stated	0.5
No	Information is not inferable	0

According to the authors Kitchenham and Charters (2007), quality verification questions must be defined. Therefore, a questionnaire was designed based on seven items that outline the quality of the study, in such a way that they were scored to know a general measure of the quality of the selected works. In this sense, the questions were adapted to the present study, and determined the relevance of these works around the deepening toward a complete reading and subsequent analysis.

The quality questions that are developed below allowed minimizing the bias of the study and maximizing both external and internal validity.

- Are application areas and target groups established in Secondary Education?
- Does the document describe the electronic devices and technological applications used for educational activities with AR?
- Does the document indicate the form of RA application conducted?
- Were the users who participated in the creation of content fully defined?
- Is the contribution of AR to student motivation clearly described and defined?
- Is achievement included as one of the main contributions of AR in Secondary Education?
- Are all research questions answered?

The quality assessment checklist describes the score based on the quality level of the article. Each of the questions was evaluated using the following information (Table 1).

Articles were included and classified as "full reading article" in the following stages if the sum of the criteria was >4 points.

Data extraction

The software used to manage the data and analyse the information of the selected works were Mendeley and Microsoft Excel.

In the case of Microsoft Excel, it was used to manage the articles resulting from the search in the scientific databases, eliminate duplicate references and classify the information of each article. The workbook is made up of several sheets, where each of the phases is documented.

TABLE 2 Acronyms to classify information.

Source	Acronym
Application areas or subjects	Biology (B)—Mathematic (M)—Languages (L)—Technology (T)—Physical Education (PE)—History (H)—Chemistry (C)—Health Education (HE)
Target groups	7° Grade (7G)–12–13 years 8° Grade (8G)–13–14 years 9° Grade (9G)–14–15 years 10° Grade (10G)–15–16 years 11° Grade (11G)–16–17 years 12° Grade (12G)–17–18 years
AR activities in educational settings	Discovery-based Learning (DL)—Objects Modeling (OM)—AR Books (B)—Skills Training (ST)—AR Gaming (G)
Electronic device	Computer (C)—Mobile Phone (MP)—Tablet (T)—Glasse (G)
AR technology software or application	Lightining Studios (LS)—Unity/Vuforia (UV)—ARDehaes toolkit (AT)—Aurasma (A)—RAVVAR (R)—Metaverse Studio (MS)
Content creation	Designed by students (DS)—Designed by teacher (DT)—Designed by external persons (DE)
Motivational level	Very high (VH)—High (H)—Medium (M)—Low (L)—Very Low (VL)
Academic performance level	Very high (VH)—High (H)—Medium (M)—Low (L)—Very Low (VL)

With respect to Mendeley, it was used to bring together the candidate articles, and highlight the highlighted information, underlining with a different color depending on the category.

The data extraction process was developed in three stages.

- Information analysis: the analysis and classification of the article information was conducted from the bottom up. The text fragments that answer the research questions were highlighted with different colors, using the Mendeley tool. This action allowed further reading and detailed analysis and classification.
- Information classification: label codes to assign a representative meaning to the highlighted elements. The information was defined synchronously with the Information Analysis stage. Table 2 shows the codes considered for each of the research questions.
- Information extraction: Each text segment highlighted in the information analysis stage is classified according to the code established in the classification stage.

TABLE 3 Records obtained.

Criteria	Filters	Scopus	Web of Science (WoS)
Restriction	Topic (title, abstract, and keywords)	621	634
Period	2012–2022	536	547
Document type	Articles and conference proceedings	473	348
Language	English	440	285
Eligibility	High School	171	173
Total		344	

A spreadsheet is needed to process the information generated at this stage https://alumnosunir-my.sharepoint.com/:x:/g/personal/antoniojesus_amores916_comunidadunir_net/EY58lmoHHVpMvxPUV5FVKj8BB-NvKu_65vT66VQ6G5Cuvg?rtime=udITTMSg2kg.

Data synthesis

The data was tabulated and displayed to represent:

- The different areas or subjects of application and target groups that participated in the articles.
- Directions of Augmented Reality in educational activities.
- Electronic devices, applications and software used in the different investigations.
- The creation of educational activities using Augmented Reality.
- Level of motivation observed in the articles.
- Degree of performance obtained thanks to the use of Augmented Reality.

Results

This section is structured in response to the research questions, after going through the analysis process developed in the previous section. For this, the protocol chosen during data extraction has been considered. Further, data extracted from the review protocol are consolidated in the spreadsheet: https://bit.ly/3znd49h.

Included and excluded studies

This first section breaks down the results obtained thanks to the search strings entered and the inclusion and exclusion criteria developed.

TABLE 4 Number of investigations chosen.

Criteria	Papers
Elected papers	43
Excluded papers	215

TABLE 5 Full reading articles included.

Criteria	Papers
Full reading papers	13
Excluded papers	30

The first step was to introduce specific search strings based on each of the scientific databases used. This process resulted in a multitude of investigations, Table 3 shows the records obtained:

It should be noted that 344 results were obtained between both databases, and the search in them occurred on July 14, 2022.

The second step was to eliminate the duplication of investigations that could be seen in both databases. For this, the Microsoft Excel tool was used, where the number of scientific articles was reduced to a total of 258 works. Next, the eligibility criteria based on incorrect titles and abstracts were considered, excluding a total of 215 investigations based on this criterion.

The investigations selected once the eligibility criteria of the systematic review were addressed are described in Table 4.

The investigations selected once the quality assessment of the systematic review was addressed are described in Table 5.

Figure 1 presents the phases and results of the number of scientific works that have been carried out in the process of systematic review of the literature, following a process of identification, review, eligibility, and inclusion (Moher et al., 2009).

Interpretation of results

In the first place, the number of published articles, based on the characteristics that are worked on in this systematic review, has obtained rapid growth in the years 2019, 2020, and 2021. Another of the fields that has been analyzed has been the origin of the scientific articles, where six (6) countries of origin have been found, of which Taiwan stands out with seven (7) publications. Finally, the works analyzed are concentrated in nine (9) journals and four (4) conference proceedings, which is 69 and 31%, respectively.

Next, the results obtained around the designed research questions are detailed, considering the evaluation of the articles together with the information analysis process.

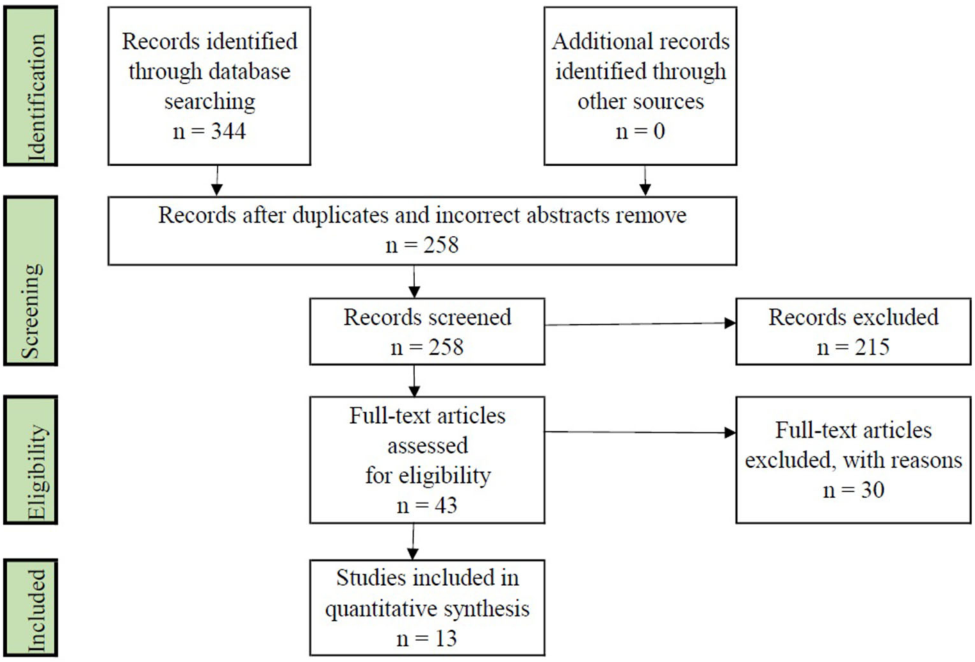


FIGURE 1
Selection process of the studies.

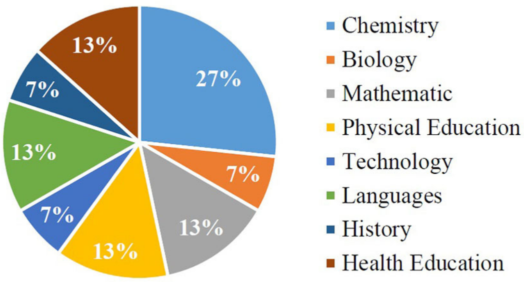


FIGURE 2
Application areas or subjects in Augmented Reality.

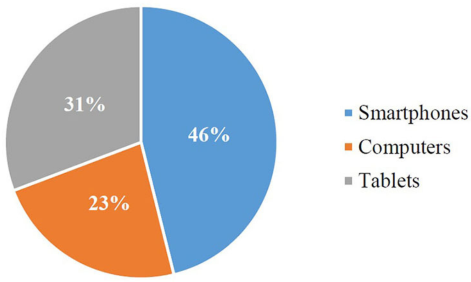


FIGURE 3
Electronic devices used.

Target areas and groups

Since the systematic review is developed in the educational stage of Secondary Education, the areas that appear are relevant in that phase. In this way, Figure 2 shows the different subjects where Augmented Reality has been implemented in the classroom.

Regarding the target groups of the activities with Augmented Reality, they are concentrated in the courses included in Secondary Education, that is, from 7 to 12th grade. In this sense, the students included in the 10 and 11th courses, whose ages range between 15 and 17 years, have accounted for 32.5% each. On the contrary, the 8 and 9th grades, whose ages vary between 13 and 15 years, have occupied 17.5% each. In short, students in the last years of Secondary Education have been exposed to investigations with Augmented Reality in 65% of the total cases investigated.

Technological devices and applications

The use of electronic devices for the implementation of educational activities based on Augmented Reality is a mandatory measure for this type of research. Based on this, the results obtained are shown in Figure 3.

As for the different technologies that have been used to create applications or software for Augmented Reality, Vuforia stands out. This development kit has been used by 54% of all developers, being its multiplatform engine Unity. The rest of the investigations have used other tools equally, which has meant 7.6% each. These applications are the following: Lighting Studios, ARDehaes toolkit, Aurasma, RAVVAR, and Metaverse Studio.

Design and forms of AR application

In reference to the design of educational activities based on Augmented Reality, it has been entirely developed by people outside the teaching staff or students who have been part of the research. In all cases, 100% of the researchers have overseen creating the content that was later applied in the different classrooms. Therefore, the role that teachers have developed has been that of guide or facilitator of the information necessary to conduct research practices. In this sense, they have had to deepen or conduct training tasks around the different Augmented Reality applications that have been implemented, in order to take them to the classroom.

On the other hand, the implementation of these contents has been conducted through different forms of application of Augmented Reality. In this sense, object modeling has been present in most of the investigations, accounting for 45.45% of the total. This was followed by AR books with 18.18% of the total, although closely followed by gamification of educational games and skills training with 13.63% of the total each. Finally, discovery-based learning techniques have been applied, accounting for 9.11% of the total.

Motivation and academic performance

Before presenting the results obtained in each of the categories, it is necessary to mention that these variables have been analyzed globally, leaving aside the subject of application, the target group, the methodology implemented, and the technological devices and applications used.

The first category represents the motivation shown by students during the process of implementation and evaluation of educational activities based on Augmented Reality. On this occasion, the different works have been analyzed to find out what levels of motivation or interest have been shown by the students who have used this educational technology compared to the students who have not used it. To do this, the Instructional Materials for Motivation (IMMS) instrument has been used to determine the indicators of attention, relevance, trust, and satisfaction on which Keller's ARCS model (1987, 2010) is based. The data obtained from this systematic review of the literature reflect that motivation levels have grown ostensibly. Specifically, 83.33% of the investigations grant an increase in the degree of interest and motivation in the students. Likewise, the works investigated have indicated through interviews or questionnaires that Augmented Reality has increased the predisposition and interest in the teaching and learning process.

The second category indicates the impact that the use of Augmented Reality has had on students in the acquisition of knowledge and therefore in their qualifications. On this occasion, it was about evaluating the knowledge acquired by the students, once this technology has been implemented in the classrooms, by means of tests or tests. In this way, the qualifications of the students have been contrasted with the purpose of obtaining an assessment of the importance of the use of this educational technology. In this sense, 77% of the works analyzed have refuted that the use of Augmented Reality has improved the grades of the students. In addition, no research work has stated that the use of this technology in the teaching and learning process has led to a drop in grades.

Discussion

In this section, the results analyzed are discussed, the research questions are answered based on the findings, and finally the conclusions drawn from this systematic review of the literature are presented.

RQ1: What subjects and groups are the recipients of educational activities with AR?

At first, it is sought that Augmented Reality has been applied only to the Secondary Education stage. This fact requires that the subjects that appear in the investigations be worked on in the curriculum of said educational stage. However, the spectrum of areas in this stage is very broad, however, the data obtained reflect that the predominant subjects are grouped in natural sciences and logical-mathematical (Chen and Liao, 2015; Lin et al., 2015; Wang et al., 2017; Chen and Chen, 2018; Hsieh and Chen, 2019; Cen et al., 2020; Tarng et al., 2021). Proof of this are the subjects of Chemistry, Biology and Mathematics, which make up more than 50% of the studies analyzed. In these investigations, the contribution of Augmented Reality is highlighted based on the visualization, understanding and acquisition of content about these subjects, which brings with it an increase in student attention during the teaching and learning process (Wang et al., 2017; Chen et al., 2021). In this sense, subjects such as Chemistry and Biology, where abstract concepts are addressed, obtain extremely low grades from students, one of the main reasons being the lack of attention produced by little or no assimilation of their contents (Chen and Liao, 2015; Chen and Chen, 2018; Tarng et al., 2021).

On the other hand, the target groups resulting from the systematic review have been the 10 and 11th grade courses, which form the ages of 15–17 years (Paredes-Velastegui et al., 2018; Cen et al., 2020; Moreno-Guerrero et al., 2020; Koç et al., 2021; Lin et al., 2021). This shows that researchers seek to work with students who have the highest ages within the stage, the main reason being the high degree of disinterest that is reflected in that age group (Amores-Valencia and De-Casas-Moreno, 2019; González-Valenzuela and Martín Ruiz, 2019).

RQ2: What technological devices have been used to generate and/or execute AR applications?

Knowing the technological devices used during the execution of educational activities with Augmented Reality is essential, since it directly affects the possibilities and inconveniences that this type of scientific work can present. In this sense, it is necessary to have an electronic device, be it a tablet, computer or mobile phone to carry out the implementation of the investigations. Therefore, the results of this systematic review reflect that almost half of the investigations have used smartphones as a technological device. The main reason is that the students were in possession of one, giving them the opportunity to develop the activities designed individually, without the need to share devices (Paredes-Velastegui et al., 2018; Hsieh and Chen, 2019; Cen et al., 2020; Moreno-Guerrero et al., 2020; Koç et al., 2021; Tarng et al., 2021). This great inconvenience has been duly notified together with the problems of time availability in the Computer Classrooms and the Internet connection (Lin et al., 2015; Wang et al., 2017; Chen and Chen, 2018).

Regarding the software or applications used for the implementation of Augmented Reality, it can be seen how the applications created through the Vuforia platform, where the developers have been able to adapt the applications to the maximum, have been the most used, occupying the half of the research papers. In this sense, the development of own applications has prevailed, where it has been particularized according to the students who were going to develop each investigation. In this way, several aspects have been considered, such as their age, language and cognitive and sensory capacity (Wei et al., 2015; Paredes-Velastegui et al., 2018; Hsieh and Chen, 2019; Cen et al., 2020). It is important to highlight that the rest of the investigations, even having used a commercial application, have prepared the educational activities based on their own students.

RQ3: How are educational activities implemented with AR in the classroom?

All the researchers have designed or adapted the content for the different scientific works. This event shows that no teacher where these educational activities with Augmented Reality have been incorporated has been part of the process or design of the applications or software used. However, teachers have been immersed in some activities design processes, corroborating the inclusion of skills and learning standards appropriate to the subject and the educational context. Therefore, the professors who work day by day with the students who have undergone this research have played the role of guide and content designer, taking into account parameters of creativity and innovation (Wei et al., 2015; Wang et al., 2017; Chen and Chen, 2018; Paredes-Velastegui et al., 2018; Chen et al., 2020; Moreno-Guerrero et al., 2020). Likewise, the students have not been involved in the design process of the activities, however, they have developed an active, participatory and collaborative attitude. In research conducted by Fernández-Robles (2017) and Gallego-Pérez (2018) in university environments, it has been confirmed that the involvement of students in the creation of content favors motivation, since they have obtained better values in attention, relevance, confidence and satisfaction present in the ARCS model of Keller (1987, 2010).

Regarding the form of application of Augmented Reality, it has been based on the five dimensions exposed by Yuen et al. (2011), where the relevance of the context of the students when implementing said technology is showed. In this matter, object modeling is considered the form of application par excellence since it has been involved in practically half of the investigations. However, Augmented Reality books, gamification or skills training cannot be ruled out. This last process is used more in very advanced and specialized educational stages, since it supplies the means to acquire skills without harming any material or human damage (Rio-Guerra et al., 2019; Gómez-Tone et al., 2020).

RQ4: What motivational impact do students have based on the use of AR?

The research analyzed showed a comparative study between experimental groups and control groups. In such a way, that the control groups developed methodologies where Augmented Reality is not implemented as an educational resource, in contrast to the experimental groups. Based on the results obtained, it can be affirmed that the students of these experimental groups have ostensibly increased their motivation or interest during the educational activities with Augmented Reality. This assertion is determined by the results obtained, where specifically 83% of the studies analyzed have determined that the levels of motivation evaluated through questionnaires based on the parameters of attention, relevance, confidence and satisfaction established in the ARCS model of Keller (1987, 2010) have grown abruptly (Chen and Liao, 2015; Lin et al., 2015, 2021; Wei et al., 2015; Chen and Chen, 2018; Paredes-Velastegui et al., 2018; Cen et al., 2020; Chen et al., 2020; Moreno-Guerrero et al., 2020; Tarng et al., 2021).

On the other hand, the data obtained in the interviews or satisfaction questionnaires conducted a posteriori in the different investigations have endorsed the previous information, since most students have confirmed that the use of educational activities with Augmented Reality has fostered their predisposition learning, capturing their attention throughout the project. Therefore, this systematic review shows that the use of Augmented Reality in the classroom generates higher levels of motivation in students, and therefore they are of immense help for teaching. These results are like those achieved by Di Serio et al. (2013), Cabero-Almenara et al. (2017), and Gallego-Pérez (2018) in other educational stages.

RQ5: How does the use of AR influence the academic performance of students?

Based on the results obtained, it can be said that students show better grades when educational activities with Augmented Reality are implemented in the classroom. This assertion is preceded by the data extracted in the investigations, where a comparison between control groups and experimental groups has been conducted. For this, different tests or tests have been conducted, before and after the implementation of this educational technology, with the purpose of contrasting the results and refuting whether the academic performance in terms of qualifications has been increased thanks to the use of Augmented Reality. During the teaching-learning process. In this sense, it has been confirmed that two out of three students have seen their grades increase very significantly (Chen and Liao, 2015; Wei et al., 2015; Wang et al., 2017; Chen and Chen, 2018; Paredes-Velastegui et al., 2018; Cen et al., 2020; Moreno-Guerrero et al., 2020; Tarng et al., 2021).

Likewise, no research has emphasized that the use of Augmented Reality in the classroom has led to a decrease in the academic grades of students, so it is a sure bet of success (Lin et al., 2015, 2021; Hsieh and Chen, 2019; Chen et al., 2020; Koç et al., 2021). These results are like those obtained by Quintero et al. (2019) where it is showed that the use of Augmented Reality improved the performance of students with visual, motor, cognitive and auditory difficulties. It is necessary to point out that other research related to the field of medicine has confirmed that the use of Augmented Reality has ostensibly improved performance in terms of navigation screen control (Cagiltay et al., 2019; Jacobsen et al., 2019).

Conclusion

This systematic review tries to value the relevance of ICT, specifically the use of Augmented Reality in the Secondary Education stage. According to (Unesco (2004), p. 30) "ICTs are a decisive tool to help students access vast resources of knowledge, collaborate with other classmates, consult experts, share knowledge, and solve complex problems using tools cognitive." The use of ICT tools such as Augmented Reality that integrate open educational resources (OER) in an organic and transversal way in face-to-face, online and blended educational contexts, is a challenge for education, including Secondary Education (Unesco, 2019).

This educational stage, so problematic due to the low or null motivation that students present, together with poor academic results, is a great challenge for teachers (Picó-Lozano, 2014; Amores-Valencia and De-Casas-Moreno, 2019; González-Valenzuela and Martín Ruiz, 2019; Jerez-Carrillo, 2021). The first of the important aspects to know are the subject's where Augmented Reality has been introduced as an educational technology. In this sense, it has been found that the studies carried out by Chen and Liao (2015), Lin et al. (2015), Wang et al. (2017), Chen and Chen (2018), Hsieh and Chen (2019), Cen et al. (2020), Tarng et al. (2021), and have been developed in the areas of natural sciences and logical-mathematical. This shows that the use of Augmented Reality is more widespread in this field, as opposed to the areas of languages, arts or social sciences. Although the results analyzed in said investigations have confirmed that there is no significant difference regardless of the educational subject where it is implanted, the use of this educational technology must be addressed in the most and least attractive fields for the students, in this way could do a comparative study, and thus know the true potential and scope (Abad-Segura et al., 2020).

Regarding the target groups where this educational technology has been used, it has been an important turning point, since the ages where the greatest signs of disinterest and motivation appear are those between 12 and 18 years old (Amores-Valencia and De-Casas-Moreno, 2019; González-Valenzuela and Martín Ruiz, 2019). For this reason, this study has been conducted entirely dedicated to Secondary Education, since older students would give results conditioned by a more mature, more concentrated and motivated behavior, which would mean better academic performance (Cabero-Almenara et al., 2019a). Similarly, Primary Education students have elevated levels of attention and motivation, so their academic results tend to have high values (Toledo-Morales and Sánchez-García, 2017; Wang, 2017; Kirikkaya and Başgül, 2019; Lai et al., 2019; López-Belmonte et al., 2019). Within the Secondary Education stage itself, almost all the articles analyzed have been developed between the ages of 15 and 17, this range being the most conflictive in terms of lack of interest, motivation and low grades (Picó-Lozano, 2014; Jerez-Carrillo, 2021). For this reason, the investigations conducted by Cen et al. (2020), Koç et al. (2021), Lin et al. (2021), Moreno-Guerrero et al. (2020), and Paredes-Velastegui et al. (2018) have opted to analyze this age range.

In reference to the use of technological devices to bring the implementation of Augmented Reality to the classroom, the smartphone has been confirmed as the tool par excellence. This fact is due to the great limitations that educational centers present due to the unavailability of computers or tablets for each student, without forgetting the complicated situation of combining the use of these devices with other activities conducted in these centers (Vedadi et al., 2017). Thus, the availability of Computer Classrooms is extremely limited for many teachers due to the great demand received. In addition, it is acceptable to highlight the Internet connection problems present in educational centers. This translates into great concern for many teachers, since it shows that it is not exactly easy to integrate this educational technology in schools (Akçayir and Akçayir, 2017).

Once the technological devices have been addressed, software applications designed for the implementation of

Augmented Reality should be looked for. In this aspect, two clearly found versions have been seen. The first refers to the applications created through the Vuforia platform, which have been duly studied and designed to fit perfectly in the students, taking into account aspects such as age, language, cognitive and sensory capacity of these (Wei et al., 2015; Paredes-Velastegui et al., 2018; Hsieh and Chen, 2019; Cen et al., 2020). About the second version, the materials created must largely adhere to the possibilities and limitations inherent in applications or software previously developed by other users. Based on this aspect, it can be affirmed that the contents developed and created by the researcher himself in full are more predisposed to success than the others (López-García et al., 2019).

The purpose of this research was to know the impact of the use of Augmented Reality in the Secondary Education stage based on motivation factors and academic performance. As for the data obtained from the different investigations addressed, they show an abrupt growth in the motivation levels of the students compared to the students who have not used this educational technology. This information has been reflected in the parameters of attention, relevance, trust and satisfaction analyzed following Keller's ARCS model (1987, 2010), which is a profound reason for the formation of teaching and learning practices that are based on the use of Augmented Reality (Chen and Liao, 2015; Lin et al., 2015, 2021; Wei et al., 2015; Wang et al., 2017; Chen and Chen, 2018; Paredes-Velastegui et al., 2018; Cen et al., 2020; Chen et al., 2020; Moreno-Guerrero et al., 2020; Tarng et al., 2021). In addition, the interviews conducted in several research works show a relationship between the use of this technology in the educational context and the increase in student motivation. This fact leads to highlight the importance of relying on the use of technologies, such as Augmented Reality, which enhance the interest of students (Keller, 2012).

Regarding the influence of this educational technology on the academic performance of the students reflected in their grades, the comparative results of the questionnaires or tests conducted show a significant difference between the students who made use of Augmented Reality in the teaching process and learning and those who tackled the activities without this technology. In this sense, the scores of the experimental groups were higher, which translates into an important reason to introduce Augmented Reality at these ages (Chen and Liao, 2015; Lin et al., 2015; Wei et al., 2015; Wang et al., 2017; Chen and Chen, 2018; Paredes-Velastegui et al., 2018; Cen et al., 2020; Chen et al., 2020; Moreno-Guerrero et al., 2020; Tarng et al., 2021). Likewise, a close relationship between motivation and academic performance is seen, since students with important levels of motivation obtain better grades (Hsieh and Chen, 2019; Koç et al., 2021; Lin et al., 2021). This information is valuable, since it is a great proposal for all teachers who want to see how the academic results of their students improve (Deigmann et al., 2015).

In relation to the limitations that have been seen, it should be noted that most of the research on Augmented Reality is not conducted in the Secondary Education stage, which has made it difficult to choose more research works. Likewise, most of these works do not jointly address the motivation and academic performance of students, and in turn, are studied from different perspectives. On the other hand, the works analyzed have not used a high number of students, due to the need to use technological devices. Regarding the possible lines of future research, it would be interesting to propose research where the teachers themselves are the designers and developers of software applications. However, this process is extremely complicated, since most teachers lack the knowledge and skills necessary to develop and apply this type of educational technology in the classroom (López-Belmonte et al., 2019). However, it can be said that students who have been immersed in the creation of Augmented Reality learning objects have shown much more satisfactory results, in terms of motivation and qualifications, than those who have only been consumers of this educational technology (Fernández-Robles, 2017; Quintero et al., 2019).

The present work aims to expand the current state of research in the field of Augmented Reality in the Secondary Education stage, grouping not only curricular aspects such as subjects, activities, method, but also two other major factors such as motivation and academic performance, with the purpose of capturing the repercussion of this educational technology for future studies.

Data availability statement

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

Author contributions

AA-V, DB, and JB-B: conceptualization, design of the study, and writing—review and editing. AA-V: formal analysis, methodology, and writing—original draft. All authors approved the definitive version of the manuscript to be published.

Acknowledgments

The authors would like to thank Research Institute for Innovation and Technology in Education (UNIR iTED), Universidad Internacional de La Rioja, Logroño, Spain, and Research Group GIDIA, Faculty of Mining, Universidad Nacional de Colombia, Medellín Campus, Colombia.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1011409/full#supplementary-material

References

Abad-Segura, E., González-Zamar, M. D., Luque-De la Rosa, A., and Morales-Cevallos, M. B. (2020). Sustainability of educational technologies: an approach to augmented reality research. *Sustainability* 12, 4091. doi: 10.3390/su12104091

Aguilar-Acevedo, F., Flores-Cruz, J. A., Hernández-Aguilar, C. A., and Pacheco-Bautista, D. (2022). Diseño e implementación de un simulador basado en realidad aumentada móvil para la enseñanza de la física en la educación superior. *Edutec. Revista Electrónica De Tecnología Educativa*. 80, 66–83. doi: 10.21556/edutec.2022.80.2509

Akçayir, M., and Akçayir, G. (2017). Advantages and challenges associated with augmented reality for education: a systematic review of the literature. *Educ. Res. Rev.* 20, 1–11. doi: 10.1016/j.edurev.2016.11.002

Akçayir, M., Akçayir, G., Pektas, H., and Ocak, M. (2016). Augmented reality in science laboratories: the effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Comput. Hum. Behav.* 57, 334–342. doi: 10.1016/j.chb.2015.12.054

Alonso-Tapia, J., and Ruiz-Díaz, M. (2007). Learning related motives and the perception of the motivational quality of the learning environment. *Psicothema* 19, 603–609.

Álvarez-Marín, A., Velázquez-Iturbide, J. A., and Castillo-Vergara, M. (2020). "Intention to use an interactive AR app for engineering education," in *Presentation in Conference. IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. New York, NY. doi: 10.1109/ISMAR-Adjunct51615.2020.00033

Amores-Valencia, A. J. (2020). Metodologías activas y TIC. Propuestas didácticas para Educación Secundaria Obligatoria. *Campus educación Revista Digital Docente* 1, 19–24. Available at: https://www.campuseducacioc.com/revista-digital-docente/numeros/17

Amores-Valencia, A. J., and De-Casas-Moreno, P. (2019). El uso de las TIC como herramienta de motivación para alumnos de enseñanza secundaria obligatoria. *Estudio de caso Español. Hamut'ay*. 6, 37–49. doi: 10.21503/hamu.v6i3.1845

Area-Moreira, M., Cepeda-Romero, O., and Feliciano-García, L. (2018). Perspectivas de los alumnos de Educación Primaria y Secundaria sobre el uso escolar de las TIC. *Revista Educatio Siglo XXI*. 36, 229–253. doi: 10.6018/j/333071

Azuma, R. T. (1997). A survey of augmented reality. *Presence* 6, 355–385. doi: 10.1162/pres.1997.6.4.355

Barca-Lozano, A., Almeida, L. S., Porto-Rioboo, A. M., Peralbo, M., and Brenlla-Blanco,. J. C. (2012). Motivación escolar y rendimiento: impacto de las metas académicas, de estrategias de aprendizaje y autoeficacia. *Anales de Psicologia* 28, 848–859. doi: 10.6018/analesps.28.3.156101

Barroso-Osuna, J., and Cabero-Almenara, J. (2013). "La escuela en la sociedad de la información," in *Nuevos escenarios digitales. Las tecnologías de la información y la comunicación aplicadas a la formación y desarrollo curricular*, ed En J. Barroso y J. Cabero (Madrid, España: Ediciones Pirámide), 21–36.

Barroso-Osuna, J., Cabero-Almenara, J. J., and Gutíerrez-Castillo, J. J. (2018). La producción de objetos de aprendizaje en realidad aumentada por estudiantes universitarios: grado de aceptación de esta tecnología y motivación para su uso. *Revista Mexicana de Investigación Educativa* 23, 1261–1283. doi: 10.51302/tce.2018.221

Boekaerts, M. (1996). Self-regulated leaming at the junction of cognition and motivation. *Eur. Psycholog.* 1, 100–112. doi: 10.1027/1016-9040.1.2.100

Borkowski, J. G., and Thorpe, P. K. (1994): "Self-regulation and motivation: a life-span perspective on underachievement," in *Self-regulation of Learning and Performance. Issues and Educational Applications*, eds D. R. Schunk and B. J. Zimmerman (Hillsdale, NJ: LEA), 45–74.

Bower, M., Howe, C., McCredie, N., Robinson, A., and Grover, D. (2014). Augmented Reality in education-cases, places and potentials. *Educ. Media Int.* 51, 1–15. doi: 10.1080/09523987.2014.889400

Buchner, J. (2021). Generative learning strategies do not diminish primary students' attitudes towards augmented reality. *Educ. Informat. Technol.* 26, 1–17. doi: 10.1007/s10639-021-10445-y

Bursztyn, N., Walker, A., Shelton, B., and Pederson, J. (2017). Assessment of student learning using augmented reality Grand Canyon field trips for mobile smart devices. *Geosphere* 13, 260–268. doi: 10.1130/GES01404.1

Bursztyn, N., Walker, A., Shelton, B., and Pederson, J. (2020). Augmented reality instructional tool in enhancing geography learners academic performance and retention in Osun state Nigeria. *Educ. Inform. Technolog.* 25, 3021–3033. doi: 10.1007/s10639-020-10099-2

Cabero-Almenara, J. (2015). Reflexiones educativas sobre las tecnologías de la información y la comunicación (TIC). *Revista Tecnología, Ciencia y Educación* 1, 19–27. doi: 10.51302/tce.2015.27

Cabero-Almenara, J., and Barroso-Osuna, J. (2016a). Ecosistema de aprendizaje con realidad aumentada: posibilidades educativas. *CEF* 5, 141–154. doi: 10.51302/tce.2016.101

Cabero-Almenara, J., and Barroso-Osuna, J. (2016b). The educational possibilities of augmented reality. *J. N. Approach. Educat. Res.* 5, 44–50. doi: 10.7821/naer.2016.1.140

Cabero-Almenara, J., Barroso-Osuna, J., and Llorente-Cejudo, C. (2019b). Augmented reality in university education. *Revista de Docencia Universitaria*. 17, 105–118. doi: 10.4995/redu.2019.11256

Cabero-Almenara, J., Barroso-Osuna, J., Llorente-Cejudo, C., and Fernández-Martínez, M. (2019a). Educational uses of augmented reality (AR): experiences in educational science. *Sustainability* 11, 4990. doi: 10.3390/su11184990

Cabero-Almenara, J., Fernández-Robles, B., and Marín-Díaz, V. (2017). Dispositivos móviles y realidad aumentada en el aprendizaje del alumnado universitario. *Revista Iberoamericana de Educación a Distancia* 20, 167–185. doi: 10.5944/ried.20.2.17245

Cabero-Almenara, J., and Marín-Díaz, V. (2018). Blended learning y realidad aumentada: experiencias de diseño docente. *Revista Iberoamericana de Educación a Distancia* 21, 57–74. doi: 10.5944/ried.21.1.18719

Cagiltay, N. E., Ozcelik, E., Berker, M., and Menekse-Dalveren, G. G. (2019). The underlying reasons of the navigation control effect on performance in a virtual reality endoscopic surgery training simulator. *Int. J. Hum. Comput. Interaction* 35, 1396-1403. doi: 10.1080/10447318.2018.1533151

Carlton, B. (2017). *Virtual Reality & Learning. A Masie Report*. Eindhoven, Netherlands: Elliot Masie.

Carrizo, D., and Moller, C. M. (2018). Estructuras metodológicas de revisiones sistemáticas de literatura en Ingeniería de Software: un estudio de mapeo sistemático. *Ingeniare. Revista chilena de ingeniería*. 26, 45–54. doi: 10.4067/S0718-33052018000500045

Castro-Marcos, O. (2022). *Minecraft educativo: una propuesta de gamificación en educación primaria a través de videojuegos y realidad aumentada. [Trabajo Fin de Grado, Universidad de Oviedo]*. Oviedo, España: Repositorio Institucional RUO.

Cen, L., Ruta, D., Al Qassem, L. M. M. S., and Ng, J. (2020). Augmented immersive reality (AIR) for improved learning performance: a quantitative evaluation. *IEEE Trans. Learn. Technol.* 13, 283–296. doi: 10.1109/TLT.2019.2937525

Chen, G. J., and Chen, R. S. (2018). "A case study of the development of using augmented reality in teaching nature and life technology to a junior high school students in Southern Taiwan," in *Presentation in Conference. Proceedings - 1st International Cognitive Cities Conference*. Okinawa. doi: 10.1109/IC3.2018.00-41

Chen, J. J., Hsu, Y., Wei,. W., and Yang, C. (2021). Continuance intention of augmented reality textbooks in basic design course. *Educ. Sci.* 11, 1–16. doi: 10.3390/educsci11050208

Chen, M. P., and Liao, B. C. (2015). "Augmented reality laboratory for high school electrochemistry course," in *Presentation in Conference. Proceedings - IEEE 15th International Conference on Advanced Learning Technologies: Advanced Technologies for Supporting Open Access to Formal and Informal Learning, ICALT*. Taiwan. doi: 10.1109/ICALT.2015.105

Chen, M. P., Wang, L. C., Zou, D., Lin, S. Y., Xie, H., and Tsai, C. C. (2020). Effects of captions and English proficiency on learning effectiveness, motivation and attitude in augmented-reality-enhanced theme-based contextualized EFL learning. *Comput. Assist. Lang. Learn*. 2019, 1704787. doi: 10.1080/09588221.2019.1704787

Coimbra, T., Cardoso, T., and Mateus, A. (2015). Augmented reality: an enhancer for higher education students in math's learning? 6th international conference on software development and technologies for enhancing accessibility and fighting infoexclusion (DSAI 2015). *Procedia Comput. Sci.* 67, 332–339. doi: 10.1016/j.procs.2015.09.277

Collazos, C.A., Fardoun, H., AlSekait, D., Pereira, C.S., and Moreira, F. (2021). Designing online platforms supporting emotions and awareness. *Electronics* 10, 1–19. doi: 10.3390/electronics10030251

Deigmann, P., Schmidt, M., Eynden, S. V., and Basten, D. (2015). "Benefits of augmented reality in educational environments – a systematic literature review," in *12th International Conference on Wirtschaftsinformatik* (Osnabrück, Germany), 1542–1556.

Del-Rio-Guerra, M. S., Martin-Gutiérrez, J., Lopez-Chao, V. A., Flores-Parra, R., and Ramírez-Sosa, M. A. (2019). AR graphic representation of musical notes for self-learning on guitar. *Appl. Sci.* 9, 4527. doi: 10.3390/app9214527

De-Paiva-Guimarães, M., and Farinazzo-Martins, V. (2014). Desafios a serem superados para o uso de Realidade Virtual e Aumentada no cotidiano do ensino. *Revista de Informática Aplicada*. 9, 1–10. doi: 10.13037/ria.vol9n1.2744

Di Serio, A., Ibáñez, M. B., and Delgado-Kloos, C. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Comput. Educ*. 68, 586–596. doi: 10.1016/j.compedu.2012.03.002

Dirin, A., Alamäki, A., and Suomala, J. (2019). Gender differences in perceptions of conventional video, virtual reality and augmented reality. *Int. J. Interact. Mobile Technolog*. 13, 93–103. doi: 10.3991/ijim.v13i06.10487

Duh, H. B. L., and Klopher, E. (2013). Augmented reality learning: new learning paradigm in co-space. *Comput. Educ.* 68, 534–535. doi: 10.1016/j.compedu.2013.07.030

Fernández-Robles, B. (2017). *Aplicación del Modelo de Aceptación Tecnológica (TAM) al uso de la realidad aumentada en estudios universitarios*. (Tesis doctoral), Universidad de Córdoba, Repositorio UCOPress, Córdoba, Spain.

Fonseca-Escudero, D., Redondo-Domínguez, E., and Valls, F. (2016). Motivación y mejora académica utilizando realidad aumentada para el estudio de modelos tridimensionales arquitectónicos. *Educ. Knowl. Soc.* 17, 45–64. doi: 10.14201/eks20161 714564

Gallego-Pérez, O. M. (2018). *Estudio y Análisis sobre las posibilidades educativas de la Realidad Aumentada como herramienta de producción de experiencias formativas por parte del alumnado universitario*. (Tesis doctoral), Universidad de Córdoba, Repositorio digital UCO, Córdoba, Spain.

García, T. (1995). "The role of motivational strategies in self-regulated leaming," in *Understanding Self-Regulated Learning*, ed P. R. Pintrich (San Francisco, CA: Jossey-Bass Publishers), 29–42. doi: 10.1002/tl.37219956306

García, T., and Pintrich, P. R. (1994). "Regulating motivation and cognition in the classroom: the role of selfschemas and self-regulatory strategies," in *Self-Regulation Oflearning and Performance. Issues and Educational Applications*, eds D. H. Schunk and B. J. Zirnmerman (Hillsdale, NJ: LEA), 127–154.

Garrido-Macías, M., Jiménez-Luque, N., Landa-Sánchez, A., Páez-Espinar, E., and Ruíz-Barranco, M. (2013). Factores que influyen en el rendimiento académico: la motivación como papel mediador en las estrategias de aprendizaje y clima escolar. *Revista electrónica de investigación Docencia Creativa* 2, 17–25. doi: 10.30827/Digibug.27620

Gómez-García, G., Rodríguez-Jiménez, C., and Marín-Marín, J. A. (2020). La trascendencia de la Realidad Aumentada en la motivación estudiantil. *Una revisión sistemática y meta-análisis. Alteridad* 15, 36–46. doi: 10.17163/alt.v15n1.2020.03

Gómez-Tone, H. C., Martin-Gutiérrez, J., Valencia-Anci, L., and Mora-Luis, C. E. (2020). International comparative pilot study of spatial skill development in engineering students through autonomous augmented reality-based training. *Symmetry* 12, 1401. doi: 10.3390/sym12091401

González-Pienda, J. A. (2003). El rendimiento escolar. un análisis de las variables que lo condicionan. *Revista Galego-Portuguesa de Psicoloxia e Educación* 7, 247–258. Available online at: https://www.psicothema.com/pdf/97.pdf

González-Pienda, J. A., Núñez-Pérez, J. C., González-Pumariega, S., and García-García, M. S. (1997). Self-concept, self-esteem and schoolleaming. *Psicothema* 9,271–289.

González-Segredo, L., and Hernández-Cabrera, J. (2022). *La Realidad Aumentada en las aulas de Educación Infantil. [Trabajo Fin de Grado, Universidad de La Laguna]. Repositorio ULL*. San Cristóbal de La Laguna.

González-Valenzuela, M. J., and Martín Ruiz, I. (2019). Rendimiento académico, lenguaje escrito y motivación en adolescentes españoles. *Universitas Psychologica* 18, 1–13. doi: 10.11144/Javeriana.upsy18-4.rale

Hanafi, H. F., Said, C. S., Wahab, M. H., and Samsuddin, K. (2017). Improving students' motivation in learning ICT course with the use of a mobile augmented reality learning environment. *IOP Conf. Ser.* 226, 1–10. doi: 10.1088/1757-899X/226/1/012114

Helmke, A., and van Aken, M. A. G. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: a longitudinal study. *J. Educ. Psychol*. 87, 624–637. doi: 10.1037/0022-0663.87.4.624

Hernandez, R. M. (2017). Impacto de las TIC en la educación: Retos y Perspectivas. *Propósitos y Representaciones* 5, 325–347. doi: 10.20511/pyr2017.v5n1.149

Hohlfeld, T., Ritzhaupt, A., and Barron, A. (2013). Are gender differences in perceived and demonstrated technology literacy significant? It depends on the model. *Educ. Technol. Res. Dev*. 61, 639–663. doi: 10.1007/s11423-013-9304-7

Hsieh, M. C., and Chen, S. H. (2019). Intelligence augmented reality tutoring system for mathematics teaching and learning. *J. Internet Technol.* 20, 1673–1681. doi: 10.3966/160792642019092005031

Hsu, T. C. (2019). Effects of gender and different augmented reality learning systems on English vocabulary learning of elementary school students. *Universal Access Inform. Soc.* 18, 315–325. doi: 10.1007/s10209-017-0593-1

Huang, T. C., Chen, C. C., and Chou, Y. W. (2016). Animating eco-education: to see, feel, and discover in an augmented reality-based experiential learning environment. *Comput. Educ.* 96,72–82. doi: 10.1016/j.compedu.2016.02.008

Huang, T. L., and Liao, S. L. (2017). Creating e-shopping multisensory flow experience through augmented-reality interactive technology. *Internet Res.* 27, 449–475. doi: 10.1108/IntR-11-2015-0321

Ibáñez, M. B., Di Serio, Á., Villarán, D., and Delgado-Kloos, C. (2014). Experimenting with electromagnetism using augmented reality: impact on flow student experience and educational effectiveness. *Comput. Educ.* 71,1–13. doi: 10.1016/j.compedu.2013.09.004

Jacobsen, M. F., Konge, L., Bach-Holm, D., la Cour, M., Holm, L., Højgaard-Olsen, K., et al. (2019). Correlation of virtual reality performance with real-life cataract surgery performance. *J. Cataract Refract. Surg.* 45, 1246–1251. doi: 10.1016/j.jcrs.2019.04.007

Jerábek, T., Rambousek, V., and Wildová, R. (2014). Specifics of visual perception of the augmented reality in the context of education. *Procedia Soc. Behav. Sci.* 159, 598–604. doi: 10.1016/j.sbspro.2014.12.432

Jerez-Carrillo, M. S. (2021). *Motivación y Rendimiento Académico en el alumnado de Educación Secundaria*. (Trabajo Fin de Máster), Universidad de Sevilla, Depósito de Investigación de la Universidad de Sevilla (idUS), Sevilla, Spain.

Juan, M. C., Charco, J. L., García-García, I., and Mollá, R. (2019). An augmented reality app to learn to interpret the nutritional information on labels of real packaged foods. *Front. Comput. Sci.* 1, 1–16. doi: 10.3389/fcomp.2019.00001

Keller, J. M. (1987). "Motivational design of instructional design: a theorical perspective," in *Instructional-Design Theories and Models: An Overview of Their Current Status*, ed C. M. Reigeluth (Mahwah, NJ: Lawrence Earlbaum Associates), 384–434.

Keller, J. M. (2010). *Motivational Design for Learning and Performance: The ARCS Model Approach*. Berlin: Springer. doi: 10.1007/978-1-4419-1250-3

Keller, J. M. (2012). "ARCS model of motivation," in *Encyclopedia of the Sciences of Learning*, ed N. M. Seel (Berlin: Springer), 304–305. doi: 10.1007/978-1-4419-1428-6_217

Kirikkaya, E. B., and Başgül, M. S. (2019). The effect of the use of augmented reality applications on the academic success and motivation of 7th grade students. *J. Baltic Sci. Educ.* 18, 362–378. doi: 10.33225/jbse/19.18.362

Kitchenham, B., and Charters, S. (2007). "Guidelines for performing systematic literature reviews in software engiieering," in *Presentation in Conference. EBSE Keele University and Durham University Join Report 001* (Keele, Staffordshire, UK).

Kitchenham, B. A., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Inform. Softw. Technol.* 51, 7–15. doi: 10.1016/j.infsof.2008.09.009

Koç, Ö., Altun, E., and Yüksel, H. G. (2021). Writing an expository text using augmented reality: students' performance and perceptions. *Educ. Inform. Technolog.* 2021, 1–22. doi: 10.1007/s10639-021-10438-x

Lai, A. F., Chen, C. H., and Lee, G. Y. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *Br. J. Educat. Technol.* 50, 232–247. doi: 10.1111/bjet.12716

Lin, H. C. K., Chen, M. C., and Chang, C. K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interact. Learn. Environ.* 23, 799–810. doi: 10.1080/10494820.2013.817435

Lin, H. C. K., Lin, Y. H., Wang, T. H., Su, L. K., and Huang, Y. M. (2021). Effects of incorporating augmented reality into a board game for high school students' learning motivation and acceptance in health education. *Sustainability* 13, 3333. doi: 10.3390/su13063333

Lin, Y. H., Koong, H. C., and Lio, I. C. (2017). "The study on the application in the combination of pervasive gaming and augmented reality in the temple tour for users with different cognitive styles," in *ICCE - 25th International Conference on Computers in Education: Technology and Innovation: Computer-Based Educational Systems for the 21st Century, Workshop Proceedings*, Taiwan.

Liu, Y. C., Lu, S. J., Kao, C. Y., Chung, L., and Tan, K. H. (2019). Comparison of AR and physical experiential learning environment in supporting product innovation. *Int. J. Eng. Bus. Manag.* 11, 1–10. doi: 10.1177/1847979019839578

López-Belmonte, J., Pozo-Sánchez, S., and Fuentes-Cabrera, A. (2019). Techno-pedagogical resources to support teaching: augmented reality as a dynamic tool for the substitute teacher. *Int. J. Educat. Res. Innov.* 12, 122–136. doi: 10.46661/ijeri.3653

López-García, A., Miralles-Martínez, P., and Maquilón, J. (2019). Design, application and effectiveness of an innovative augmented reality teaching proposal through 3P model. *Appl. Sci.* 9, 5426. doi: 10.3390/app9245426

Macías-González, L., and Manresa-Yee, C. (2013). Mayores y nuevas tecnologías: motivaciones y dificultades. *Ariadna* 1, 6–11. doi: 10.6035/Ariadna.2013.1.2

Marín-Díaz, V. (2016). Posibilidades de uso de la realidad aumentada en la educación inclusiva. *Estudio de caso. Ensayos, Revista de la Facultad de Educación de Albacete*. 31, 57–68. doi: 10.21071/edmetic.v6i1.5804

Marín-Díaz, V., Cabero-Almenara, J., and Gallego-Pérez, O. M. (2018). Motivación y realidad aumentada: Alumnos como consumidores y productores de objetos de aprendizaje. *Motivat. Augment. Reality* 47, 337. doi: 10.17811/rifie.47.3.2018.337-346

Marsh, H. W., and Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: structural equation models of longitudinal data. *J. Educat. Psychol.* 89, 41–54. doi: 10.1037/0022-0663.89.1.41

Martín, E., Martínez-Arias, R., Marchesi, A., and Pérez, E. M. (2008). Variables that predict academic achievement in the Spanish Compulsory Secondary Educational System: a longitudinal, multilevel analysis. *Spanish J. Psychol.* 11, 400–413. doi: 10.1017/S113874160000442X

Martín-Gutiérrez, J., and Meneses-Fernández, M. D. (2014). Applying augmented reality in engineering education to improve academic performance and student motivation. *Int. J. Eng. Educ.* 30, 625–635. Available online at: https://dialnet.unirioja.es/servlet/articulo?codigo=7372823

Mascarenhas, S., Almeida, L. S., and Barca, A. (2005). Atribuições causais e rendimento escolar: impacto das habilitaçoes escolares dos pais e do género dos alunos. *Revista Portuguesa de Educação*. 18, 77–91. Available online at: https://www.redalyc.org/articulo.oa?id=37418105

McLean, G., and Wilson, A. (2019). Shopping in the digital world: examining customer engagement through augmented reality mobile applications. *Comput. Hum. Behav.* 101, 210–224. doi: 10.1016/j.chb.2019.07.002

Mestre-Navas, J. M., Pérez, N., González-De-La-Torre, G., Núñez-Lozano, J. M., and Guil-Bozal, R. (2017). El desarrollo de la inteligencia emocional a través de la optimización de las capacidades cognitivas a través de la educación obligatoria. *Contextos educativos* 20, 57–75. doi: 10.18172/con.3023

Miñano-Pérez, P., and Castejón-Costa, J. L. (2011). Variables cognitivas y motivacionales en el rendimiento académico en Lengua y Matemáticas. *Revista de Psicodidáctica* 16, 203–230. Available online at: https://ojs.ehu.eus/index.php/psicodidactica/article/download/930/1585

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6, 97. doi: 10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systemat. Rev.* 4, 1–9. doi: 10.1186/2046-4053-4-1

Moreno-Guerrero, A. J., Alonso-García, S., Navas-Parejo, M. R., Campos-Soto, M. N., and Gómez-García, G. (2020). Augmented reality as a resource for improving learning in the physical education classroom. *Int. J. Environ. Res. Public Health* 17, 3637. doi: 10.3390/ijerph17103637

Núñez, J. C. (2009). "Motivación, aprendizaje y rendimiento académico," in *Acta do X Congreso Internacional Galego-Portugués de Psicopedagogía*. Braga: Universidad do Minho.

Paredes-Velastegui, D., Lluma-Noboa, A., Olmedo-Vizueta, D., Avila-Pesantez, D., and Hernandez-Ambato, J. (2018). "Augmented reality implementation as reinforcement tool for public textbooks education in Ecuador," in *IEEE Global Engineering Education Conference, EDUCON*. Santa Cruz de Tenerife. doi: 10.1109/EDUCON.2018.8363372

Park, C., 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

Picó-Lozano, M. (2014). *La importancia de la motivación en el rendimiento académico de los estudiantes de Educación Secundaria Obligatoria*. (Trabajo Fin de Máster), Universidad de las Islas Baleares, Repositorio Institucional de la UIB, Palma, Spain.

Pintrich, P. R. (1994). Continuities and discontinuities: future directions for research in educational psychology. *Educat. Psychologist* 29, 137–148. doi: 10.1207/s15326985ep2903_3

Pintrich, P. R., Marx, R. W., and Boyle, R. A. (1993). Beyond cold conceptual change: the role ofmotivational beliefs and classroom contextual factors in the process of conceptual change. *Rev. Educat. Res.* 63, 167–199. doi: 10.3102/00346543063002167

Pressley, M., Harris, K. R., and Guthrie, J. T. (1992). *Promoting Academic Competence and Literacy in School*. Cambridge, MA: Academic Press, Inc.

Quintero, J., Baldiris, S., Rubira, R., Cerón, J., and Velez, G. (2019). Augmented reality in educational inclusion. A systematic review on the last decade. *Front. Psychol.* 10, 1835. doi: 10.3389/fpsyg.2019.01835

Redondo-Domínguez, E., Fonseca-Escudero, D., Sánchez-Riera, A., and Navarro-Delgado, I. (2014). Mobile learning en el ámbito de la arquitectura y la edificación. *Análisis de casos de estudio. Revista de Universidad y Sociedad del Conocimiento* 11, 152–174. doi: 10.7238/rusc.v11i1.1844

Reinoso-Peinado, R. (2016). "Realidad aumentada posibilidades y usos en educación," in *Recursos educativos aumentados. Una oportunidad para la inclusión*, eds S. M. Baldiris, N. Darío, D. J. Salas, J. C. Bernal, R. Fabregat, R. Mendoza, et al. (Cartagena: Sello editorial Tecnológico Comfenalco), 8–25.

Rezaee, S., Sadeghi, A., Shakeri, M., and Choi, S. M. (2021). Personalized augmented reality based tourism system: big data and user demographic contexts. *Appl. Sci.* 11, 36047. doi: 10.3390/app11136047

Rio-Guerra, M. S., Martin-Gutiérrez, J., Lopez-Chao, V. A., Flores-Parra, R., and Ramírez-Sosa, M. A. (2019). AR graphic representation of musical notes for self-learning on guitar. *Appl. Sci.* 9, 4527.

Sánchez-Bolado, J. (2017). El potencial de la realidad aumentada en la enseñanza de español como lengua extranjera. *Edmetic* 6, 68–80. doi: 10.21071/edmetic.v6i1.5808

Schaffernak, H., Moesl, B., Vorraber, W., and Koglbauer, I. V. (2020). Potential augmented reality application areas for pilot education: an exploratory study. *Educ. Sci.* 10, 40086. doi: 10.3390/educsci10040086

Schunk, D. H., and Zimmennan, B. J. (1994). *Self-Regulation Oflearning and Performance. Issues and Educational Applications.* Mahwah, NJ: LEA.

Solak, E., and Cakir, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *J. Educat. Onl.* 13, 50–72. doi: 10.9743/JEO.2015.2.5

Sommerauer, P., and Müller, O. (2014). Augmented reality in informal learning environments: a field experiment in a mathematics exhibition. *Comput. Educ.* 79, 59–68. doi: 10.1016/j.compedu.2014.07.013

Tarng, W., Lin, Y., and Ou, K. L. (2021). A virtual experiment for learning the principle of daniell cell based on augmented reality. *Appl. Sci.* 11, 762. doi: 10.3390/app11020762

Toledo-Morales, P., and Sánchez-García, J. M. (2017). Realidad Aumentada en Educación Primaria: efectos sobre el aprendizaje / Augmented Reality in Primary Education: effects on learning. *Revista Latinoamericana de Tecnología Educativa* 16, 79–92. doi: 10.17398/1695-288X.16.1.79

Unesco (2004). *Las Tecnologías de la Información y la Comunicación en la formación docente: guía de planificación.* Paris: División de Educación Superior – Unesco.

Unesco (2019). *Recommendation on Open Educational Resources (OER).* Paris: General Conference, 40th session.

Urrútia, G., and Bonfill, X. (2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clinica* 135, 507–511. doi: 10.1016/j.medcli.2010.01.015

Usán-Supervía, P., and Salavera-Bordás, C. (2018). Motivación escolar, inteligencia emocional y rendimiento académico en estudiantes de educación secundaria obligatoria. *Actualidades en Psicología* 32, 95–112. doi: 10.15517/ap.v32i125.32123

Vedadi, S., Abdullah, Z. B., Arsovski, S., Cheok, A. D., and Aris, B. B. (2017). The impact of augmented reality on vocabulary acquisition in Iran. *Adv. Sci. Lett.* 23, 9098–9100. doi: 10.1166/asl.2017.10031

Wang, S. M., Chen, K. T., Hou, H. T., and Li, C. T. (2017). "A science history educational board game with augmented reality integrating collaborative problem solving and scaffolding strategies," in *ICCE - 25th International Conference on Computers in Education: Technology and Innovation: Computer-Based Educational Systems for the 21st Century, Workshop Proceedings* (Christchurch, New Zealand).

Wang, Y. H. (2017). Using augmented reality to support a software editing course for college students. *J. Comput. Assist. Learn.* 33, 532–546. doi: 10.1111/jcal.12199

Wei, X., Weng, D., Liu, Y., and Wang, Y. (2015). Teaching based on augmented reality for a technical creative design course. *Comput. Educ.* 81, 221–234. doi: 10.1016/j.compedu.2014.10.017

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychol. Rev.* 92, 548–573. doi: 10.1037/0033-295X.92.4.548

Weiner, B. (1986). *An Attributional Theory of Motivation and Emotion*. Berlin: Springer Verlag. doi: 10.1007/978-1-4612-4948-1

Wojciechowski, R., and Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Comput. Educ.* 68, 570–585. doi: 10.1016/j.compedu.2013.02.014

Yuen, S., Yaoyuneyong, G., and Johnson, E. (2011). Augmented reality: an overview and five directions for AR in education. *J. Educat. Technol. Dev. Exchange* 4, 119–140. doi: 10.18785/jetde.0401.10

Yuen, S. C. Y., Yaoyuneyong, G., and Johnson, E. (2013). *Augmented Reality and Education: Applications and Potentials*. Berlin: Springer Heidelberg. doi: 10.1007/978-3-642-32301-0_17

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 18 October 2022
DOI 10.3389/fpsyg.2022.1032680

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University,
China

REVIEWED BY
Xiaoxian Liu,
Henan Normal University,
China
Xin Zhao,
The University of Sheffield, United Kingdom
Zhu Ke,
Henan Normal University,
China

*CORRESPONDENCE
Zheng Zheng
z2015104@shufe-zj.edu.cn

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

RECEIVED 31 August 2022
ACCEPTED 27 September 2022
PUBLISHED 18 October 2022

CITATION
Xia Q, Ke L and Zheng Z (2022) Is video creation more effective than self-exercise in motor skill learning?
Front. Psychol. 13:1032680.
doi: 10.3389/fpsyg.2022.1032680

Is video creation more effective than self-exercise in motor skill learning?

Qiudong Xia[1], Lu'an Ke[1] and Zheng Zheng[2,3]*

[1]Department of Physical Education, Zhejiang Gongshang University, Hangzhou, China, [2]College of Business, Shanghai University of Finance and Economics, Shanghai, China, [3]Physical Education Teaching and Research Office of Public Foundation Department, Zhejiang University of Finance and Economics, Hangzhou, Zhejiang, China

Parallel to the tremendous growth and expansion of video technology, it is easy and enjoyable for students to create a video as a learning activity. However, most previous studies primarily focused on declarative knowledge learning (e.g., language learning, science learning) rather than motor skill learning. The current study aimed to investigate whether creating and sharing a video with classmates would be more effective than merely creating a video and self-exercise to learn a motor skill in terms of intrinsic motivation, perseverance in learning, learning satisfaction, and roller-skating skill. Partially consistent with our hypothesis, we found that creating and sharing a video with classmates increased students' intrinsic motivation, perseverance in motor tasks, and learning satisfaction, but not roller-skating skill, followed by merely creating a video and then self-exercise. The findings have an important implication for motor skills learning: during teaching motor skills, teachers can use encourage students to create and share a video with classmates as a homework activity to increase students' intrinsic motivation, perseverance in motor tasks, and learning satisfaction.

KEYWORDS
video creation, self-exercise, motor skill learning, intrinsic motivation, video learning

Introduction

Technology has developed significantly, and its impacts are inevitable on education in multiple disciplines (e.g., language learning, science learning, motor skills; Aguilar and Pifarre Turmo, 2019; Chang et al., 2020; Zainuddin and Che' Lah, 2022). As the nature of dynamic and multichannel presenting learning content, learning from a video is immensely popular and an essential ingredient of many contemporary instructional approaches, such as massive open online courses, flipped classrooms, and even traditional classrooms (Noetel et al., 2021; Rekik et al., 2021). However, previous studies primarily consider videos to deliver learning content to students (Pi et al., 2021, 2022).

Parallel to the tremendous growth and expansion of video technology, it is easy and enjoyable for students to create a video as a learning activity (Greene and Crespi, 2012; Hoogerheide et al., 2019). Some emerging studies are interested in the effects of creating videos by students (Sari et al., 2020; Zainuddin and Che' Lah, 2022). Creating a video offers students a chance to exercise and self-assessment. During creating a video, students can develop their interpretation, capture, review, and evaluate their performance on their own or by seeking their peers' views (Huang, 2015; Zainuddin and Che' Lah, 2022).

Recent research on learning strategies has shown that creating a video can be an effective strategy for learning (Hoogerheide et al., 2019; Erlangga, 2021; Zainuddin and Che' Lah, 2022). For example, the study by Zainuddin and Che' Lah (2022) tested the effects of video creation in English classrooms. They found that video creation increased students' confidence and eventually improved their speaking skills. Similar results were obtained by Hoogerheide et al. (2019) with different learning content (i.e., biology). Again, students creating a video consistently experienced more enjoyment, which may stimulate engagement factors such as perseverance in learning, and showed higher learning satisfaction and better learning performance compared to restudy (Garland, 1982). Perseverance in learning refers to persistence at a learning task, which is evidenced to can be promoted by task enjoyment (Leonard and Weitz, 1971). The benefits of video creation have since been replicated in other studies (Huang, 2015; Sari et al., 2020; Erlangga, 2021; Jung, 2021).

However, most previous studies primarily focused on declarative knowledge learning (e.g., language learning, science learning) rather than motor skill learning. There are some differences between declarative knowledge and motor skill learning (Anderson, 1982). Specifically, declarative knowledge is 'know-what' knowledge (e.g., facts, concepts, and theory); motor skill is 'know-how' knowledge (e.g., roller-skating, playing ping-pong, and swimming). Therefore, motor skill learning includes learning declarative knowledge and mastering how to enact the skill *via* exercise (van Abswoude et al., 2019). Physical educators struggle with the challenges of creating opportunities for all students to engage in exercise and self-assessment daily, such as roller-skating (Finkenberg et al., 2005). Creating a video as a learning activity can address these challenges by encouraging students to show motor skills, develop their interpretation, capture, review, and evaluate their performance on their own (Huang, 2015; Zainuddin and Che' Lah, 2022).

Researchers have proposed several potential mechanisms for the effects of creating video. First, the exercise effect postulates that exercise has beneficiary effects on motor skill consolidation (Chen et al., 2019; Khan et al., 2022). Exercise helps students acquire motor memory, which is stored in the form of cortical spine plasticity (Yang et al., 2009; Hayashi-Takagi et al., 2015). During creating a video to show the learned motor skill, students engage in exercise that is effective for motor skill learning (Chen et al., 2019; Khan et al., 2022).

Second, the self-assessment effect builds on empirical studies and suggests that creating a video enables students to self-assess related to their learning and performance by viewing a created video (Paul et al., 1998; Huang, 2015; Ritchie, 2016; Zainuddin and Che' Lah, 2022). Consequently, self-assessment enhances their self-reflection and adjustment. Some studies have shown that creating video raise students' awareness of self-assessment by viewing a created video and improves learning performance in the oral presentation (Ritchie, 2016; Tailab and Marsh, 2020).

Third, the social facilitation effect in motor tasks argues that the mere presence of others increases students' arousal level, intrinsic motivation, and elicits the perception of evaluation by others, and thus influences their performance in motor tasks (Cottrell, 1968; Strauss, 2002). Intrinsic motivation is an essential component of learning, which refers to motivation that originates from within students and consists of spontaneous interest and mastery (Ryan and Deci, 2000). It is a common phenomenon for students to create videos and share them with their classmates later (Huang, 2015, 2021). Therefore, students are aware of the potential audience (i.e., peers) while creating videos. Research has shown that merely believing that someone else is watching you (i.e., a fictitious audience) can evoke arousal (Somerville et al., 2013; Hoogerheide et al., 2018). Based on empirical studies on the social facilitation hypothesis in motor tasks, the mere presence of others facilitates motor skills that have been formulated (Zajonc, 1965; Strauss, 2002).

Although accumulative studies confirmed the benefits of creating a video on students' intrinsic motivation, engagement, and learning performance (Huang, 2015; Hoogerheide et al., 2018; Erlangga, 2021), it is yet unclear whether creating a video is an effective strategy in motor skills and which working mechanism is responsible for such effects. Specifically, it is unclear whether the benefits of creating a video result from engaging in exercise, self-assessment, or others presence.

The present study tested whether creating and sharing a video with classmates or not would improve students' motor skill learning (i.e., roller-skating) compared to the self-exercise condition. To control time on task, students created videos or did self-exercise in a physical class. Merely creating a video and self-exercise were used as control conditions, which also encourages exercise and self-assessment by viewing the created video but lacks others presence. Furthermore, compared to creating and sharing a video with classmates, self-exercise lacks self-assessment and others presence. We examined effects on intrinsic motivation, perseverance in motor tasks, learning satisfaction, and roller-skating skill. Based on previous studies and theories, we proposed the following hypotheses.

> *H1*: Students who create and share a video with classmates will show higher intrinsic motivation, followed by those who merely create a video and, finally, those who do self-exercise.
>
> *H2*: Students who create and share a video with classmates will show higher perseverance in motor tasks, followed by those who merely create a video and, finally, those who do self-exercise.
>
> *H3*: Students who create and share a video with classmates will show better motor skills, followed by those who merely create a video and, finally, those who do self-exercise.

Materials and methods

Participants and design

A total of 160 students (79 females and 81 males, mean age = 20.08, *SD* = 1.28, Range: 18–22 years) from six classes in one Chinese university participated in the study. None of their majors are physics or physics related. We adopted a between-subjects design and thus randomly assigned two classes of students one of three conditions to learn roller-skating. In the creating and sharing a video with classmates condition, there were 53 students (27 females and 26 males, mean age = 20.04 years, *SD* = 1.26, range: 18–22); In the merely creating a video condition, there were 53 students (27 females and 26 males, mean age = 20.32 years, *SD* = 1.33, range: 19–22); In the self-exercise condition, there were 54 students (25 females and 29 males, mean age = 19.87 years, *SD* = 1.23, range: 18–22). All the students were informed of the learning tasks and volunteered to participate in the study.

Martials

Physical fitness tests

Physical fitness tests included squatting against the wall, standing on one leg with eyes closed, sitting and bending forward, 50 m running, standing long jump, and sit-ups. Each test is related to roller skating learning. Specifically, squatting against the wall tests the basic strength of the legs; standing on one leg with eyes closed tests the balance, sitting and bending forward tests the flexibility of the body; 50 m running tests the displacement speed of the body; standing long jump tests the lower limb explosive force; and sit-ups test the waist strength and core control ability. These tests are from the National Student Physical Health Standards in China (Zhang, 2019).

Intrinsic motivation scale

The intrinsic motivation scale was developed by Lin et al. (2020). The scale contains 12 items and includes four dimensions: interest (e.g., "I find this learning activity very interesting"), competence (e.g., "I think I am good at this learning activity"), pressure (e.g., "I'm not nervous at all when doing this study activity"), and value (e.g., "I think this class is very useful to me"). Participants indicated their endorsement of each item on a 7-point scale (e.g., 1 = strongly disagree, 7 = strongly agree). High scores indicated high levels of learning motivation. The Cronbach's alpha coefficients for the four dimensions in the current study were.54, 0.94, 0.94, and.74, respectively.

Perseverance in motor tasks questionnaire

The perseverance in motor tasks questionnaire was adapted from Zhou (1991). The questionnaire measures the degree of students' persistence when encountering learning difficulties, learning disabilities, or external stimuli in motor skills learning. It includes 14 items, such as "When you find a problem that you could not solve in the past during practice, you will try to solve it again within two or 3 days." The participants answered each item based on their actual situation using a five-point Likert scale (1 = completely inconsistent; 5 = completely consistent). The higher the score, the higher the persistence of the participants. The Cronbach's alpha coefficient of the questionnaire in the current study was 0.91.

Learning satisfaction scale

The learning satisfaction scale was adapted from Wang (2014). The scale contains 17 items and four dimensions: Teaching, course content, teacher-student interaction, and learning environment and equipment. Teaching contains six items, such as "Teachers' professional knowledge is very rich, and I have learned new knowledge and skills"; course content contains five items, such as "The task in learning is challenging"; teacher-student interaction contains three items, such as "Through mutual learning, I can fully participate in the learning process"; learning environment and equipment contains three items, such as "Satisfied with the normal operation of teaching." The students needed to evaluate every item according to their feelings using a five-point Likert scale (1 = completely disagree; 5 = completely agree). The higher the score, the more satisfied students are with teaching. The Cronbach's alpha coefficients for the four dimensions in the current study were 0.95, 0.96, 0.97, and 0.98, respectively.

Roller-skating test

Students completed a total of 26 m roller skating in a roller-skating test. The test included three parts. In part one, students slid forward 10 m from the starting point. In part two, students turned left and slid forward for 6 m. In part three, students turned left and slid forward for 10 m to get the ending point. An instructor at the ending point recorded time. The finish time of 26 m was the roller-skating achievement for students. Shorter the finish time, the better the roller-skating achievement. The test map can be seen in Figure 1.

Procedure

Figure 2 shows the procedure. First, participants completed physical fitness tests. After that, each class was randomly assigned

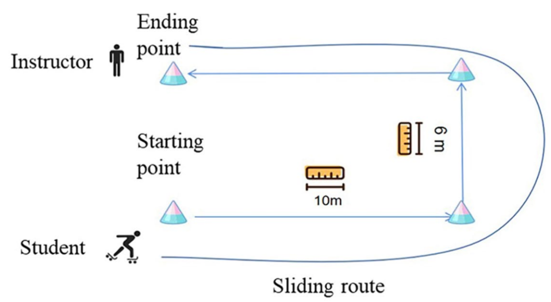


FIGURE 1
Sliding route of the roller-skating test.

to one experimental condition, and a male instructor taught the exact content of roller-skating, including standing and sliding. Participants did exercises following the instructor. Then, in the creating and sharing a video condition, the instructor asked the students to create an explaining video in a physical class and share it with their classmates in the WeChat group (i.e., a popular social media in China). In the merely creating a video condition, the instructor asked the students to create an explaining video in a physical class. In the self-exercise condition, students watched the explaining video created by the instructor rather than creating their videos and did exercises themselves in a physical class. At last, all students completed the roller-skating test. These steps were repeated three times. Finally, after the third class, all students finished the questionnaires.

Data analysis

We first tested whether there were differences in the physical fitness tests among the three conditions. We conducted between-group comparisons to test H1, H2 and H3: intrinsic motivation, perseverance in motor tasks, and performance on the roller-skating test. If there were differences in the physical fitness tests among the three conditions, to eliminate the potential influence of the participants' physical fitness on the results, we conducted these comparisons with the analysis of covariance (ANCOVA), in which we used the scores on the physical fitness tests as covariates in our models. All analyses were carried out on SPSS 26.

Result

Table 1 shows the means and standard deviations (*SD*) of all variables in the current study.

Physical fitness tests

We first tested the differences in physical fitness tests among the three conditions. The results of one-way ANOVAs showed significant differences in squatting against the wall, $F(2,157) = 5.27$, $p = 0.006$, $\eta_p^2 = 0.06$. *Post hoc* comparison (*LSD*) showed that participants' time in the creating and sharing a video condition was significantly shorter than those in the self-exercise condition ($MD = -33.38$, $p = 0.002$, 95% CI [−53.83, −12.92]).

There were significant differences in standing on one leg with eyes closed, $F(2,157) = 4.36$, $p = 0.014$, $\eta_p^2 = 0.05$. *Post hoc* comparison showed that participants' time in the creating and sharing a video condition was significantly shorter than those in the self-exercise condition ($MD = -35.41$, $p = 0.004$, 95% CI [−59.13, −11.68]).

There were marginally significant differences in sitting and bending forward, $F(2,157) = 2.41$, $p = 0.094$, $\eta_p^2 = 0.03$. *Post hoc* comparison showed that participants' distance in the creating and sharing a video condition was significantly longer than those in the merely creating a video condition ($MD = 2.813$, $p = 0.032$, 95% CI [0.26, 5.37]);

There were significant differences in 50 m running, $F(2,157) = 5.39$, $p = 0.005$, $\eta_p^2 = 0.06$. *Post hoc* comparison showed that participants' time in the creating and sharing a video condition was significantly longer than those in the self-exercise condition ($MD = 0.62$, $p = 0.001$, 95% CI [0.25, 1.00]).

There were significant differences in standing long jump, $F(2,157) = 5.34$, $p = 0.006$, $\eta_p^2 = 0.06$. *Post hoc* comparison showed that participants' distance in the creating and sharing a video condition was significantly shorter than those in the merely creating a video condition ($MD = -14.81$, $p = 0.038$, 95% CI [−28.75, −0.86]), and was significantly shorter than those in the self-exercise condition ($MD = -22.64$, $p = 0.002$, 95% CI [−36.52, −8.75]);

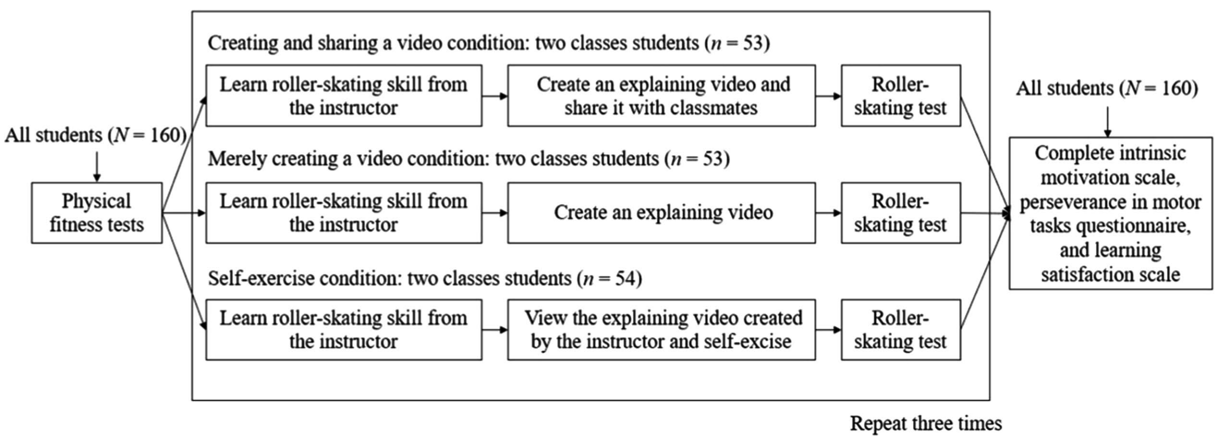


FIGURE 2
The procedure of the current study.

TABLE 1 Means and standard deviations (*SD*) of all variables.

	Creating and sharing a video		Merely creating a video		Self-exercise	
	Mean	***SD***	**Mean**	***SD***	**Mean**	***SD***
Physical fitness tests						
Squatting against the wall (s)	87.70	45.85	108.02	58.80	121.08	55.16
Standing on one leg with eyes closed (s)	65.08	55.96	81.02	63.87	100.48	66.04
Sitting and bending forward (cm)	17.41	6.27	14.60	6.65	16.31	7.02
50 m running (s)	8.51	0.95	8.14	0.93	7.89	1.08
Standing long jump (cm)	186.04	32.98	200.85	33.03	208.67	42.14
Sit-ups	45.70	9.11	41.43	9.30	39.83	9.68
Intrinsic motivation						
Interest	5.20	1.17	5.01	1.02	5.12	1.37
Competence	5.37	1.54	5.04	1.35	4.98	1.55
Pressure	5.06	1.33	4.86	1.21	5.04	1.57
Value	5.86	1.26	5.34	1.09	5.38	1.31
Perseverance in motor tasks	4.07	0.48	3.59	0.68	3.69	0.79
Learning satisfaction						
Teaching	4.40	0.81	4.17	0.79	3.98	0.92
Classroom content	4.35	0.82	4.07	0.78	3.96	0.95
Teacher-student interaction	4.43	0.80	4.07	0.80	4.04	1.00
Learning environment and equipment	4.40	0.92	4.06	0.86	4.01	1.01
Roller skating tests (s)	23.29	14.16	20.50	10.60	20.85	9.30

There were significant differences in sit-ups, $F(2,157) = 5.59$, $p = 0.004$, $\eta_p^2 = 0.07$, and *post hoc* comparison showed that participants' number in the creating and sharing a video condition was significantly more than those in the merely creating a video condition ($MD = 4.26$, $p = 0.020$, 95% CI [0.67, 7.86]), and was significantly more than those in the self-exercise condition ($MD = 5.87$, $p = 0.001$, 95% CI [2.29, 9.44]).

In sum, the above results indicated that participants in the three conditions were different in physical fitness. Therefore, we conducted ANCOVAs to test the effects of creating video on motor skill learning, the scores on the physical fitness tests as covariates.

Intrinsic motivation

For the interest, competence, and pressure dimensions, there was no significant difference between the three conditions, $F(2,151) = 0.59$, $p = 0.556$, $\eta_p^2 = 0.01$; $F(2,151) = 0.63$, $p = 0.536$, $\eta_p^2 = 0.01$; $F(2,151) = 0.50$, $p = 0.607$, $\eta_p^2 = 0.01$.

For the value dimension, there were significant differences between the three conditions, $F(2,151) = 3.15$, $p = 0.046$, $\eta_p^2 = 0.04$. The *post hoc* tests showed that participants in the creating and sharing a video condition reported a significantly higher score than those in the merely creating a video condition ($MD = 0.60$, $p = 0.016$, 95% CI [0.11, 1.10]) and a higher score, in a marginally significant level, than those in the self-exercise condition ($MD = 0.50$, $p = 0.060$, 95% CI [−0.02, 1.03]). See Figure 3.

Perseverance in learning

The ANCOVA showed that there were significant differences in the score of the perseverance in motor tasks questionnaire between the three conditions, $F(2,151) = 6.42$, $p = 0.002$, $\eta_p^2 = 0.08$. The *post hoc* tests showed that participants in the creating and sharing a video condition reported a significantly higher score than those in the merely creating a video condition ($MD = 0.49$, $p < 0.001$, 95% CI [0.22, 0.76]) and the self-exercise condition ($MD = 0.35$, $p = 0.017$, 95% CI [0.06, 0.64]). See Figure 4.

Learning satisfaction

For the teaching dimension, there were differences in a marginally significant level between the three conditions, $F(2,151) = 2.60$, $p = 0.078$, $\eta_p^2 = 0.03$. The *post hoc* test showed that participants in the creating and sharing a video condition reported a significantly higher score than those in the self-exercise condition ($MD = 0.41$, $p = 0.027$, 95% CI [0.05, 0.77]). See Figure 5.

For the teacher-student interaction dimension, there were a difference in a marginally significant level between the three conditions, $F(2,151) = 2.65$, $p = 0.074$, $\eta_p^2 = 0.03$. The *post hoc* test showed that participants in the creating and sharing a video condition reported a significantly higher score than those in the merely creating a video condition ($MD = 0.40$, $p = 0.028$, 95% CI [0.04, 0.75]), and a higher score, in a marginally significant level, than those in the self-exercise condition ($MD = 0.34$, $p = 0.08$, 95% CI [−0.04, 0.71]). See Figure 6.

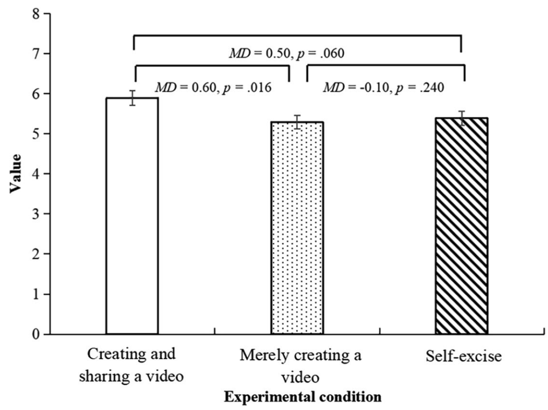


FIGURE 3
The bar chart for the value dimension.

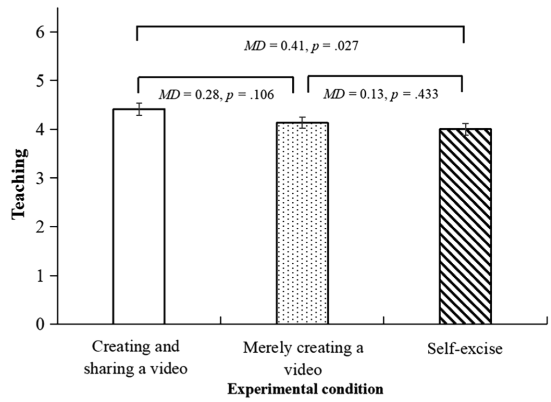


FIGURE 5
The bar chart for the teaching dimension.

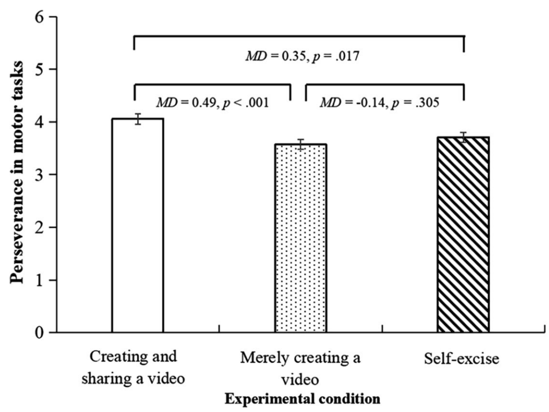


FIGURE 4
The bar chart for perseverance in motor tasks.

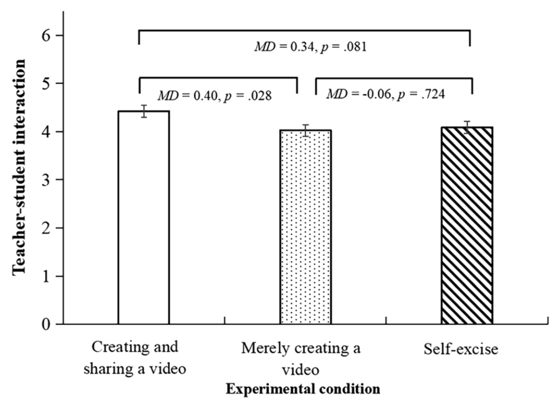


FIGURE 6
The bar chart for the teacher-student interaction dimension.

For the course content and the learning environment and equipment dimensions, there was no significant difference between the three conditions, $F(2,151) = 2.23$, $p = 0.111$, $\eta_p^2 = 0.03$; $F(2,151) = 1.95$, $p = 0.146$, $\eta_p^2 = 0.03$.

Roller-skating test

We calculated the average time of three roller-skating tests as the final scores. The result showed no significant difference between the three conditions, $F(2,151) = 0.68$, $p = 0.508$, $\eta_p^2 = 0.01$.

Discussion

The current study aimed to investigate whether creating and sharing a video with classmates would be more effective than merely creating a video and self-exercise to learn a motor skill in terms of intrinsic motivation, perseverance in learning, learning satisfaction, and roller-skating skill. Partially consistent with our hypothesis, we found that creating and sharing a video with classmates increased students' intrinsic motivation, perseverance in motor tasks, and learning satisfaction, but not roller-skating skill, followed by merely creating a video and then self-exercise. The results broadened our understanding of the effects of creating a video on motor skills learning.

In line with our first and second hypotheses that creating a video improved intrinsic motivation and perseverance in motor tasks compared with self-exercise, we confirmed the benefits of creating a video on students' intrinsic motivation (i.e., the value dimension) and perseverance in motor tasks compared to self-exercise. The results were consistent with previous studies, showing creating a video increases students' intrinsic motivation (i.e., value) and learning satisfaction (Greene and Crespi, 2012; Yang and Wu, 2012; Huang, 2015; Zainuddin and Che' Lah, 2022). Previous studies have shown that creating a video encourages students' engagement and supports support authentic learning (Kearney and Schuck, 2003; Schuck and Kearney, 2004; Greene and Crespi, 2012). Some researchers defined behavioral engagement as effort and perseverance in learning (Lee, 2014).

Furthermore, intrinsic motivation is positively associated with enhanced perseverance in learning (Larson and Rusk, 2011). Students may have been more inclined to persevere with the difficult material due to intrinsic motivation (Logan et al., 2011).

In contrast to our third hypothesis that creating and sharing a video with classmates facilitated roller-skating skills most, we did not observe differences in the roller-skating skills among the three conditions. The results suggested that self-exercise was enough to lead to good performance in the roller-skating test. Hoogerheide et al. (2019) found that creating videos did not outperform the summarizing condition, like the self-exercise condition in the present study. Additionally, the study by Sari et al. (2020) also did not observe students' English skills/proficiency was improved by creating video sharing on YouTube. Therefore, exercise, self-assessment, or other presence seem to not contribute to the mechanisms underlying the benefits of creating video. In contrast, teaching expectancy might play an essential role in the benefits of creating videos observed in previous studies (Hoogerheide et al., 2018, 2019). The studies on teaching by creating videos have shown that students in the teaching on video condition consistently outperform and report more engagement (Hoogerheide et al., 2016, 2018, 2019; van Brussel et al., 2021). Future work should test whether a teaching expectancy enhances the benefits of creating a video on learning motor skills.

This study has the following limitations. First, this study did not test the long-term effect of creating and sharing a video with classmates. We observed that creating and sharing a video increased students' intrinsic motivation and perseverance in motor tasks. Intrinsic motivation is a powerful "engine" of learning (Larson and Rusk, 2011). Intrinsic motivation is associated with enhanced engagement and perseverance in learning but with greater use of meta-cognitive strategies and deeper information processing (Larson and Rusk, 2011). Therefore, the benefits of creating and sharing a video with classmates might be exhibited over time. Future studies should test the long-term effect of creating and sharing a video with classmates. Second, we did not analyze the videos created by students. Previous studies on generative learning have shown that the quantity and quality of generative learning activities mediate the effect of generative learning strategies (e.g., self-explaining) on learning performance (Baars et al., 2018; Hefter, 2021; Bichler et al., 2022). Future research should further investigate the mediating role of the quantity and quality of the videos created by students.

Although there were some limitations, the present study advances our understanding of the various effects of two types of creating video compared to self-exercise. With the rapid development of video technology, it is easy for students to create a video as an enjoyable activity (Greene and Crespi, 2012; Hoogerheide et al., 2019). The present study tested the various effects of video creation by including three conditions: creating and sharing a video with classmates, merely creating a video, and self-exercise. Additionally, we measured students' subjective perceptions and their objective performance on the roller-skating test. Our results suggest that video creation enhances students' intrinsic motivation, perseverance in motor tasks, and learning satisfaction compared to self-exercise; sharing with classmates can further enhance their intrinsic motivation, perseverance in motor tasks, and learning satisfaction; self-exercise guarantees good performance on motor skill.

In conclusion, the main finding of the present study was that students who created and shared a video with classmates reported higher intrinsic motivation, perseverance in motor tasks, and learning satisfaction, but not the roller-skating skill, followed by those who merely creating a video and then those who self-exercise. With the easy availability of video production software, the findings have an important implication for motor skills learning: during teaching motor skills, teachers can use encourage students to create and share a video with classmates as a homework activity to increase students' intrinsic motivation, perseverance in motor tasks, and learning satisfaction.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Zhejiang Gongshang University. The participants provided their written informed consent to participate in this study.

Author contributions

QX: conceptualization, methodology, writing – original draft, and writing-reviewing and editing. LK: investigation, data curation, visualization, and methodology. ZZ: supervision and conceptualization. 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.

References

Aguilar, D., and Pifarre Turmo, M. (2019). Promoting social creativity in science education with digital technology to overcome inequalities: a scoping review. *Front. Psychol.* 10:1474. doi: 10.3389/fpsyg.2019.01474

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychol. Rev.* 89, 369–406. doi: 10.1037/0033-295X.89.4.369

Baars, M., Leopold, C., and Paas, F. (2018). Self-explaining steps in problem-solving tasks to improve self-regulation in secondary education. *J. Educ. Psychol.* 110, 578–595. doi: 10.1037/edu0000223

Bichler, S., Stadler, M., Bühner, M., Greiff, S., and Fischer, F. (2022). Learning to solve ill-defined statistics problems: does self-explanation quality mediate the worked example effect? *Instr. Sci.* 50, 335–359. doi: 10.1007/s11251-022-09579-4

Chang, K.-E., Zhang, J., Huang, Y.-S., Liu, T.-C., and Sung, Y.-T. (2020). Applying augmented reality in physical education on motor skills learning. *Interact. Learn. Environ.* 28, 685–697. doi: 10.1080/10494820.2019.1636073

Chen, K., Zheng, Y., Wei, J., Ouyang, H., Huang, X., Zhang, F., et al. (2019). Exercise training improves motor skill learning via selective activation of mTOR. Science. *Advances* 5:eaaw1888. doi: 10.1126/sciadv.aaw1888

Cottrell, N. B. (1968). "Performance in the presence of other human beings: mere presence, audience, and affiliation effects," in *Social Facilitation and Imitative Behavior*. eds. E. C. Simmel, R. A. Hoppe and S. A. Milton (Boston, MA: Allyn and Bacon).

Erlangga, R. A. (2021). Students' motivation in learning English through English learning videos making. *Soc. Sci. Stud.* 1, 177–189. doi: 10.47153/sss13.2542021

Finkenberg, M. E., Fiorentino, L. H., and Castelli, D. (2005). Creating a virtual gymnasium. *J. Phys. Educat. Recreat. Dance* 76, 16–18. doi: 10.1080/07303084.2005.10608231

Garland, H. (1982). Goal levels and task performance: a compelling replication of some compelling results. *J. Appl. Psychol.* 67, 245–248. doi: 10.1037/0021-9010.67.2.245

Greene, H., and Crespi, C. (2012). The value of student created videos in the college classroom – an exploratory study in marketing and accounting. *Int. J. Arts Sci.* 5, 273–283.

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y. I., Loshbaugh, A. L., et al. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature* 525, 333–338. doi: 10.1038/nature15257

Hefter, M. H. (2021). Web-based training and the roles of self-explaining, mental effort, and smartphone usage. *Technol. Knowl. Learn.* doi: 10.1007/s10758-021-09563-w [Epub ahead of print]

Hoogerheide, V., Deijkers, L., Loyens, S. M. M., Heijltjes, A., and Van Gog, T. (2016). Gaining from explaining: learning improves from explaining to fictitious others on video, not from writing to them. *Contemp. Educ. Psychol.* 44–45, 95–106. doi: 10.1016/j.cedpsych.2016.02.005

Hoogerheide, V., Renkl, A., Fiorella, L., Paas, F., and Van Gog, T. (2018). Enhancing example–based learning: teaching on video increases arousal and improves problem–solving performance. *J. Educ. Psychol.* 111, 45–56. doi: 10.1037/edu0000272

Hoogerheide, V., Visee, J., Lachner, A., and Van Gog, T. (2019). Generating an instructional video as homework activity is both effective and enjoyable. *Learn. Instr.* 64:101226. doi: 10.1016/j.learninstruc.2019.101226

Huang, H. (2015). The effects of video projects on EFL learners' language learning and motivation: an evaluative study. *Int. J. Comput. Assist. Lang.* 5, 53–70. doi: 10.4018/IJCALLT.2015010104

Huang, H.-W. (2021). Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement. *Australas. J. Educ. Technol.* 37, 18–40. doi: 10.14742/ajet.6623

Jung, C. D. (2021). Perceptions of collaborative video projects in the language classroom: a qualitative case study. *Int. J. Instr.* 14, 301–320. doi: 10.29333/iji.2021.14418a

Kearney, M., and Schuck, S. (2003). "Authentic learning through the use of digital video." in *Proceedings of the Australian Computers in Education Conference*. eds. W. Au and B. White. Adelaide, Australia.

Khan, A., Singh, J., Neary, J. P., and Mang, C. S. (2022). Exercise effects on motor skill consolidation and intermuscular coherence depend on practice schedule. *Brain Sci.* 12:436. doi: 10.3390/brainsci12040436

Larson, R. W., and Rusk, N. (2011). Intrinsic motivation and positive development. *Adv. Child Dev. Behav.* 41, 89–130. doi: 10.1016/B978-0-12-386492-5.00005-1

Lee, J. S. (2014). The relationship between student engagement and academic performance: is it a myth or reality? *J. Educ. Res.* 107, 177–185. doi: 10.1080/00220671.2013.807491

Leonard, S., and Weitz, J. (1971). Task enjoyment and task perseverance in relation to task success and self-esteem. *J. Appl. Psychol.* 55, 414–421. doi: 10.1037/h0031706

Lin, L., Ginns, P., Wang, T., and Zhang, P. (2020). Using a pedagogical agent to deliver conversational style instruction: what benefits can you obtain? *Comput. Educ.* 143:103658. doi: 10.1016/j.compedu.2019.103658

Logan, S., Medford, E., and Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learn. Individ. Differ.* 21, 124–128. doi: 10.1016/j.lindif.2010.09.011

Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., et al. (2021). Video improves learning in higher education: a systematic review. *Rev. Educ. Res.* 91, 204–236. doi: 10.3102/0034654321990713

Paul, S., Dawson, K. P., Lanphear, J. H., and Cheema, M. Y. (1998). Video recording feedback: a feasible and effective approach to teaching history-taking and physical examination skills in undergraduate paediatric medicine. *Med. Educ.* 32, 332–336. doi: 10.1046/j.1365-2923.1998.00197.x

Pi, Z., Zhang, Y., Zhou, W., Xu, K., Chen, Y., Yang, J., et al. (2021). Learning by explaining to oneself and a peer enhances learners' theta and alpha oscillations while watching video lectures. *Br. J. Educ. Technol.* 52, 659–679. doi: 10.1111/bjet.13048

Pi, Z., Zhu, F., Zhang, Y., Chen, L., and Yang, J. (2022). Complexity of visual learning material moderates the effects of instructor's beat gestures and head nods in video lectures. *Learn. Instr.* 77:101520. doi: 10.1016/j.learninstruc.2021.101520

Rekik, G., Belkhir, Y., Mezghanni, N., Jarraya, M., Chen, Y. S., and Kuo, C.-D. (2021). Learning basketball tactical actions from video modeling and static pictures: when gender matters. *Children* 8:1060. doi: 10.3390/children8111060

Ritchie, S. M. (2016). Self-assessment of video-recorded presentations: does it improve skills? *Act. Learn. High. Educ.* 17, 207–221. doi: 10.1177/1469787416654807

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68

Sari, A. B. P., Dardjito, H., and Azizah, D. M. (2020). EFL students' improvement through the reflective you tube video project. *Int. J. Instr.* 13, 393–408. doi: 10.29333/iji.2020.13425a

Schuck, S., and Kearney, M. (2004). Students in the director's seat: teaching and learning across the school curriculum with student-generated video (Research Report). Available at: http://epress.lib.uts.edu.au/research/bitstream/handle/10453/14209/SchuckKearney.04.pdf?sequence=1

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., and Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychol. Sci.* 24, 1554–1562. doi: 10.1177/0956797613475633

Strauss, B. (2002). Social facilitation in motor tasks: a review of research and theory. *Psychol. Sport Exerc.* 3, 237–256. doi: 10.1016/S1469-0292(01)00019-X

Tailab, M., and Marsh, N. (2020). Use of self-assessment of video recording to raise students' awareness of development of their oral presentation skills. *High. Educ. Stud.* 10, 16–28. doi: 10.5539/hes.v10n1p16

van Abswoude, F., Van Der Kamp, J., and Steenbergen, B. (2019). The roles of declarative knowledge and working memory in explicit motor learning and practice among children with low motor abilities. *Human Kinetics Journals* 23, 34–51. doi: 10.1123/mc.2017-0060

van Brussel, S., Timmermans, M., Verkoeijen, P., and Paas, F. (2021). Teaching on video as an instructional strategy to reduce confirmation bias—a pre-registered study. *Instr. Sci.* 49, 475–496. doi: 10.1007/s11251-021-09547-4

Wang, X. (2014). *Research on the Influence of Information Presentation and Self-Efficacy on Learning Effect*. Liaoning: Liaoning normal university.

Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462, 920–924. doi: 10.1038/nature08577

Yang, Y., and Wu, I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: a year-long experimental study. *Comput. Educ.* 59, 339–352. doi: 10.1016/j.compedu.2011.12.012

Zainuddin, K., and Che' Lah, C.-F. (2022). Students' perceptions on the implementation of recorded oral presentation. *J. Administr. Sci.* 19, 233–246.

Zajonc, R. B. (1965). Social facilitation. *Science* 149, 269–274. doi: 10.1126/science.149.3681.269

Zhang, Y. (2019). Physical literacy-oriented "National Student Physical Health Standards" (Primary stage) formulation analysis and implementation research. *Unplished Doctoral Dissertation. Hunan Normal University.*

Zhou, B. (1991). *Academic Adjustment Test (AAT)*. Beijing: Beijing Normal University.

frontiers | Frontiers in Psychology

TYPE Curriculum, Instruction, and Pedagogy
PUBLISHED 07 December 2022
DOI 10.3389/fpsyg.2022.1076013

How data knowledge transformation affects teachers' precision teaching ability: The mediating effect of data consciousness

Jia Tao[1], Chenchen Fan[1]*, Wenwen Wu[1] and Yongyuan Zhu[2]

[1]School of Educational Science, Anhui Normal University, Wuhu, Anhui, China, [2]Party and Government Office, Anhui Medical University, Hefei, Anhui, China

Check for updates

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University, China

REVIEWED BY
Qinna Feng,
Central China Normal University, China
Gege Li,
Hokkaido University of Education, Japan

*CORRESPONDENCE
Chenchen Fan
1921010039@ahnu.edu.cn

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

RECEIVED 21 October 2022
ACCEPTED 14 November 2022
PUBLISHED 07 December 2022

CITATION
Tao J, Fan C, Wu W and Zhu Y (2022) How data knowledge transformation affects teachers' precision teaching ability: The mediating effect of data consciousness. *Front. Psychol.* 13:1076013. doi: 10.3389/fpsyg.2022.1076013

With the development of smart education, teachers' precision teaching ability has increasingly become an important component of their professional ability. Based on the Socialization, Externalization, Combination, Internalization (SECI) model of knowledge management, this research constructs a theoretical model of data knowledge transformation, data consciousness, and teachers' precision teaching ability, and uses Partial Least Squares to analyze the interaction and influence among the elements. The empirical results show that the transformation of internal knowledge related to data has a significant positive impact on the improvement of teachers' precision teaching ability and data consciousness; the transformation of external knowledge related to teaching has a positive effect on the teaching management ability of teachers' precision teaching ability, but the effect is not significant. Data consciousness plays a partial mediating effect between internal knowledge transformation and teachers' precision teaching ability, and a complete mediating effect between external knowledge transformation and teaching management ability.

KEYWORDS

data knowledge, teachers' precision teaching ability, mediating effect, knowledge transformation, data consciousness

Introduction

Precision teaching is a representative of technology-enabled education in the era of big data, and is a major structural change in the field of teaching and learning, which brings many new challenges to the academia (Ministry of Education of the People's Republic of China, 2021). The incongruity and mismatch between teachers' previous teaching experience and data knowledge transformation is one of the challenges teachers face in promoting precision teaching (Cui and Zhang, 2022). Some studies have reported that teachers' data knowledge transformation implies teachers' ability to recognize, use, and evaluate data technology (Mandinach and Gummer, 2013).

Teachers who have better data knowledge transformation skills are more successful in implementing precision teaching (Cabero et al., 2022). It can be argued that teachers' lack of attention to or misconceptions about the development of data knowledge transformation skills can directly affect the improvement of teachers' precision teaching skills. To examine exactly how data knowledge transformation affects teachers' precision teaching ability, we first need to understand what teachers' precision teaching ability is.

Research on teachers' precision teaching ability reveals that scholars have affirmed the importance of precision teaching, while also raising a series of questions during the inquiry. Hayes et al. (2018) noted that precision teaching realizes the automation of teaching but weakens the pedagogical significance of teachers' experience. Fletcher et al. (2013) take the smart campus as the context and reveal the current increased workload of teachers in precision teaching, such as the diversification of roles and the refinement of division of functions. Kubina and Yurich (2009) verified that the practice of precision teaching is positively related to teachers' personal ability to use data through precision teaching research. Moreover, some scholars investigated the definition of teachers' precision teaching ability. For example, Peng and Zhu (2018) proposed that teachers' precision teaching ability allows them to use learning data to enhance teaching, propose teaching strategies based on evidence, and implement teaching practice. Pierce et al. (2014) argued that to achieve precision teaching, teachers should have professional (statistics) and situational (data generation) knowledge. Theall (2001) argued that teachers' precision teaching ability is a high-level ability to gather practical teaching experience and data analysis skills, and to learn scientific theories. However, Mandinach (2012) posited that teachers' precision teaching ability is basically consistent with their data literacy; Mandinach proposes that teachers' precision teaching ability is a comprehensive ability that can transform data into teaching strategies to meet students' specific needs. Scholars have differing views on the definition of teachers' precision teaching ability.

Data consciousness is a concept at the cognitive level, which requires teachers to know that data is meaningful and valuable, and be aware of the possible harm caused by poor data management. From the existing studies, there seems to be a lack of empirical research on the relationship between data knowledge transformation, teachers' precision teaching ability and data consciousness. Therefore, this study aims to verify the mediating effect model of teachers' data knowledge transformation, data consciousness, and precision teaching ability through questionnaire data under the framework of knowledge management theory, and provide evidence to clarify the mechanism of action among the three. The research questions of this study are as follows:

- (RQ1)Does teachers' data knowledge learning and transformation have an effect on the improvement of their precision teaching ability? How does it affect?
- (RQ2)What is the role of teachers' data consciousness in the relationship between data knowledge transformation and teachers' precision teaching ability?

Literature review and theoretical assumptions

Analyzed from the perspective of pedagogy, the teaching practice mainly contains four dimensions: learning situation analysis, teaching management, teaching decision, and teaching evaluation (Srikoom et al., 2018). Therefore, teachers' precision teaching ability can be seen as a comprehensive ability of teachers to use data to advance the four dimensions of learning analysis, teaching management, teaching decision making, and teaching evaluation. From the perspective of the socialization, externalization, combination, internalization (SECI) model of knowledge management, the improvement of teachers' precision teaching ability mainly includes explicit knowledge and implicit knowledge, explicit knowledge is directly related to data and can be shared in a standardized and systematic language, implicit knowledge mainly includes data applications, understanding, data-driven teaching and learning, and other knowledge related to beliefs, metaphors, and thinking patterns (Revuelta-Domínguez et al., 2022). These elements have interacted and changed together during the transformation of explicit and implicit knowledge to create a whole knowledge innovation. This process of the externalization, integration, internalization, and socialization of knowledge ultimately points to the improvement of teachers' precision teaching ability (Weinberger and Green, 2022).

The relationship between data knowledge transformation and teachers' precision teaching ability

The transformation of data knowledge corresponds to teachers' precision teaching ability. Based on the SECI model, teachers' data knowledge can be divided into two categories. One is internal knowledge transformation, which focuses on teachers' learning of data-related technologies and resources; the other is external knowledge transformation, which focuses on the combination of teacher experience and data knowledge, and runs through the processes of teaching management, evaluation, decision-making, and others.

Internal knowledge transformation and teachers' precision teaching ability

Explicit knowledge of data mainly includes development and settings of the platform, reading and calculation of the data, use and processing of software, and data security (Duraiswami and Krishnamurthy, 2020), The absorption of internal knowledge mainly comes from thematic learning. Before teachers use technical means to carry out precision teaching, they need to learn a lot of relevant data knowledge, especially dealing with technology. Research has found that the effects of precision teaching are often limited by teachers' mastery of platforms, software, and other data tools (Lyons, 1984). The accumulation of teachers' internal knowledge can effectively improve their understanding of precision teaching, reduce the costs of trial and error, and play an important role in teachers' active precision teaching changes (Wang et al., 2011). Therefore, this paper asserts that the transformation of internal knowledge can effectively promote the improvement of teachers' precision teaching ability, and we propose the following hypotheses:

> *H1: There is a positive correlation between internal knowledge transformation and teachers' precision teaching ability.*
> *H1a: There is a positive correlation between internal knowledge transformation and student analysis ability.*
> *H1b: There is a positive correlation between internal knowledge transformation and teaching management ability.*
> *H1c: There is a positive correlation between internal knowledge transformation and educational decision-making ability.*
> *H1d: There is a positive correlation between internal knowledge transformation and instructional evaluation ability.*

External knowledge transformation and teachers' precision teaching ability

Tacit data knowledge mainly includes data analysis and evaluation, educational decision-making and integration, and conclusion dissemination, among others (Kabir and Carayannis, 2013). In the process of introducing external knowledge, teachers actively participate in data processing, which can improve their ability to engage in data-driven teaching. While absorbing external knowledge, teachers can improve their own educational and teaching experience through creation, exchange, integration, and other activities. In the application stage of external knowledge, teachers promote the sharing of new knowledge generated by precision teaching practice to recreate new knowledge and promote the personalized reform of education. This process is the ultimate goal of precision teaching (Haigh, 2008). Therefore, this paper posits that the transformation of external knowledge has a positive impact on improving teachers' precision teaching ability and promoting precision teaching practice. We offer the following hypotheses:

> H2: There is a positive correlation between external knowledge transformation and teachers' precision teaching ability.
> *H2a: There is a positive correlation between external knowledge transformation and student analysis ability.*
> *H2b: There is a positive correlation between external knowledge transformation and teaching management ability.*
> *H2c: There is a positive correlation between external knowledge transformation and educational decision-making ability.*
> *H2d: There is a positive correlation between external knowledge transformation and instructional evaluation ability.*

The relationship between data knowledge transformation and data consciousness

Internal knowledge transformation and data consciousness

According to existing research, teachers can have clear conflicts with the data generated in teaching (Du and Su, 2021). There are three main reasons for such conflicts. First, it is difficult to correctly interpret different forms of measurement data; second, it is difficult to convert existing information into appropriate teaching strategies to solve the problem of data representation; and third, it is difficult to interpret large data sets, resulting in a certain sense of powerlessness and fear. This sense of teaching burden generated by the data is one of the major challenges to the in-depth practice of precision teaching at this stage. Internal knowledge related to data can be learned through centralized learning, communication, text training, and other explicit tasks (Friedman et al., 1979). Therefore, this paper asserts that through the study of internal knowledge, such as data, teachers' fears of data can be effectively alleviated to deepen their knowledge and understanding of data, which has a positive impact on the cultivation of teachers' data consciousness. Therefore, the following hypothesis is proposed:

H3: There is a positive correlation between internal knowledge transformation and data consciousness.

External knowledge transformation and data consciousness

For teachers, learning about knowledge related to educational data is meaningless if it is separated from the educational context. However, the results of our survey show that at this stage, the internal knowledge learning of relevant data is mainly controlled by administrative departments, such as education authorities and schools, which characterized by policy and compulsion. We also note that after learning about these topics, teachers hardly use data knowledge; the reason for this is that teachers lack understanding of the value of data (Williamson, 2012). Training for precision teaching focuses on the internal knowledge level, and teachers equate this with technology learning, without paying much attention to it. In the learning and transformation of external knowledge, strengthening deduction skills, exchange, data sharing, and data-driven education practice can help teachers realize the advantages of data-driven precision teaching. Therefore, this paper argues that the transformation of external knowledge also has a positive impact on improving teachers' data consciousness. The following hypothesis is proposed:

H4: There is a positive correlation between external knowledge transformation and data consciousness.

The relationship between data consciousness and teachers' precision teaching ability

From the research on the professional quality of pedagogical students' education and teachers' post-service education, data-related knowledge is not the focus of training, according to the Professional Standards for Teachers in Kindergartens, Primary Schools, and Middle Schools, which were customized by the Ministry of Education in 2012. Rather, the professional quality of teachers in China mainly includes three aspects: teacher ethics, professional knowledge, and professional ability (Ministry of Education of the People's Republic of China, 2012). Based on the current general trend of smart education development, data-driven decision-making has gradually been integrated into all aspects of teachers' ethics, professional knowledge, and professional ability requirements. Teachers often face difficulties in learning and transforming data-related knowledge before applying this knowledge in precision teaching practice. In the current training program, although teachers have learned new data knowledge through professional development, they still know little about data comprehension, technology, and new content (McDowall et al., 2021). From the learning of data knowledge to learning how to apply data in teaching and learning, teachers need to awaken their personal interest in data learning. Therefore, this paper suggests that the cultivation of teachers' data consciousness has a positive impact on improving teachers' precision teaching ability, and the following hypotheses are proposed:

H5: There is a positive correlation between data consciousness and teachers' precision teaching ability.

H5a: There is a positive correlation between data consciousness and student analysis ability.

H5b: There is a positive correlation between data consciousness and teaching management ability.

H5c: There is a positive correlation between data consciousness and educational decision-making ability.

H5d: There is a positive correlation between data consciousness and instructional evaluation ability.

The mediating role of data consciousness

As internal and external knowledge are transformed, teachers have changed focus from the learning of explicit statistical knowledge to the implicit formation of an individual understanding of data. The latter has in turn transformed into information through internalization and been applied to teaching practice to guide the processes of student analysis, teaching management, educational decision-making, and instructional evaluation. The transformation of explicit and implicit knowledge is a two-way, cyclical accumulation process (Wieser, 2016). Individual teachers acquire the ability and teaching experience working with data and statistics through learning or practice, guided by data consciousness; the educational significance of the raw data is identified through the process of connection and the internalization of explicit knowledge. The teacher's movement from understanding data to summarizing information and decision-making in teaching practice is the process of discovering practical problems, creating new knowledge, assimilating existing implicit knowledge, and innovating and integrating to generate new explicit knowledge under the impetus of data consciousness. Whether it is the transformation of internal knowledge or the transformation of external knowledge, this process ultimately points to the improvement of teachers' precision teaching ability.

Knowledge is a component of data consciousness (Filderman et al., 2022), and data consciousness is the cornerstone of cultivating teachers' precision teaching ability. Therefore, this paper argues that data consciousness plays a mediating role between internal and external knowledge transformation and teachers' precision teaching ability. We propose the following hypotheses:

H6: Data consciousness plays a mediating role between data knowledge and teachers' precision teaching ability.

H6a: Data consciousness plays a mediating role between internal knowledge and student analysis ability.

H6b: Data consciousness plays a mediating role between internal knowledge and teaching management ability.

H6c: Data consciousness plays a mediating role between internal knowledge and educational decision-making ability.

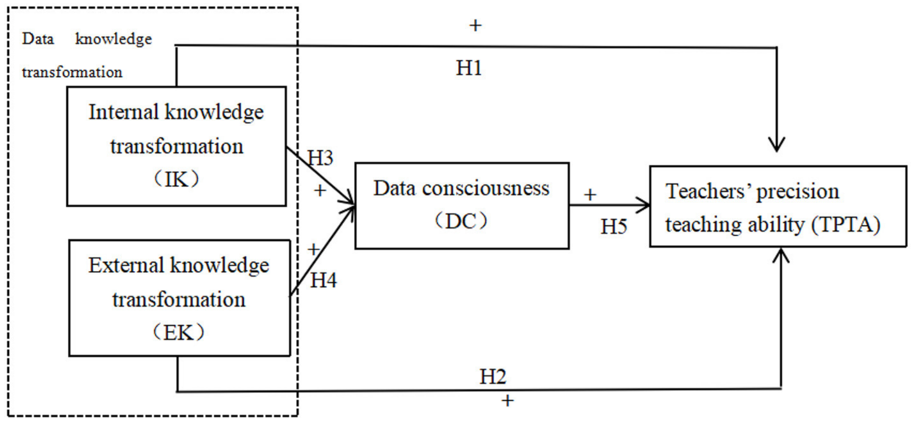


FIGURE 1
A theoretical model of data knowledge transformation, data awareness, and teachers' precision teaching ability.

H6d: Data consciousness plays a mediating role between internal knowledge and instructional evaluation ability.
H6e: Data consciousness plays a mediating role between external knowledge and student analysis ability.
H6f: Data consciousness plays a mediating role between external knowledge and teaching management ability.
H6g: Data consciousness plays a mediating role between external knowledge and educational decision-making ability.
H6h: Data consciousness plays a mediating role between external knowledge and instructional evaluation ability.

On the basis of related research and the above theoretical discussion, we propose the following theoretical model (Figure 1).

Research method

The survey instrument

The scales that measure data consciousness were developed by the research team based on a series of publications by Chinese scholars. The scales comprise three dimensions, which were informed by the theoretical assumptions and research results of the extant literature. Moreover, on the basis of the existing research dimensions, the measures provide response options on a five-point Likert scale, as shown in Table 1.

After developing the scale, the measures were repeatedly revised and discussed to ensure understanding, accuracy, and the representativeness of specific items. Ten front-line teachers who have participated in the knowledge training for the precision teaching system were invited to conduct a pre-test of the questionnaire to check question wording and the order of specific items. Case descriptions were added to the questions about the dimensions of teachers' precision teaching ability to help respondents understand the specific aspects of their ability. In addition, before the questionnaire was administered, we confirmed that the teachers participating in the survey understood the differences between internal knowledge transformation and external knowledge transformation. We also added instructions to the questionnaire guidelines.

Participants

The participants in this study are front-line teachers who have participated in smart education, precision teaching projects, competitions, training, and other activities. Respondents are from Shanghai, Anhui, Zhejiang, Jiangsu, Guangdong, and other provinces. The survey was administered from June 2021 to August 2021. First, 98 questionnaires were distributed (and all were recovered; recovery and effective rates were both 100%) with the help of the characteristics of the school and its major, and with the help of the interpersonal and working relationships of graduates and teachers. Second, we entrusted the principals of certain cooperative schools to distribute 150 questionnaires (of which 147 were recovered); the recovery rate is 98% and the effective rate is 96.7%. In total, 243 valid questionnaires were collected from teachers at 45 schools, including 6-year primary schools, 3-year junior high schools, 3-year high schools, 9-year consistent schools, and 6-year full middle schools. Among them, more than 60% of the respondents had more than 10 years of work experience, and 73% of the respondents had participated in the training and carried out precision teaching practice.

TABLE 1 Reference source and items of measuring scale.

Scale dimension			Item	References
First-level dimensions	Second-level dimensions	Third-level dimensions		
Internal knowledge transformation		Data positioning	4	Zhang and Li, 2015
		Data acquisition		
		Data analysis		
		Data interpretation		
External knowledge transformation		Find problems	4	Ruan and Zheng, 2016
		Make educational decision		
		Monitor teaching development		
		Reflect and evaluate		
Teachers' precision teaching ability	Student analysis ability	Sensitivity	4	Chen, 2012
		Reason		
		Thinking		
		Culture		
Teachers' precision teaching ability	Teaching management ability	Ability to manage teachers	3	Liu, 2006
		Ability to manage students		
		Ability to manage teaching environment		
	Educational decision-making ability	Teaching plan	3	Song and Li, 2011
		Teaching interaction		
		Teaching behavior		
	Instructional evaluation ability	Evaluation method	3	Zhao, 2003
		Evaluation content		
		Evaluation function		
Data consciousness		Student analysis	4	Liu et al., 2021
		Teaching management		
		Educational decision-making		
		Teaching evaluation		

Analysis on reliability and validity of the survey instrument

We used IBM SPSS version 25 to analyze the reliability and validity of the questionnaire to ensure its accuracy.

Reliability analysis

To test the reliability of the generated items and the scale, Cronbach's α is used for detection; the test results are shown in Table 2. When values are >0.8, the scale is considered acceptable. The questionnaire is divided into four categories: internal knowledge transformation, external knowledge transformation, teachers' precision teaching ability, and data consciousness. The reliability of each sub-scale is 0.918, 0.926, 0.964, and 0.872, respectively. Because these values are higher than 0.8, the reliability is considered high. Therefore, reliability analysis can be conducted.

To ensure that the data obtained from the questionnaire are suitable for a factor analysis, the KMO (Kaiser-Meyer-Olkin) and Bartlett tests were conducted on the sub-scales and the total scale. When KMO values are larger than 0.7 and approach 1, the partial correlation between variables is strong, and factor analysis is more accurate. When the significance level of Bartlett's sphericity test is <0.05 and the correlation between variables is strong, factor analysis can be conducted. The test results are shown in Table 3. The KMO values of the sub-scales and the total scale were >0.7, and the p-values of the Bartlett test was <0.05; thus, this questionnaire is suitable for factor analysis (Liang et al., 2020).

TABLE 2 Scales and Cronbach's α values.

Scale	Cronbach's α
Internal knowledge transformation	0.918
External knowledge transformation	0.926
Teachers' precision teaching ability	0.964
Data consciousness	0.872

TABLE 3 KMO and Bartlett test.

Scale		Data knowledge transformation	Teachers' precision teaching ability	Data consciousness	Whole scale
KMO		0.942	0.947	0.768	0.959
Bartlett test of sphericity	Approximate chi square	1,860.923	3,084.682	681.121	8,075.281
	Degrees of freedom	28	66	10	406
	p-value	0.000	0.000	0.000	0.000

KMO, Kaiser-Meyer-Olkin.

Validity analysis

The study conducted a CFA (Confirmatory Factor Analysis) on seven factors and 24 analysis items. The effective sample capacity was 243, or 10 times more than the number of analysis items. This level is moderate, so the data can undergo convergence validity and discriminant validity analyses. The factor loading of items reflects the correlations between the factors and analysis items. When the standard load coefficient is >0.7, there is a strong correlation between factors and analysis items. As shown in Table 4, the factor loading coefficients of items are >0.7. The composite reliability (CR) and average variance extracted (AVE), can comprehensively characterize the convergent validity of the scale. Generally, when the CR value is >0.7 and the AVE is >0.5, the convergence validity of the scale is high. As shown in Table 4, the AVE values corresponding to the seven factors are all >0.5, and the CR values are all higher than 0.7, indicating that the analysis and data have good convergence validity.

CFA can be used to distinguish validity results. The diagonal lines in Table 4 are AVE square root values, and the other cells are correlation coefficients. The value of the square root of the AVE represents the convergence of factors, and the correlation coefficient represents the correlation between items. If the factor has strong convergence (i.e., obviously stronger than the absolute value of the correlation coefficient with other factors), it can indicate that it has discriminant validity. The results are shown in Table 5. The square root of the AVE of each factor is greater than the absolute value of the correlation coefficient of this factor and other factors; therefore, all factors have good discriminant validity.

Data analysis

We selected SmartPLS 3 as the test tool for the research model to verify its collinearity, path relationships, model interpretation ability, and prediction ability of the structural equation model, and to improve and modify the model according to the verification results.

Collinearity means that there is a high correlation between independent variables, which leads to unreliable prediction results of the model. When the model has serious collinearity, it will lead to a deviation in the statistical significance of the results. Generally, if the variance inflation factor (VIF) value of each variable is <10, then the variables of the model do not exhibit collinearity. The results of this test are shown in Table 6. The VIF of each variable is <10, which eliminates the possibility of multicollinearity.

Path analysis

Path analysis is a statistical method used to verify the causal relationship between multiple variables and its strength to obtain the accuracy and reliability of the causal model (Bernacki et al., 2015). The path analysis results for the variables obtained through the PLS Algorithm in SmartPLS 3 are shown in Figure 2. In the figure, IK refers to internal knowledge transformation, EK refers to external knowledge transformation, and DC refers to data consciousness. Teachers' precision teaching ability includes four potential variables: student analysis ability (SA), teaching management ability (TM), educational decision-making ability (ED), and instructional evaluation ability (IE).

The direct effects between factors are shown in Table 7. We find that H2b [External knowledge transformation has a positive impact on teaching management ability (EK $\rightarrow$ TM)] is not supported, while the other hypotheses are supported. The transformation of internal knowledge related to data has a significant positive impact on teachers' precision teaching ability and data consciousness, while the transformation of external knowledge related to teaching has no such impact on teachers' teaching management ability in precision teaching. Both internal knowledge transformation and external knowledge transformation have a significant positive impact on data consciousness, and data consciousness has a positive impact on the improvement of teachers' precision teaching ability.

The results of the indirect effect path relationship are shown in Table 8, which indicates that the indirect path relationship decisions of the factors are valid. Specifically, there is an indirect

TABLE 4 Results of confirmatory factor analysis.

Latent variable	Measurement items	Factor loading	CR	AVE
Internal knowledge transformation	IK1	0.850	0.918	0.736
	IK2	0.847		
	IK3	0.859		
	IK4	0.874		
External knowledge transformation	EK1	0.877	0.927	0.760
	EK2	0.876		
	EK3	0.851		
	EK4	0.882		
Student analysis ability	SA1	0.848	0.885	0.658
	SA2	0.857		
	SA3	0.777		
Teaching management ability	TM1	0.758	0.909	0.769
	TM2	0.916		
	TM3	0.819		
Educational decision-making ability	ED1	0.893	0.919	0.792
	ED2	0.880		
	ED3	0.908		
Instructional evaluation ability	IE1	0.882	0.883	0.717
	IE2	0.931		
	IE3	0.836		
Data consciousness	DC1	0.766	0.839	0.639
	DC2	0.886		
	DC3	0.85		
	DC4	0.639		

CR, composite reliability; AVE, average variance extracted; IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

TABLE 5 Distinguishing validity: Pearson correlations and AVE square root.

	IK	EK	SA	TM	ED	IE	DC
IK	0.858						
EK	0.791	0.871					
SA	0.783	0.766	0.812				
TM	0.728	0.779	0.608	0.876			
ED	0.699	0.618	0.587	0.595	0.890		
IE	0.578	0.549	0.582	0.599	0.501	0.839	
DC	0.764	0.715	0.732	0.644	0.549	0.671	0.799

AVE, average variance extracted on diagonal.

relationship between internal knowledge transformation, external knowledge transformation, data consciousness, and teachers' precision teaching ability. To illustrate the mediating effect of data consciousness, this effect is evaluated based on the indirect path relationship established in Table 8. The overall effect between the factors is calculated by the sum of the path coefficients of the direct effect and the indirect effect. However, the significance of the indirect path relationship cannot directly reflect the intermediary effect of the factors, and the variant allele fraction (VAF) of factors should be considered. Generally, if the VAF is <20%, the factor has no intermediary effect; if it is between 20 and 80%, it has

TABLE 6 Collinearity diagnostics.

Measurement items	VIF
IK1	4.318
IK2	3.927
IK3	4.467
IK4	5.476
EK1	4.826
EK2	4.830
EK3	4.613
EK4	5.965
SA1	4.865
SA2	4.673
SA3	3.708
TM1	3.791
TM2	4.750
TM3	3.465
ED1	4.578
ED2	6.803
ED3	7.007
IE1	6.665
IE2	7.761
IE3	4.436
DC1	3.065
DC2	3.981
DC3	3.622
DC4	5.781

VIF, variance inflation factor; IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

some intermediate effect; and if it is >80%, it has a complete mediating effect. The results for the overall effect and VAF are shown in Table 9.

Evaluation of model detection and prediction ability

Table 10 displays the results for the model's interpretation and prediction ability; the explanatory power of the model includes the explanatory effect of the exogenous variables and endogenous variables of the model. Among them, the interpretation of the effect of the exogenous variables on the endogenous variables is determined by f^2, and the effect of the endogenous variables is determined by the coefficient R^2. When f^2 is <0.15, the explanatory power of the exogenous variables is weak; when it is >0.15 and <0.35, it is moderate; when it is >0.35, it is strong. As shown in Table 10, the explanatory effect of the exogenous variables in this model is mainly between 0.15 and 0.35, with moderate explanatory power. In this study, the coefficient of determination is around 0.75; thus, the explanatory power of the endogenous variables is relatively strong.

Finally, the predictive ability of the model is measured by the value of Q^2 obtained by blind-folding. If the values of Q^2 are >0, the model has the ability to predict the endogenous variables; the larger the value, the stronger the predictive ability of the model. The values of Q^2 in this model are high, so the predictive ability is significant.

Based on the comprehensive analysis, the interpretation and prediction capabilities of this model are relatively significant.

Model updating

To verify the path relationship, the direct effect assumes that EK → TM fails the significance test, so it is deleted from the model. The modified model is shown in Figure 3.

In testing the intermediary effects, we assume that data consciousness is not an intermediary variable in the path: EK → DC → ED, EK → DC → SA, EK → DC → IE. On this basis, the mediating effects of the modified theoretical model were checked, as shown in Table 11.

Discussion

Most of the existing studies have concentrated on analyzing the definition of teachers' precision teaching ability or strategies to enhance them. However, the different roles that explicit and invisible data knowledge learning and transformation can play in enhancing teachers' precision teaching ability and the interrelationships between them remain to be clarified. Therefore, this study constructs a mediating effect model between data knowledge transformation and teachers' precision teaching competencies based on knowledge management theory.

The results of this study suggest that data knowledge learning and transformation play an important role in the improvement of teachers' precision teaching ability, and also reveal some important differences.

Regarding the impact of teachers' learning and transformation of data knowledge on improving their teachers' precision teaching ability (RQ1), our findings add to the literature. The learning and transformation of data knowledge can be divided into two categories: internal knowledge (related to data) and external knowledge (related to teaching). Among them, the former has a positive impact on teachers' precision teaching ability. While the latter has a positive effect on student analysis ability (SA), educational decision-making ability (ED), and instructional evaluation ability (IE), and has no significant positive correlation with the ability to manage teaching (TM).

The results of this study show that data consciousness plays a specific mediating effect between data knowledge transformation and teachers' precision teaching ability (RQ2).

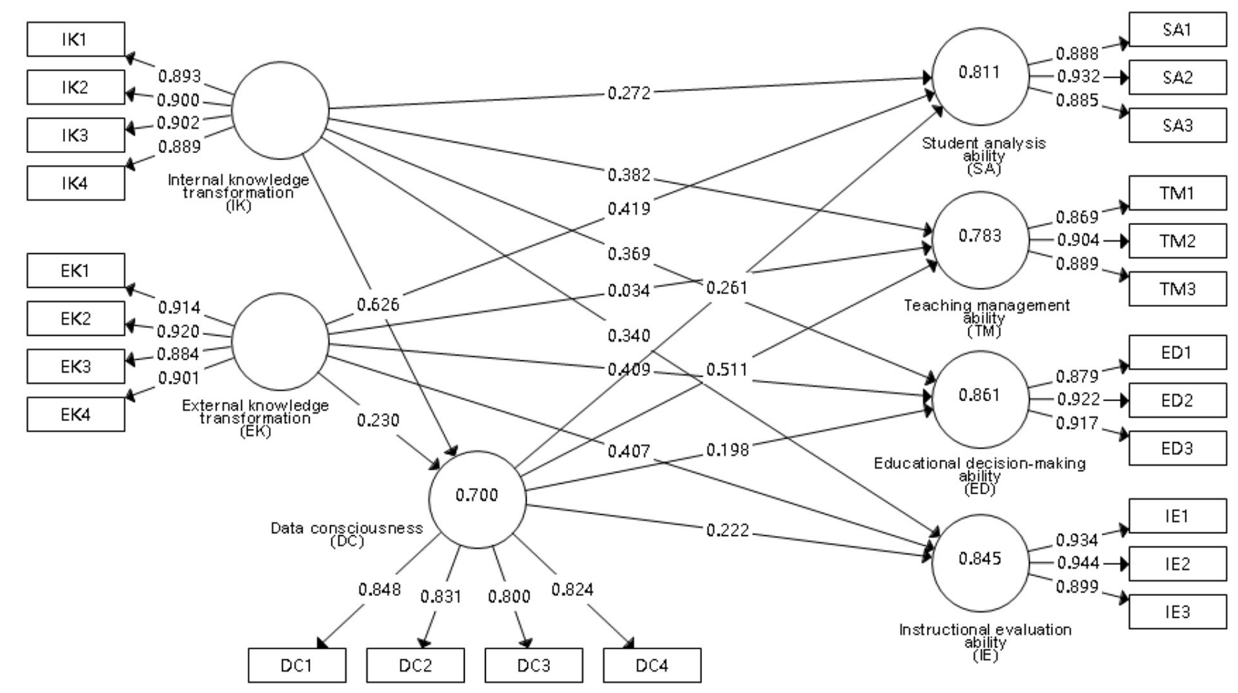


FIGURE 2
Path coefficient diagram.

TABLE 7 Verification of direct effect path relationships.

Assumption	Relationship	Path coefficient	*t*-value	*p*-value	Decision
H5c	DC → ED	0.198	3.950	0.000	Supported
H5d	DC → IE	0.222	3.897	0.000	Supported
H5a	DC → SA	0.261	4.277	0.000	Supported
H5b	DC → TM	0.511	7.264	0.000	Supported
H4	EK → DC	0.230	2.716	0.007	Supported
H2c	EK → ED	0.409	5.704	0.000	Supported
H2d	EK → IE	0.407	5.306	0.000	Supported
H2a	EK → SA	0.419	5.355	0.000	Supported
H2b	EK → TM	0.034	0.380	0.704	Unsupported
H3	IK → DC	0.626	7.562	0.000	Supported
H1c	IK → ED	0.369	4.366	0.000	Supported
H1d	IK → IE	0.340	3.880	0.000	Supported
H1a	IK → SA	0.272	3.007	0.003	Supported
H1b	IK → TM	0.382	3.995	0.000	Supported

IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

Among them, data consciousness partially mediates the effect between internal knowledge transformation (data-related) and teachers' precision teaching competencies. Between external knowledge transformation (teaching-related) and teachers' ability to teach accurately, data consciousness only fully mediated the effect between external knowledge transformation and instructional management ability.

In detail, the study found the following points. First, teachers' increased data consciousness relies on systematic learning of data knowledge. Teachers prefer activities

TABLE 8 Verification of the indirect path relationships.

Relationship	Path coefficient	*t*-value	*p*-value	Decision
EK → DC → TM	0.117	2.368	0.018	Supported
IK → DC → SA	0.163	3.401	0.001	Supported
EK → DC → ED	0.045	2.275	0.023	Supported
IK → DC → ED	0.124	3.149	0.002	Supported
EK → DC → SA	0.060	2.291	0.022	Supported
IK → DC → TM	0.320	5.187	0.000	Supported
EK → DC → IE	0.051	2.032	0.043	Supported
IK → DC → IE	0.139	3.389	0.001	Supported

IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

TABLE 9 Evaluation of the mediation effect.

Relationship	Direct effect	Indirect effect	Overall effect	VAF	Intermediary effect
EK → TM	0.034	0.117	0.151	77.5%	Part
IK → SA	0.272	0.163	0.435	37.5%	Part
EK → ED	0.409	0.045	0.454	9.9%	Null
IK → ED	0.369	0.124	0.493	25.2%	Part
EK → SA	0.419	0.060	0.479	12.5%	Null
IK → TM	0.382	0.320	0.702	45.6%	Part
EK → IE	0.407	0.051	0.458	11.1%	Null
IK → IE	0.340	0.139	0.479	29.0%	Part

VAF, variant allele fraction; IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

such as smart education and precision teaching that revolve around knowledge of data collection, statistics, and analysis.

Second, teachers have a high level of recognition of data-related knowledge. For teachers, the learning and transformation of internal knowledge can effectively enhance their data consciousness, thus improving teachers' precision teaching ability.

Third, teachers pay less attention to training on teaching-related knowledge. Teachers generally believe that learning analysis, instructional decision making and instructional evaluation in data-driven teaching process are mainly influenced by educational experience, while data consciousness cannot be used as a mediating factor.

Finally, the group of instructional managers among teachers perceived data consciousness as having a significant mediating effect. They believe that both learning related to internal knowledge (data-related) and external knowledge (teaching-related) are needed at the instructional management level. Both of them work together to effectively enhance data consciousness and ultimately improve precision instructional management skills.

Conclusion and implications

The findings of this study are of significant practical value as it provides indications for future targeted training of teachers in precision teaching competencies.

We need to understand the data confusion of teachers and focus on the development of teachers' precision teaching ability. At this stage, teachers' data confusion remains an important issue limiting the development of teachers' precision teaching ability. The research results show that teachers' confusion about the intrinsic knowledge of data (related to data) is much greater than the applied knowledge confusion of extrinsic knowledge (related to teaching). Teachers have many problems with the knowledge of data acquisition, statistics, organization, and analysis.

In this way, the activities focusing on the development of teachers' precision teaching ability should slow down the process of the activities. Before the activity is carried out, we should effectively understand the data confusion of the learning targets, focus on solving teachers' data problems first, and improve teachers' understanding of data value, data consciousness, and precision teaching. Learning about data as a way to enhance teachers' data confidence and reduce their sense of burden in precision teaching. Whether the learning activities related to precision teaching are organized by education departments or schools, the learning topics should be tailored to teachers' actual data confusion and provide appropriate learning materials.

At this stage, teachers' data confusion remains an important issue that restricts the development of teachers' precision teaching ability. Teachers as the executors of precision teaching, the implementation effect of precision teaching mainly depends on teachers' ability to understand, analyze and apply data. Whether the learning activities related to precision teaching are organized by education departments or schools, the learning topics should be tailored to teachers' actual data confusion and provide appropriate learning materials.

We must emphasize the development of data consciousness and optimize data-related knowledge translation mechanisms. The results of the study found that most teachers believe that data-related knowledge transformation can promote the development of data consciousness, which can ultimately enhance their own data literacy. The results of the study found that most teachers believe that data-related knowledge transformation can promote the development of data consciousness, which can ultimately enhance teachers' precision teaching ability.

Training activities should enhance the process of developing from technology to consciousness as teachers engage in internal knowledge of data-related platform development and setup, reading and computing, software use and processing, and data security. It should not be limited to text-based learning,

TABLE 10 Verification of the explanatory and predictive abilities of the model.

Relationship	f^2	R^2	Q^2
DC → ED	0.084	0.861	0.847
EK → ED	0.240		
IK → ED	0.159		
DC → IE	0.095	0.845	0.827
EK → IE	0.211		
IK → IE	0.121		
DC → SA	0.108	0.811	0.786
EK → SA	0.184		
IK → SA	0.063		
DC → TM	0.360	0.783	0.698
EK → TM	0.001		
IK → TM	0.109		
EK → DC	0.036	0.700	0.694
IK → DC	0.268		

IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

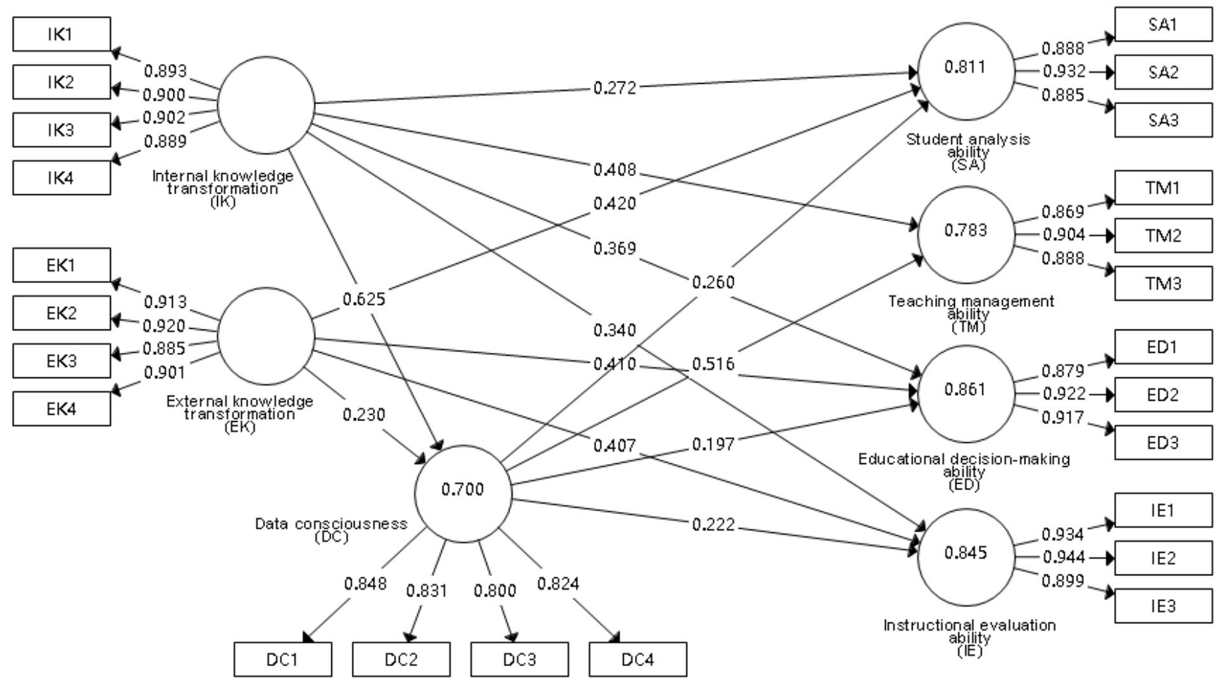


FIGURE 3
Revised theoretical model and path coefficient.

TABLE 11 Revised mediation effect verification.

Assumption	Relationship	VAF	Intermediary effect	Decision
H6a	IK → DC → SA	27%	Part	Supported
H6b	IK → DC → TM	44%	Part	Supported
H6c	IK → DC → ED	20%	Part	Supported
H6d	IK → DC → IE	23%	Part	Supported
H6f	EK → DC → TM	100%	Complete	Supported

VAF, variant allele fraction; IK, internal knowledge; EK, external knowledge; SA, student analysis ability; TM, teaching management ability; ED, educational decision-making ability; IE, instructional evaluation ability; DC, data consciousness.

but should focus on teachers' operational learning. Teachers should have a practical experience of data reliability, ease of operation, and its importance to teaching and learning during the training activities.

The learning of data technology knowledge can deepen teachers' understanding of data, reduce their sense of powerlessness and rejection of data, and promote the development of teachers' data consciousness. It allows teachers to understand the goal, process, and meaning of precision teaching at the cognitive level, thus enhancing their initiative to learn and apply data knowledge and optimize their transformation process.

In the training, we want to create realistic scenarios for application and focus on knowledge transformation mechanisms for data-driven teaching and learning processes. A large number of real teaching cases should be provided for the participants in the teachers' participation in data knowledge training, and virtual courses for accurate teaching and learning. To cope with the transformation of external knowledge (related to teaching), teachers not only need to be familiar with mastering data-related knowledge, but also need a lot of teaching and learning knowledge and experience in order to complete the transfer and transformation of knowledge. Create real scenario application, guide teachers to experience the process of data collection, collation, analysis, evaluation and prediction in a real environment, and feel the process and meaning of application in the process of learning. This is very essential.

Finally, we need to develop strategies for the development of group roles. Because different group roles require different focuses on teachers' precision teaching ability. Training should be designed and implemented to respond to different needs.

Limitations and future research

Of course, there are still some shortcomings in this study, such as the limited sample size, insufficient diversity in the samples, the influence of subjective factors in the questionnaire design process, and the lack of qualitative data, such as observations and interviews. In subsequent research, we should increase the proportion of qualitative data collected and expand the sample size appropriately.

In addition, it is important to suggest future research directions for those interested in the unifying theme of this study. In the course of our research, we found that there are multiple groups of teachers, including teaching, management, and administration, and that the needs and perceptions of different groups differ significantly in terms of precision teaching competencies. It would be helpful if future research could delve into this study with specific categories of teacher groups.

Data availability statement

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

Ethics statement

Ethical review and approval was not required for the study involving human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants in accordance with the national legislation and the institutional requirements.

Author contributions

JT: project administration and data curation. CF: methodology and writing review and editing. WW: analysis data and perfect the thesis. YZ: revise the thesis and translation. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the State Key Laboratory of Cognitive Intelligence 2022 Smart Education Open

Project: Practical Research on Big Data Driven Precision Teaching (iED2022-007) and Key Research Project of Anhui Provincial Quality Engineering Teaching Research: Research on the Construction of High-Quality Teacher Education Talent Training System Based on Output Orientation (2021jyxm1437).

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

Bernacki, M. L., Nokes-Malach, T. J., and Aleven, V. (2015). Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. *Metacogn. Learn.* 10, 99–117. doi: 10.1007/s11409-014-9127-x

Cabero, A. J., Guillén, G. F. D., Ruiz, P. J., and Palacios, R. A. (2022). Teachers' digital competence to assist students with functional diversity: identification of factors through logistic regression methods. *Br. J. Educ. Technol.* 53, 41–57. doi: 10.1111/bjet.13151

Chen, L. (2012). The reconstruction of Chinese class from the view of learning—based on the case study of learning situation analysis by Chinese teachers. *Curric. Teach. Mater. Method*. 4, 42–48. doi: 10.3969/j.issn.1000-419X.2012.12.004

Cui, Y., and Zhang, H. (2022). Integrating teacher data literacy with TPACK: a self-report study based on a novel framework for teachers' professional development. *Front. Psychol.* 13, 966575. doi: 10.3389/fpsyg.2022.966575

Du, Z., and Su, J. (2021). Analysis of the practice path of the flipped classroom model assisted by big data in english teaching. *Sci. Program*. 2021, 1–12. doi: 10.1155/2021/1831892

Duraiswami, R. A., and Krishnamurthy, P. (2020). Report on the teacher training workshop in earth science education, UGC-HRDC, Savitribai Phule University of Pune (SPPU), Pune, India. *J. Geol. Soc. India.* 96, 420–421. doi: 10.1007/s12594-020-1571-4

Filderman, M. J., Toste, J. R., Didion, L., and Peng, P. (2022). Data literacy training for K—12 teachers: a meta-analysis of the effects on teacher outcomes. *Remed. Special Educ.* 43, 328–343. doi: 10.1177/07419325211054208

Fletcher, T., Mandigo, J., and Kosnik, C. (2013). Elementary classroom teachers and physical education: change in teacher-related factors during pre-service teacher education. *Phys. Educ. Sport Pedagogy*. 18, 169–183. doi: 10.1080/17408989.2011.649723

Friedman, P., Berg, S. V., and Duncan, J. (1979). External vs. internal knowledge acquisition: joint venture activity and RandD intensity. *J. Econ. Bus.* 31, 103.

Haigh, A. (2008). Teaching with precision. *TES Magazine* 4805, 51. doi: 10.12968/sece.2008.10.1306

Hayes, B., Heather, A., Jones, D., and Clarke, C. (2018). Overcoming barriers to using precision teaching with a web-based programme. *Educ. Psychol. Pract.* 34, 166–174. doi: 10.1080/02667363.2018.1433129

Kabir, N., and Carayannis, E. (2013). Big data, tacit knowledge and organizational competitiveness. *J. Intellig. Stud. Bus.* 3, 54–62. doi: 10.37380/jisib.v3i3.76

Kubina, J. R. M., and Yurich, K. K. L. (2009). Developing behavioral fluency for students with autism: a guide for parents and teachers. *Interv. Sch. Clin.* 44, 131–138. doi: 10.1177/1053451208326054

Liang, Y., Ge, X., Yue, F., and Cao, B. (2020). The construction of the professional satisfaction index model of the western industry characteristic universities–taking five universities in Chongqing as an example. *Educ. Teach. Forum*. 11, 78–80. Available online at: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=JYJU202011035&DbName=CJFQ2020

Liu, G., Chen, Y., and Qin, Y.,(2021). Training mode and teachers' data literacy. *J. Teach. Educ*. 3, 21–30. doi: 10.1186/s12909-021-02993-1

Liu, Q. (2006). *The exploring research of the middle school teachers' classroom teaching management ability*. (Master's thesis). [Chongqing, China]: Southwest University. Available online at: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=2006118372.nh&DbName=CMFD2006

Lyons, D. (1984). A computer aided implementation of precision teaching. *Computer Assisted Instruction*. 27. Available online at: https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED249676&lang=zh-cn&site=ehost-live

Mandinach, E. (2012). A perfect time for data use: using data-driven decision making to inform practice. *Educ. Psychol.* 47, 71–85. doi: 10.1080/00461520.2012.667064

Mandinach, E. B., and Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educ. Res.* 42, 30–37. doi: 10.3102/0013189X12459803

McDowall, A., Mills, C., Cawte, K., and Miller, J. (2021). Data use as the heart of data literacy: an exploration of pre-service teachers' data literacy practices in a teaching performance assessment. *Asia-Pacific J. Teach. Educ.* 49, 487–502. doi: 10.1080/1359866X.2020.1777529

Ministry of Education of the People's Republic of China (2012). *Notice of the Ministry of Education on Printing and Distributing the Professional Standards for Kindergarten Teachers (trial), the Professional Standards for Primary School Teachers (trial) and the Professional Standards for Middle School Teachers (trial).* Available online at: http://www.moe.gov.cn/srcsite/A10/s6991/201209/t20120913_145603.html (accessed September 13, 2012).

Ministry of Education of the People's Republic of China (2021). *Guidance of the Ministry of Education and Other Six Departments on Promoting the Construction of New Educational Infrastructure and Building a High-Quality Education Support System.* Available online at: https://hudong.moe.gov.cn/srcsite/A16/s3342/202107/t20210720_545783.html (accessed July 8, 2021).

Peng, H., and Zhu, Z. (2018). Human–machine collaborated data wisdom mechanisms: the alchemy of data value for smart education. *Open Educ. Res.* 2, 41–50. doi: 10.13966/j.cnki.kfjyyj.2018.02.005

Pierce, R., Chick, H., Watson, J., Les, M., and Dalton, M. (2014). A statistical literacy hierarchy for interpreting educational system data. *Aust. J. Educ.* 58, 195–217. doi: 10.1177/0004944114530067

Revuelta-Domínguez, F.-I., Guerra-Antequera, J., González-Pérez, A., Pedrera-Rodríguez, M.-I., and González-Fernández, A. (2022). Digital teaching competence: a systematic review. *Sustainability* 14, 6428. doi: 10.3390/su14116428

Ruan, S., and Zheng, Y. (2016). Components, functions and development strategies of teachers' data literacy. *Modern Distance Educ.* 1, 60–65. doi: 10.13927/j.cnki.yuan.2016.0008

Song, D., and Li, S. (2011). Investigation and analysis on the present conditions of teachers teaching planing decision-making. *Teach. Educ. Res.* 4, 51–56. doi: 10.13445/j.cnki.t.e.r.2011.04.004

Srikoom, W., Faikhamta, C., and Hanuscin, D. L. (2018). Dimensions of effective STEM integrated teaching practice. *K-12 STEM Educ.* 4, 313–330. doi: 10.1111/ssm.12199

Theall, M. (2001). Can we put precision into practice? Commentary and thoughts engendered by Abrami's "improving judgments about teaching effectiveness using teacher rating forms." *New Dir. Inst. Res.* 109, 89. doi: 10.1002/ir.5

Wang, T.-I., Su, C.-Y., and Hsieh, T.-C. (2011). Accumulating and visualising tacit knowledge of teachers on educational assessments. *Comput. Educ.* 57, 2212–2223. doi: 10.1016/j.compedu.2011.06.018

Weinberger, A. B., and Green, A. E. (2022). Dynamic development of intuitions and explicit knowledge during implicit learning. *Cognition*. 222, 105008. doi: 10.1016/j.cognition.2021.105008

Wieser, C. (2016). Teaching and personal educational knowledge – conceptual considerations for research on knowledge transformation. *Eur. J. Teach. Educ.* 39, 588–601. doi: 10.1080/02619768.2016.1253673

Williamson, R. (2012). *A Study of Effective Teachers' Knowledge and Use of Intersectional Data to Improve Student Academic Performance [ProQuest LLC]*. Memphis: ProQuest LLC.

Zhang, J., and Li, B. (2015). The connotation, value and development path of teachers' data literacy in the context of big data. *e-Educ. Res.* 7, 14–19. doi: 10.13811/j.cnki.eer.2015.07.002

Zhao, M. (2003). Research on classroom instructional evaluation. *Edu. Res. Exper*. 2, 60–64. Available online at: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=YJSY200302013&DbName=CJFQ2003

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 December 2022
DOI 10.3389/fpsyg.2022.1066164

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University, China

REVIEWED BY
Tao Yang,
University of Birmingham,
United Kingdom
Liao Mengyi,
Central China Normal University, China
Xunzhou Ma,
Southwest Minzu University, China

*CORRESPONDENCE
Changcheng Wu
wucc@sicnu.edu.cn

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

RECEIVED 10 October 2022
ACCEPTED 24 November 2022
PUBLISHED 12 December 2022

CITATION
Mou Y, Jing B, Li Y, Fang N and Wu C (2022) Interactivity in learning instructional videos: Sending danmaku improved parasocial interaction but reduced learning performance.
Front. Psychol. 13:1066164.
doi: 10.3389/fpsyg.2022.1066164

Interactivity in learning instructional videos: Sending danmaku improved parasocial interaction but reduced learning performance

Ya Mou[1,2], Bin Jing[3], Yichun Li[3], Nanyang Fang[4] and Changcheng Wu[3]*

[1]Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, Hubei, China, [2]Shuang Liu Middle School, Chengdu, Sichuan, China, [3]School of Computer Science, Sichuan Normal University, Chengdu, Sichuan, China, [4]Qingtai Mountain Middle School Attached to Sichuan Normal University, Chengdu, Sichuan, China

Introduction: The instructional video is considered to be one of the most distinct and effective virtual learning tools. However, one of its biggest drawbacks is the lack of social interaction that occurs. This study tested the impact of participants sending zero danmaku (sending messages on the screen), three danmaku sending, and unlimited danmaku as an instructional video plays on learning performance.

Methods: We assessed learners' retention and transfer scores, as well as self-report scores for cognitive load and parasocial interaction. This study sample comprised 104 participants who were randomly assigned to learn from one of three instructional videos on the topic of the heart.

Results: The results showed that sending danmaku improved learners' parasocial interaction, while significantly increasing their cognitive load and also hindering learning performance. The observed increase in cognitive load reported by learners was also caused by increased levels of parasocial interaction.

Discussion: Our findings suggest that by sending danmaku, learners can promote interactive learning, but that this has a negative impact on learning performance and the process of video learning.

KEYWORDS

instructional video, danmaku, learning performance, cognitive load, parasocial interaction

Introduction

The instructional video is considered to be one of the most distinct and effective virtual learning tools as it provides a sensory learning environment that has been shown to support learners to understand more and help them better recall knowledge (Palaigeorgiou and Papadopoulou, 2019). Furthermore, instructional videos also appear to have a significant effect on improving learners' learning performance, engagement, interest, motivation to learn new subjects, and autonomy (Giannakos et al., 2015). However, instructional videos are not a panacea. Linear instructional videos run the risk of becoming a passive, television-like experience for learners. This type of experience can lead to superficial learning or inadequacy of the overall learning performance, also known as the "couch-potato" phenomenon (Van den Bulck, 2000; Palaigeorgiou and Papadopoulou, 2019). One of the most notable weaknesses of instructional videos is that learners cannot interact with it, so instructional videos often lack cognitive appeal and challenges (Palaigeorgiou and Papadopoulou, 2019). This interaction has been considered to be an important factor in improving social presence and learners' attention (Rice and Redcay, 2016; Yang et al., 2019). In recent years, the production technology of instructional videos has grown rapidly, enabling more interactive functions, such as online forums or real-time in-video text messages, known as danmaku. Danmaku, as an emerging video interaction technology, allows users' comments to be posted and displayed instantly (anonymously) on the video screen, thus becoming an integral part of the video; every time the video is played, it also plays all past comments (Zhou and Zhou, 2022). Danmaku were initially popular in China, Japan, and other Asian countries, and became sought after and loved by the majority of video-based learners (Teng and Chan, 2022). Especially in China, when learners are viewing a video, the playback platform typically has the "danmaku display" function turned on by default, allowing people to interact with other users while watching in text form (Bai et al., 2021). While such interactions are commonplace, little is known about the effects of seeing others' danmaku in such videos on users' attention and experience.

Most existing research on danmaku has been from the perspective of "learners seeing danmaku sent by others on the screen" (e.g., Pi et al., 2020), focusing on the potential value of the danmaku itself and their impact on learners' attention as they watch the video lectures (Bai et al., 2019; Pi et al., 2020). For instance, learners' interaction with danmaku-enabled videos can not only provide immediate feedback to video creators on how they might improve the quality of the video, but they also greatly enhance the viewing experience of online learners (Bai et al., 2021). Zhang et al. (2019) discussed the potential of danmaku and proposed to include them in their instructional videos, whether the content is synchronous or asynchronous. With danmaku, even an asynchronous experience can seem like a synchronous experience. It is a potentially useful tool for use in online learning as learners feel they are learning alongside peers and teachers (Zhang et al., 2019). However, Pi et al. (2020) showed through empirical research that seeing others' danmaku during instructional videos had a negative impact on learners' attention and learning.

Compared to seeing others' danmaku onscreen, the biggest difference for learners in sending their own danmaku is interactivity (Ma and Cao, 2017; Pi et al., 2020). Learners communicate synchronously or asynchronously with others by sending danmaku. Thus, learners can have a "pseudo-synchronic" communication and viewing experience by replying to others' danmaku comments, creating a sense of a "live" experience through the sense of shared virtual time between learners and learners, as well as between learners and instructors (Johnson, 2013). This "pseudo-synchronicity" not only changes the way learners experience instructional videos from a passive and isolated experience to a more active and social experience, but it also provides them with a sense of a shared and synchronous experience. Research has shown that this type of interaction appears to give learners an instant experience of immersion, similar to if they were having a conversation directly with other learners, experiencing a "mental sensation of engagement," also known as parasocial interaction (Shin, 2019; Lim et al., 2020; Yang et al., 2022). To our knowledge, few studies have explored how learners actively sending danmaku affects their learning process and performance. Theoretically, the benefits of sending danmaku during the learning experience are consistent with the theory of social constructivism (Vygotsky and Cole, 1978). This theory proposes that learners are not only passive receivers of knowledge and information, but also active generators and constructors of knowledge. The creation of knowledge requires active interaction in a social environment. Social interaction must then exist prior to the process of learners internalizing knowledge, and as such, social interaction is essential for the creation of new knowledge, and interaction with others is conducive to learners' learning in the zone of proximal development (Vygotsky and Cole, 1978). Specifically, as a tool to reflect social interaction in the learning process, danmaku-sending provides scaffolding for learners and helps them to reach their potential level of development. The basis of social interaction is the exchange of messages (Vygotsky and Cole, 1978; Pi et al., 2019). Numerous studies on instructional video-based learning have shown that learners benefit from exchanging messages regarding the information shared through the videos (Li et al., 2014; Hung and Chen, 2018). Real-time exchange of messages in the classroom when learners are able to pause videos has been shown to have a positive impact on attention and learning performance (Weisz et al., 2007; Li et al., 2014). When learners send danmaku, they also exchange messages within the videos, which we infer would promote deeper processing and internalization of knowledge.

Although the benefit of exchanging messages has been demonstrated under particular conditions, not all instructional video formats have yet been tested (Pi et al., 2020). In previous studies, learners sent messages after learning from instructional videos, rather than exchanging messages simultaneously during learning (i.e., seeing and sending danmaku; Weisz et al., 2007; Li et al., 2014). Unfortunately, in a danmaku-enabled video, the real-time exchange of messages does appear to negatively impact learning (e.g., Pi et al., 2020). Several reasons have been suggested for this. First, learners must pay attention to danmaku content while also following and understanding the learning content. This phenomenon tends to make learners lose themselves in receiving and responding to danmaku comments, thus neglecting to absorb and process the learning content itself, resulting in excessive cognitive load (Yang et al., 2019). This situation is likely to trigger a split-attention effect, in which the learner must mentally integrate information from multiple sources simultaneously (Ayres and Sweller, 2005). Second, the danmaku sent by fellow learners may also have nothing to do with video learning knowledge, and may also be negative. This also increases learners' cognitive load and is not conducive to learning (Yang et al., 2019). Furthermore, an earlier study showed that creativity is inhibited if learners must pay attention to constant interactions with others (Diehl and Stroebe, 1987). In the case of instructional videos, this would apply to learners sending danmaku (Herrmann et al., 2013; Pi et al., 2020).

To sum up, in previous studies, learners have generally only played the role of a "bystander" in the exploration of the effects of danmaku during learning. Existing research has suggested that danmaku can be helpful in instructional videos (Yang et al., 2019), but it is important to understand how to further improve the quality of instructional video learning. Therefore, the current study investigated whether the sending and not sending of danmaku brings about differences in learning from instructional videos, and whether differences in the number of danmaku sent might optimize the effectiveness of video learning in terms of learning performance (i.e., retention and transfer), self-reported cognitive load (i.e., intrinsic, extraneous, and germane cognitive loads), and parasocial interaction after the instructional video is over.

Participants were asked to watch instructional videos on the subject of the heart. In the control condition, learners only watched the instructional videos and were not able to send danmaku (No Danmaku-Sending group; NDS). Under the experimental conditions, learners viewed the instructional video, with the difference between the experimental conditions being only the number of danmaku sent by the learners themselves, specifically: (1) the number of danmaku sent during the video was set at three (Three Danmaku Sent group; TDS), and (2) the number of danmaku sent was unlimited, whereby the learner was required to send more than three danmaku during the video (Unlimited Danmaku Sent group; UDS). For the TDS group, we found through the pre-experiment test, when the number of danmaku learners could send was allowed to be unlimited, that participants would send an average of three danmaku during the process of learning from a video. In the actual experiment, each participant in the TDS group sent the same number of danmaku (i.e., three), which facilitated scientific observation and analysis by fixing the experimental conditions.

We wanted to verify whether learning was affected by the number of danmaku sent. Therefore, based on the above literature review, this study proposed the following hypotheses:

> **Hypothesis 1:** Among the three instructional video conditions, learners in the NDS group will have the best retention and transfer scores, followed by the TDS group, and finally by the UDS group.

> **Hypothesis 2:** In the UDS condition, learners will report having the highest cognitive load (including intrinsic, extraneous, and germane cognitive load), followed by those in the TDS condition, and finally by those in the NDS condition.

> **Hypothesis 3:** Learning from instructional video under the UDS condition, learners will report the highest level of parasocial interaction, followed by learners in the TDS condition, and finally by learners in the NDS condition.

> **Hypothesis 4:** The number of danmaku sent by learners will be negatively associated with learning performance, and this relationship is mediated by parasocial interaction and cognitive load.

Materials and methods

Participants and design

The participants in the current study were 104 grade 10 high school students (64 females; age: M = 15.45, SD = 0.54) recruited from one high school in Chengdu, Sichuan, China, and were selected randomly from four classes. None of the participants had been taught in the school setting about the heart. The students were informed about the purpose and safeguards of the experiment and written informed consent was obtained prior to participation. All participants were native Chinese and naive to this study. The study was approved by the ethics committee of the school where the study was conducted, and the study followed ethical standards of care for human subjects.

A one-factor between-subjects design was adopted in the present study with students assigned randomly to one of the three groups: 35 students (22 females; age: M = 15.43, SD = 0.56) to the NDS group, 35 students (24 females; age: M = 15.46, SD = 0.56) to the TDS group, and 34 students (18 females; age: M = 15.47, SD = 0.51) to the UDS group. There were no significant differences between these three groups in mean age [F(2,101) = 0.05, p = 0.947], their mean pre-test scores

[$F(2,101) = 0.15$, $p = 0.858$], or on the proportion of males and females, $\chi^2(2) = 1.82$, $p = 0.403$.

Materials

The instructional video ran for 6 min and 18 s, and the content was on "Heart and blood circulation." We used PowerPoint and Adobe Premiere to create the video. The instructional video consisted of three parts: danmaku, learning content, and voice. Regarding the danmaku, when we were making the video, we use the "four classes of communicative approach" to design four kinds of questions (Scott et al., 2006). These questions were used to create the pseudo-danmaku that would be used to answer the questions. In this experiment, we designed a total of 30 pseudo-danmaku. These danmaku ran from the right of the screen to the left at a consistent speed, positioned in the top fifth of the screen above the video and scrolling past until they disappeared. The questions were also intended to guide learners to send their own danmaku in response to the questions raised in the videos. Some of the questions used are shown in **Table 1**. Regarding the learning content, the instructional video was made up of 12 pages of slides, with the key instructing content presented in the form of text, accompanied by an illustration to help learners understand the content. Finally, regarding the voice, the audio content was recorded by a male instructor with a neutral Chinese accent. All materials went through multiple rounds of screening and revision by two master students before they were combined to create the final instructional video.

The video was shown using software called "Dandan Play."[1] This software allows learners to send danmaku in the instructional videos over a local area network without having to register and account. The software also supports changing the color, size, and screen location of the danmaku. Using intelligent identification of video content, past danmaku saved on the server can be synchronized to appear in the video on the local computer to enhance the viewing experience. The danmaku are shown in the top fifth of the screen, while the rest of the screen is taken up by the video content (see **Figure 1**).

1 www.dandanplay.com

Measurements

Pre-questionnaire

The pre-questionnaire consisted of three parts: first, demographic information (e.g., gender and age) was collected. Second, contextualizing danmaku in learners' experience, both viewing (e.g., "If there are danmaku in the instructional video, do you think the danmaku will improve or interfere with your learning ability? Why?") and sending (e.g., "If you send danmaku while learning from the instructional video, do you think this behavior will improve or interfere with your learning? Why?"). Finally, learners' general knowledge about the heart was also tested, using seven multiple-choice items. An example item was: "Which of the following is not part of the blood circulatory system? A. Blood; B. Blood vessel; C. Heart; D. Lung; E. I don't know." Each item had five choices, and only one was the correct answer. The fifth option was "I don't know" to avoid participants simply guessing the answer (Pi et al., 2022). Two points were awarded for each correct answer. The total score could range from 0 to 14. As there were relatively fewer items used in this study than in previous studies, the reliability analysis was measured using Spearman–Brown reliability (Kelley, 1925). The prior knowledge test showed moderate internal consistency: $r_{kk} = 0.63$ (LeBreton and Senter, 2008). The learners' age, gender, and prior knowledge scores were used to determine whether the groups differed in their basic characteristics. These variables did not have any significant differences among the three groups.

Learning performance

Tests were created to assess learning performance, in terms of both knowledge retention and transfer. The knowledge retention test evaluated learners' direct memory of knowledge after they had finished watching the learning instructional video. The retention test showed moderate internal consistency: $r_{kk} = 0.60$ (LeBreton and Senter, 2008). It contained 12 items with a possible total of 30 points. There were eight multiple-choice items with only one correct response (e.g., "Which of the following statements about the function of the human heart is false? A. The heart provides adequate blood flow to the body's organs and tissues; B. The heart supplies oxygen to the body; C. The heart provides various nutrients to the body; D. The heart can bring carbon dioxide, urea, and so on to the cells; E.

TABLE 1 Specific questions used to encourage social interaction: four classes of communicative approach.

Number	Type	Sample question
1	Dialogic-interactive	*Q:* Do you know which two types of cells the cardiac muscle contain? *A:* Yes, the cardiac muscle cells contain cardiomyocytes and special cardiomyocytes.
2	Non-dialogic-interactive	*Q:* Coronary veins can also provide blood for the heart. Do you know the difference between coronary veins and arteries? *A:* Yes, the coronary veins can also provide blood for the heart.
3	Dialogic-non-interactive	The heart is a complex organ, we should not only learn the different locations of its parts but also the function of each part. For example, this is the part that keeps the heart beating, and this is the part that supplies the blood.
4	Non-dialogic-non-interactive	The cardiac muscle is the part of the heart that we normally see, and it makes up the outline of our heart.

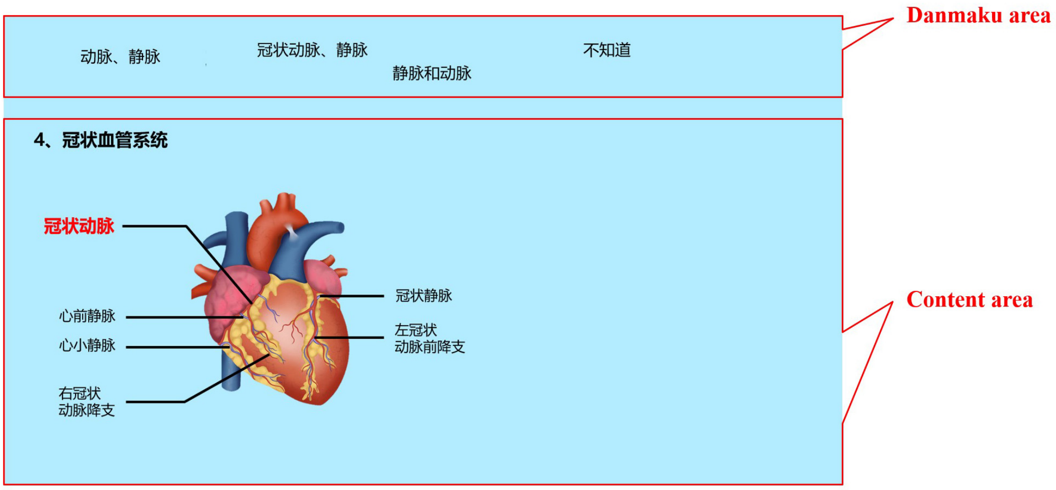


FIGURE 1
Danmaku and video content positioning on instructional video screen in Dandan Play player.

I don't know."), two multiple-choice items with several correct responses (e.g., "Which of the following statements about the heart are correct? A. The heart is positioned with the lungs on both sides; B. Above the heart is the diaphragm and below the heart are the large blood vessels that enter and exit the heart; C. The heart is located inside the middle mediastinum; D. The heart is located inside the posterior mediastinum; E. I don't know."), and two short-answer items regarding the coronary artery. Respondents received two points for each correct answer for the single-correct-choice items, three for the several-correct-choice items, and four points for short-answer items. The short-answer responses were scored separately by two raters and showed appropriate rater agreement ($rs > 0.60$, $ps < 0.001$, based on Pearson's correlation).

The knowledge transfer test evaluated learners' ability to use the presented material in novel situations after learning from the instructional videos. The transfer test had moderate internal consistency: $r_{kk} = 0.59$ (LeBreton and Senter, 2008). It contained eight multiple-choice items with one single correct response (e.g., "What happens when the coronary arteries in the human body harden? A. People get coronary heart disease; B. There is an embolism in the coronary arteries; C. There is a spasm in the coronary arteries; D. The cardiac muscle cannot get nutrients and oxygen; E. I don't know."), two multiple-choice items with several correct responses (e.g., "Which of the following symptoms of myocardial infarction are correct? A. Insufficient blood supply to the venous blood vessels; B. Decreased elasticity of the walls of the large arteries; C. A blockage in the terminal arteries of the heart; D. Necrosis in the area of coronary arteries; E. I don't know."), and a short-answer item about heart disease. The points for each item were set the same as those of the knowledge retention test (with eight points for the short-answer item) for a possible maximum of 30 points. Short-answer responses were scored separately by two raters and showed appropriate rater agreement ($r = 0.67$, $p < 0.001$, based on Pearson's correlation).

Cognitive load

The cognitive load questionnaire asked learners to rate three subjective items intended to measure their perceptions concerning their intrinsic cognitive load (i.e., "How much effort was required to learn the material?"), extraneous cognitive load (i.e., "How difficult was the material?"), and germane cognitive load (i.e., "Did you think the material was easy to learn?") during learning. The cognitive load items have been shown to be reliable and valid indications of cognitive load in previous studies (Moreno et al., 2001; Yung and Paas, 2015). The first two items were adapted from Craig and Schroeder (2017) and translated into Chinese by Yang (2014). The last item was adapted from Gerjets et al. (2009) and translated into Chinese by Zhao (2014). Learners rated each item on a nine-point scale ranging from 1 (very little) to 9 (very much).

Parasocial interaction

The parasocial interaction questionnaire was developed by Beege et al. (2017) and is used to test learners' perception of the conversational "give and take" with the instructor. It consists of seven items (e.g., "I feel like the instructor is giving me a lesson") and each item is rated using a seven-point Likert scale ranging from 1 (I do not agree at all) to 7 (I totally agree). The final score is the average of all item scores. To adapt the questionnaire for use in a Chinese context, all items were translated into Chinese and revised by a professional college English instructor (Pi et al., 2021). Cronbach's alpha for the questionnaire in the current study was 0.94.

Procedure

This study was conducted in a computer room and took approximately 45 min to complete. Before starting the experiment, each computer was checked to ensure that the headphones were plugged and the volume was adjusted to avoid interference from external conditions. The make and model of all computers and monitor screen sizes were identical. All computers were uniformly controlled by the researcher's aide, and the learners wore headphones to avoid interfering with one another. When the experiment began, learners were first assigned randomly to the NDS, TDS, or UDS group, and each individual was told they would perform the tasks individually. The learners were then asked to sit in front of the computer and to complete the pre-questionnaire at their own pace. Next, the learners watched the instructional video for their specific group, and afterward completed the retention and transfer tests without having to adhere to a time limit. Finally, they were asked to complete a post-questionnaire to measure their perception of the learning process. After the experiment, all learners were compensated in the form of small gifts (i.e., facial tissue and pen) for their participation to thank them for their time and effort.

Results

Figure 2 depicts the learners' screen setup when seeing danmaku in the instructional videos. As shown in Figure 2A, 82.69% of the learners indicated that they had paid attention to the danmaku sent by other learners while viewing the instructional video. This percentage was significantly higher than those who reported "not paying attention" (17.31%). Figure 2B shows that more than half of the learners (54.81%) reported that they would open the danmaku while the instructional video played, which was slightly higher than those who reported "closed the danmaku" (38.46%), and only 6.73% of the learners reported not paying any attention to the danmaku. Figure 2C shows that nearly half of the learners (45.19%) believed that the danmaku had no influence on their learning, while the proportion of learners who believed that the danmaku would improve their learning (31.73%) or that they would interfere with their learning (23.08%) were similar. The reasons suggested as to why the danmaku might improve their learning ability were that danmaku can supplement knowledge points, they present an opportunity for communication and discussion with others, and danmaku would expand the learner's own thinking abilities. However, the reasons why danmaku might

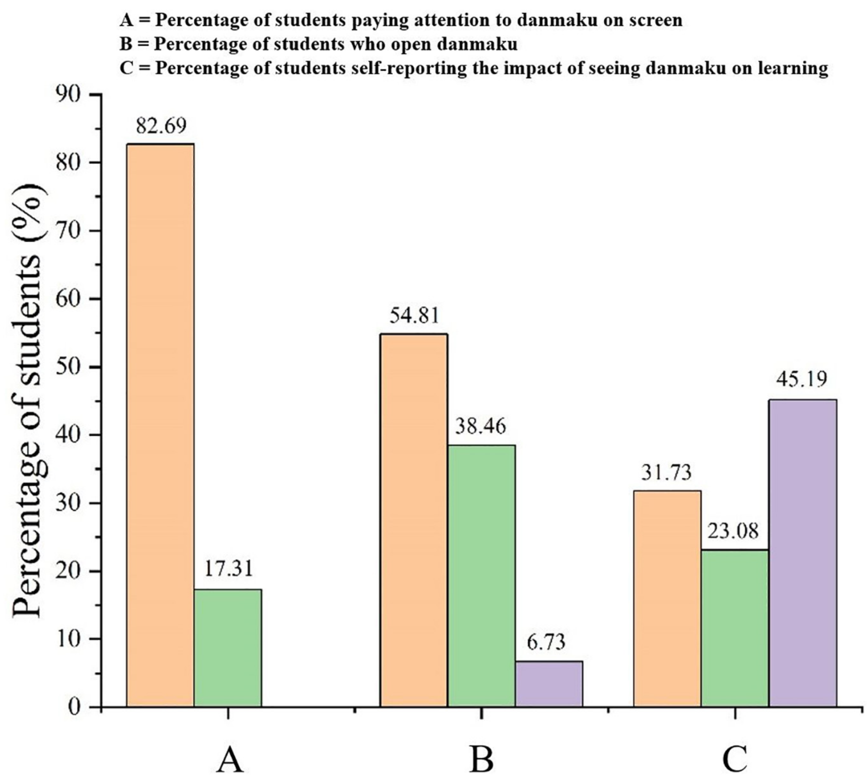


FIGURE 2
Learner expectations regarding the impact of viewing danmaku during instructional videos. **(A)** Differences in intrinsic cognitive load across experimental conditions; **(B)** differences in extraneous cognitive load across experimental conditions; and **(C)** differences in cognitive load across experimental condition.

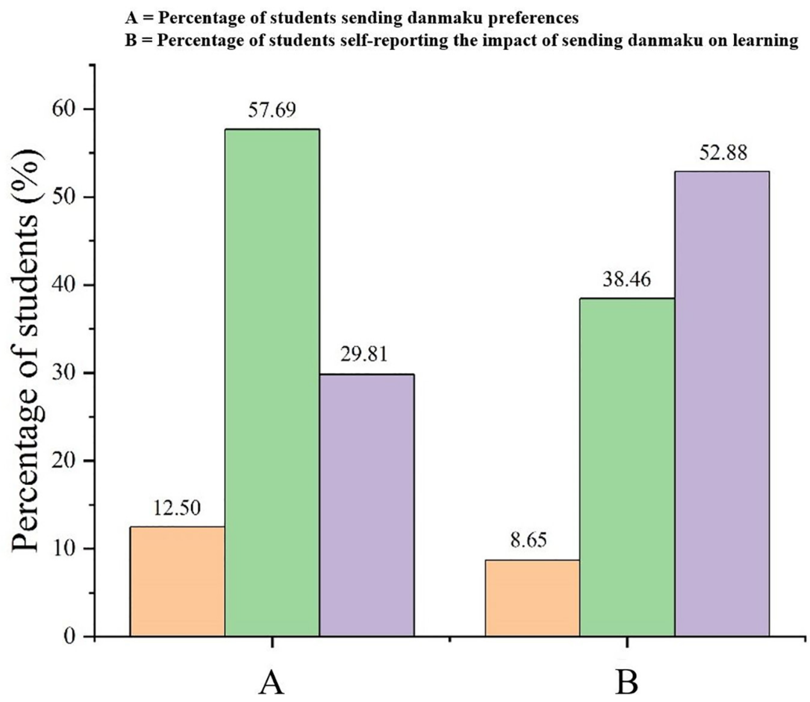


FIGURE 3
Learner expectations regarding the impact of sending danmaku during instructional videos. **(A)** Differences in retention across experimental conditions and **(B)** differences in transfer across experimental conditions.

TABLE 2 Means and standard deviations of all variables by group.

Dependent variables	NDS		TDS		UDS				
	M	SD	*M*	SD	*M*	SD	*F*	*p*	η_p^2
Learning performance									
Retention	18.84	4.07	15.79	3.21	13.50	3.40	19.37***	<0.001	0.28
Transfer	17.10	4.60	14.96	3.33	13.37	4.66	6.75**	0.002	0.12
Cognitive load	5.76	0.92	6.08	1.38	6.97	1.40	8.60***	<0.001	0.15
Intrinsic cognitive load	6.51	1.69	7.06	1.63	8.15	1.33	9.81***	<0.001	0.16
Extraneous cognitive load	4.43	1.80	4.57	1.98	6.09	2.04	7.72**	0.001	0.13
Germane cognitive load	6.34	1.83	6.60	2.35	6.68	2.24	0.23	0.796	0.01
Parasocial interaction	3.89	0.99	4.20	1.38	4.84	1.37	5.14**	0.007	0.09

NDS, no danmaku sent; TDS, three danmaku sent; UDS, unlimited danmaku sent; *M*, mean; SD, standard deviation; *F*, *F*-statistic; η_p^2, effect size. $^{**}p < 0.01$, $^{***}p < 0.001$.

interfere with learning were that the danmaku would block the screen and distract them.

Figure 3 presents the preference of learners sending danmaku in the instructional videos. Figure 3A showed that most learners (57.69%) did not like sending danmaku during the learning process; a small number of learners (29.81%) did not have any preference about sending danmaku; while only 12.50% of learners enjoyed sending danmaku. Only a very small percentage of learners (8.65%) believed that sending danmaku would improve their learning, reporting that sending danmaku would be a way to communicate with others, which would help them participate more actively in class interactions, enabling them to understand the problems and solve them quicker. In contrast, a significant number of the learners (38.46%) believed that sending danmaku would hinder their learning (see Figure 3B), thinking that it would be a waste of time or that it would distract them. Finally, 52.88% of the learners thought that sending danmaku would have no impact on their learning.

Descriptive statistics for all variables in the three groups are reported in Table 2. To test the differences across the different groups, one-way ANOVA was conducted for all variables using SPSS 24.0. Mplus 8.3 was used to analyze the mediating effect [bootstrap analysis, repeat sampling, 2,000 times following recommendations by DiCiccio and Efron (1996)]. In the current research, statistical significance was set at $p < 0.05$. In ANOVA, the effect size was represented by η_p^2; an η_p^2 of 0.01 indicated a small effect size, an η_p^2 of 0.06 indicated a moderate effect size, and an η_p^2 of 0.14 indicated a large effect size (Cohen, 2013).

Learning performance

ANOVA showed that there was a significant difference across the three experimental conditions in retention [$F(2,101) = 19.37$, $p < 0.001$, $\eta_p^2 = 0.28$; see **Figure 4A**]. *Post hoc* tests (least significant difference; LSD) showed that the retention level of learners in the NDS group was significantly different from those in the TDS and UDS groups (MD = 3.06, $p = 0.001$, 95% CI [1.36, 4.75]; MD = 5.34, $p < 0.001$, 95% CI [3.63, 7.05]; respectively). The learners in the TDS group also showed better retention than the learners in the UDS group (MD = 2.29, $p = 0.009$, 95% CI [0.58, 4.00]). The results showed that the number of danmaku sent by learners while viewing the instructional video affected their learning retention.

ANOVA also revealed a significant difference across the three experimental conditions in transfer [$F(2,101) = 6.75$, $p < 0.001$, $\eta_p^2 = 0.12$; see **Figure 4B**]. *Post hoc* tests (LSD) showed that students in the NDS group had a better transfer level than those in the TDS and UDS groups (MD = 2.14, $p = 0.037$, 95% CI [0.13, 4.15]; MD = 3.73, $p < 0.001$, 95% CI [1.71, 5.76]; respectively). There was no significant difference between the TDS and UDS groups ($p > 0.05$). The results indicated that learners who did not send danmaku on the screen during video learning had the best transfer. Therefore, Hypothesis 1 was supported.

Cognitive load

First, ANOVA found a significant difference in intrinsic cognitive load across the three experimental conditions [$F(2,101) = 9.81$, $p < 0.001$, $\eta_p^2 = 0.16$; see **Figure 5A**]. *Post hoc* tests (LSD) revealed that the intrinsic cognitive load of the UDS group was significantly different from that of both the TDS and NDS groups (MD = 1.09, $p = 0.004$, 95% CI [0.35, 1.83]; MD = 1.63 $p < 0.001$, 95% CI [0.89, 2.38]; respectively). Second, ANOVA showed that there was a significant difference in extraneous cognitive load across the experimental conditions [$F(2,101) = 8.60$, $p < 0.001$, $\eta_p^2 = 0.15$; see **Figure 5B**]. *Post hoc* tests (LSD) showed that the self-reported extraneous cognitive load of learners in the UDS group was significantly different from that in the TDS and NDS groups (MD = 1.52, $p = 0.002$, 95% CI [0.59, 2.44]; MD = 1.66, $p = 0.001$, 95% CI [0.73, 2.59]; respectively). However, no statistically significant difference was found compared to the ANOVA test on learners' self-reported germane cognitive load. Finally, ANOVA also revealed a significant difference in cognitive load across the three experimental conditions [$F(2,101) = 8.60$, $p < 0.001$, $\eta_p^2 = 0.15$; see **Figure 5C**]. *Post hoc* tests (LSD) demonstrated that learners in the UDS group reported higher cognitive load than both the TDS and NDS groups (MD = 0.89, $p = 0.004$, 95% CI [0.30, 1.49]; MD = 1.21, $p < 0.001$, 95% CI [0.61, 1.81]; respectively). These self-report results suggest that sending danmaku onto the screen can influence learners' cognitive load, as evidenced by the fact that learners in the UDS group had a higher cognitive load (including intrinsic and extraneous cognitive load) than reported by the learners in the other two groups. Thus, Hypothesis 2 was partially supported.

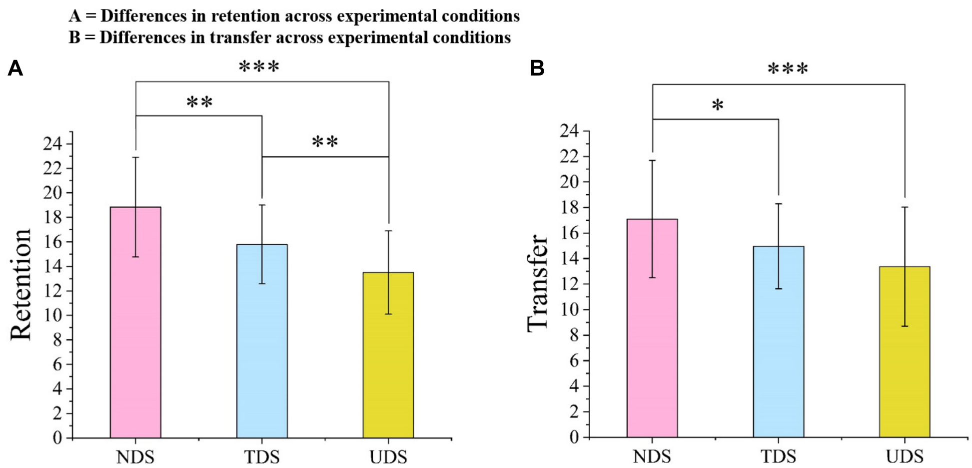


FIGURE 4
Differences in retention and transfer across experimental conditions. **(A)** Differences in retention across experimental conditions and **(B)** differences in transfer across experimental conditions. *$p < 0.05$, **$p < 0.001$, ***$p < 0.001$.

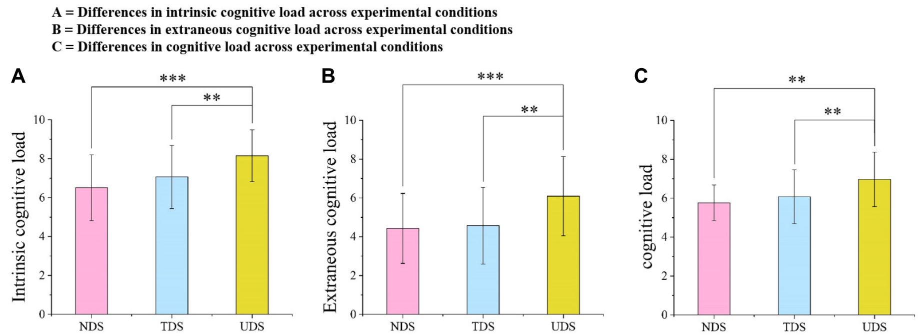


FIGURE 5
Differences in cognitive load across the experimental conditions. **(A)** Differences in intrinsic cognitive load across experimental conditions; **(B)** differences in extraneous cognitive load across experimental conditions; and **(C)** differences in cognitive load across experimental conditions. **$p < 0.01$, ***$p < 0.001$.

Parasocial interaction

ANOVA showed that there was a significant difference in learners' parasocial interaction across the three experimental conditions [$F(2,101) = 5.14$, $p = 0.007$, $\eta_p^2 = 0.09$; see Figure 6]. *Post hoc* tests (LSD) revealed a statistically significant difference in the UDS group relative to the TDS (MD = 0.64, $p = 0.036$, 95% CI [0.04, 1.25]) and NDS (MD = 0.95, $p = 0.002$, 95% CI [0.35, 1.56]) groups. It was concluded that learners who sent danmaku on the screen with no limit to how many they could send had higher parasocial interaction scores compared to those of the learners in the other two groups. Thus, Hypothesis 3 was supported.

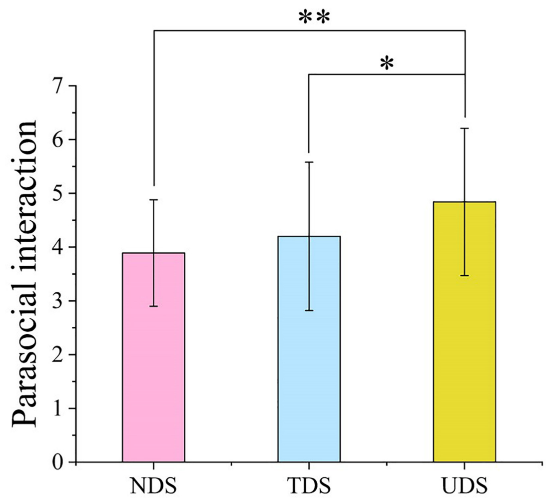


FIGURE 6
Differences in parasocial interaction across the experimental conditions. *$p < 0.05$, **$p < 0.01$.

Regression analysis and mediating effect

Intercorrelations among research variables

The Pearson correlation coefficients between the research variables and their significance levels are shown in Table 3. In terms of learning performance, learners' knowledge retention scores were significantly negatively linked with conditions and parasocial interaction ($rs = -0.53, -0.21$, $ps < 0.05$). Knowledge transfer scores were significantly positively linked with retention scores ($r = 0.54$, $p < 0.001$), but significantly negatively linked with conditions ($r = -0.34$, $p < 0.001$). In terms of learning process experience, cognitive load was significantly positively linked with conditions, parasocial interaction, intrinsic, extraneous, and germane cognitive loads ($0.37 < rs < 0.72$, $ps < 0.01$). Extraneous cognitive load was significantly positively linked with conditions and intrinsic cognitive load ($rs = 0.29, 0.33$, $ps < 0.01$). Intrinsic cognitive load was significantly positively linked with conditions ($r = 0.40$, $p < 0.001$). Parasocial interaction was also significantly positively linked with conditions ($r = 0.30$, $p = 0.002$). There were no significant links among the remaining variables.

Mediating effect analysis

The results of ANOVA showed that the retention and transfer scores of learners under different conditions were opposite to their self-reported cognitive load and parasocial interaction, and Pearson correlation analysis also found similar results. Therefore, this study constructed a mediation model (see Figure 7) to test whether the effects of the conditions on retention and transfer could be explained by parasocial interaction and cognitive load. The mediating effects were analyzed by standardizing all scores (Z-score) and performing structural equation modeling. To ensure conciseness in the model, all insignificant path coefficients and confidence

TABLE 3 Pearson correlations for research variables.

Variables	1	2	3	4	5	6	7
1. Conditions	–						
2. Parasocial interaction	0.30**	–					
3. Intrinsic cognitive load	0.40***	0.16	–				
4. Extraneous cognitive load	0.33**	0.17	0.29*	–			
5. Germane cognitive load	0.06	0.33**	0.17	0.26*	–		
6. Cognitive load	0.37***	0.33**	0.61***	0.72***	0.72***	–	
7. Retention	−0.53***	−0.04	−0.21*	−0.18	0.01	−0.18	–
8. Transfer	−0.34***	0.004	−0.11	0.01	0.17	0.05	0.54***

Pearson correlation coefficients and significance levels reported. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

intervals were deleted from the initial model. This model demonstrated a good data fit (χ^2/df = 1.44, comparative fit index = 0.96; Tucker-Lewis index = 0.95, root mean square error of approximation = 0.07, standardized residual mean root = 0.08). **Figure 7** shows the hypothesis testing results of the direct and indirect path coefficients of the mediation model. The results showed that all conditions had a significant direct effect on retention and transfer ($\beta = -0.56$, $p < 0.001$; $\beta = -0.39$, $p < 0.001$; respectively). Conditions not only significantly positively influenced parasocial interaction ($\beta = 0.31$, $p < 0.001$), but parasocial interaction also significantly positively influenced cognitive load ($\beta = 0.31$, $p = 0.021$). However, parasocial interaction and cognitive load did not influence retention and transfer ($ps > 0.05$). Therefore, Hypothesis 4 was not supported.

Discussion

The primary result of this study was that learners in the NDS group scored higher on both retention and transfer than those in the TDS and UDS groups. However, learners' self-reported cognitive load and parasocial interaction scores contradicted the finding on learning performance. There was also evidence that the more danmaku sent, the higher the learners' level of parasocial interaction, which leads to a higher cognitive load. However, excessive parasocial interaction and cognitive load were not factors in the poor learning performance of learners. In the remainder of this discussion, we consider some of the differences between our findings and previous work on social interaction in learning, and then discuss some of the limitations of this evidence.

Our findings contrast with some reports in the existing literature that learners who participated in real-time exchange of messages in the classroom performed better in their learning (Weisz et al., 2007; Li et al., 2014). This study found that the more danmaku were sent, the less attentional impact they had on the learners. The survey data also indicated that only a small minority of learners preferred to send danmaku while learning. Indeed, a considerable number of learners believed that sending danmaku would hinder their learning ability. In their opinion, there were two main reasons for this expected negative impact: distraction and wasting time, both of which would greatly reduce their efficiency in learning. Learners did waste a lot of time on sending danmaku, ignoring the learning content itself, particularly the learners in the UDS group. The content of danmaku sent by the learners in the UDS group may have been random, too frequent, or dense, and may have also interfered with learners' attention (Yang et al., 2019). Therefore, the more danmaku were sent, the easier it was to induce the split-attention effect (Ayres and Sweller, 2005). However, when learners interacted with others in the process of video learning, they had to be paying constant attention to and replying to the danmaku information scrolling along the top of the screen while also processing the learning content information. The learners had to integrate both the learning content and the danmaku information with their existing relevant knowledge, which led to the UDS group reporting a higher cognitive load in the process of video learning. In theory, higher cognitive load could lead to a lower learning performance (Ayres and Sweller, 2005; Yang et al., 2019). However, the mediation model constructed in this study did not support this assumption. Instead, it suggested that cognitive load was not an indirect cause of poor learning performance. Future research should further clarify the impact mechanism at play from the aspect of attentional impact. Furthermore, regarding the type of danmaku sent, some of the danmaku sent were social or novelty danmaku, which may have had little impact on the cognitive process of the learners. Regarding the learning content, the learning material was information about the heart. Learners needed to learn a great deal of knowledge in a short period of time, which may have been too challenging or difficult for the learners, making it easy to create a "floor effect" (Yang et al., 2019). Moreover, most of the learners in the UDS group continued to interact with each other in the form of danmaku. This phenomenon may have inhibited their creativity, as it has already been established that it is not conducive to the retention and transfer of new knowledge (Diehl and Stroebe, 1987).

Why did the UDS condition improve parasocial interaction? The purpose of learners sending danmaku was to cause them to interact more effectively with the others in their group,

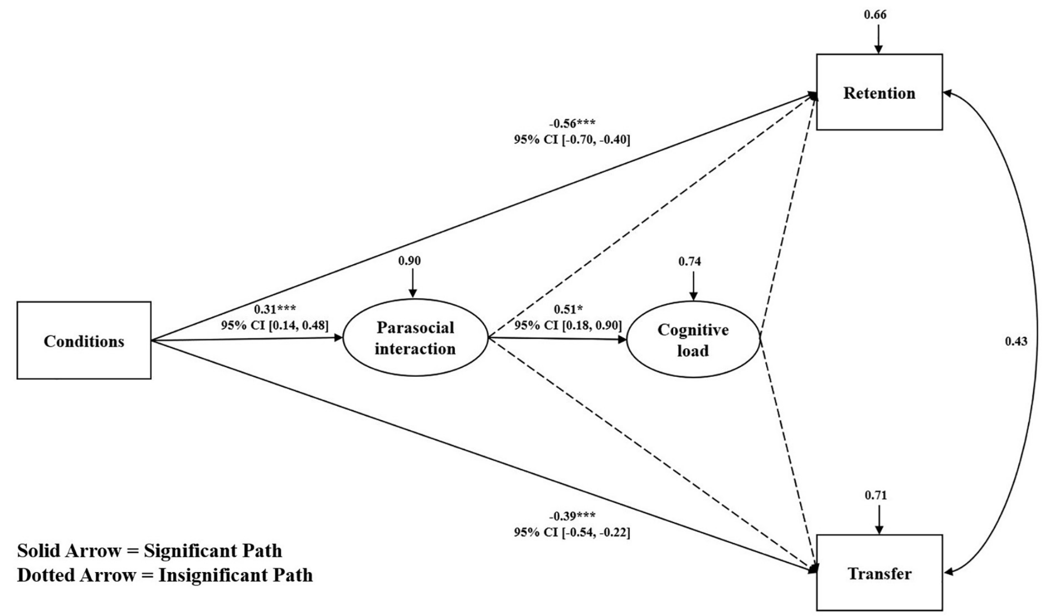


FIGURE 7
Mediation analysis with path coefficients. CI, confidence interval. *$p < 0.05$, ***$p < 0.001$.

changing them from passive recipients to active participants in learning. One of the biggest drawbacks of video learning is its lack of social interaction, especially between fellow learners (Palaigeorgiou and Papadopoulou, 2019). By sending danmaku, the learners were given the illusion of being connected in a digital world, as if they were communicating in real-time with others. In addition, the interactivity of sending danmaku is related to the cognitive process of learners (Zhang et al., 2019). Looking at the level or degree of learner interaction from the group with the lowest (NDS) to the highest (UDS) amount of communication reflected the learners' degree of mental engagement and the fact that the learners are guided into being more active participants (Patwardhan and Murthy, 2015). Vorderer et al. (2004) viewed parasocial interaction as a core precondition for media enjoyment and stated that this phenomenon facilitates learners' subsequent engagement with their respective content. Thus, stronger parasocial interaction seem to promote better learning performance in an instructional video settings (Beege et al., 2017). Parasocial interactions did not appear to be directly linked to learning performance in the present findings. Future work should therefore also investigate other variables such as arousal, affective, or motivational variables to gain insight into the underlying variables influencing learning performance (Beege et al., 2017). Interestingly, the mediation model constructed in this study showed that parasocial interaction increased cognitive load, making it clear that parasocial interaction has an impact on cognitive processes. Therefore, future studies should also provide more empirical evidence of this. Cognitive load theory and cognitive theory of multimedia learning have both been introduced as ways to expand the understanding of parasocial interaction (Ayres and Sweller, 2005; Mayer, 2005).

More generally, what stood out in the current study was that sending danmaku while learning from instructional videos not only hindered learners' learning performance, but also, the more danmaku sent, the greater the negative impact on learning. Although this type of social interaction can effectively improve learners' level of parasocial interaction, it can also increase their cognitive load.

The findings of the current study lead to some practical advice for instructional video designers, specifically, that danmaku have the characteristics of timely feedback and strong interaction, and it can meet the personalized learning needs of learners and improve their learning experience (Yang et al., 2019). However, if instructors want their learners to achieve better learning performance, they should not ask learners to send danmaku while the instructional videos are in progress. Therefore, if instructors want learners to have an interactive learning experience, we suggest that these interactions should take place after students are finished learning from the instructional videos (Pi et al., 2020).

Three features of this work limit the conclusions we can draw about the effects of learners sending danmaku in the process of their learning. First, this study only measured behavioral data which reflected the learning process of learners, and it is unclear whether other more objective and reasonable measurement techniques would produce comparable results. Future research may consider introducing eye-tracking and fMRI techniques to further reveal the cognitive and neural mechanisms of the impact of sending danmaku on learners' learning from an interdisciplinary perspective. Specifically, portable eye-tracking technology has been used to explore whether variables such as learners' preference for danmaku sending or danmaku style would have a moderating effect. fMRI and physiological polygraph have been used to measure blood flow changes caused by learners' neuron activity and the changes they experience in academic emotions at the physiological level, to seek out reasons that can be explained from a neurological and physiological perspective (Guo et al., 2019). Second, the findings of this study should not be generalized beyond the high school student population or other populations with similar demographic characteristics. In future studies, a wider range of data from across different disciplines should be collected from more participants from different schools and at different age levels (Wu et al., 2021). Finally, knowledge type may be an important moderator variable in the learning experience. In terms of declarative knowledge, danmaku did not play a role in improving this. However, for procedural knowledge, danmaku can in fact improve knowledge transfer (Yang et al., 2019). Therefore, future work can start from an understanding of different knowledge types in order to conduct further experiments.

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

CW contributed to the study conception and design. BJ, YL, and NF performed the material preparation and data collection. BJ and YM performed the data analysis and wrote the first draft of the manuscript. All authors commented on previous versions of the manuscript, read, and approved the final manuscript.

Funding

This work was funded by the International Joint Research Project of Huiyan International College, Faculty of Education, Beijing Normal University (ICER201902).

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 LM declared a shared parent affiliation with the author, YM to the handling editor at time the of review.

Publisher's note

References

Ayres, P., and Sweller, J. (2005). The split-attention principle in multimedia learning. *Cambridge Handbook Multimedia Learn.* 2, 135–146. doi: 10.1017/CBO9780511816819.009

Bai, Q., Hu, Q. V., Ge, L., and He, L. (2019). Stories that big danmaku data can tell as a new media. *IEEE Access* 7, 53509–53519. doi: 10.1109/ACCESS.2019.2909054

Bai, Q., Wei, K., Zhou, J., Xiong, C., Wu, Y., Lin, X., et al. (2021). Entity-level sentiment prediction in Danmaku video interaction. *J. Supercomput.* 77, 9474–9493. doi: 10.1007/s11227-021-03652-4

Beege, M., Schneider, S., Nebel, S., and Rey, G. D. (2017). Look into my eyes! Exploring the effect of addressing in educational videos. *Learn. Instr.* 49, 113–120. doi: 10.1016/j.learninstruc.2017.01.004

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences.* Milton Park: Routledge. doi: 10.4324/9780203771587

Craig, S. D., and Schroeder, N. L. (2017). Reconsidering the voice effect when learning from a virtual human. *Comput. Educ.* 114, 193–205. doi: 10.1016/j.compedu.2017.07.003

DiCiccio, T. J., and Efron, B. (1996). Bootstrap confidence intervals. *Stat. Sci.* 11, 189–228. doi: 10.1214/ss/1032280214

Diehl, M., and Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *J. Pers. Soc. Psychol.* 53:497. doi: 10.1037/0022-3514.53.3.497

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., and Eysink, T. H. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Comput. Hum. Behav.* 25, 360–370. doi: 10.1016/j.chb.2008.12.015

Giannakos, M. N., Chorianopoulos, K., and Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *Int. Rev. Res. Open Distrib. Learn.* 16, 260–283. doi: 10.19173/irrodl.v16i1.1976

Guo, J. J., Zhou, R., Zhao, L. M., and Lu, B. L. (2019). "Multimodal emotion recognition from eye image, eye movement and EEG using deep neural networks," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (New York, NY: IEEE), 3071–3074. doi: 10.1109/EMBC.2019.8856563

Herrmann, T., Nolte, A., and Prilla, M. (2013). Awareness support for combining individual and collaborative process design in co-located meetings. *Comput. Support. Cooperative Work* 22, 241–270. doi: 10.1007/s10606-012-9179-x

Hung, I.-C., and Chen, N.-S. (2018). Embodied interactive video lectures for improving learning comprehension and retention. *Comput. Educ.* 117, 116–131. doi: 10.1016/j.compedu.2017.10.005

Johnson, D. (2013). Polyphonic/pseudo-synchronic: Animated writing in the comment feed of Nicovideo. *Jpn. Stud.* 33, 297–313. doi: 10.1080/10371397.2013.859982

Kelley, T. L. (1925). The applicability of the Spearman-Brown formula for the measurement of reliability. *J. Educ. Psychol.* 16, 300–303. doi: 10.1037/h0073506

LeBreton, J. M., and Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organ. Res. Methods* 11, 815–852. doi: 10.1177/1094428106296642

Li, N., Verma, H., Skevi, A., Zufferey, G., Blom, J., and Dillenbourg, P. (2014). Watching MOOCs together: Investigating co-located MOOC study groups. *Distance Educ.* 35, 217–233. doi: 10.1080/01587919.2014.917708

Lim, J. S., Choe, M.-J., Zhang, J., and Noh, G.-Y. (2020). The role of wishful identification, emotional engagement, and parasocial relationships in repeated viewing of live-streaming games: A social cognitive theory perspective. *Comput. Hum. Behav.* 108:106327. doi: 10.1016/j.chb.2020.106327

Ma, X., and Cao, N. (2017). "Video-based evanescent, anonymous, asynchronous social interaction: Motivation and adaption to medium," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (New York, NY: ACM), 770–782. doi: 10.1145/2998181.2998256

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge, MA: Cambridge university press. doi: 10.1017/CBO9780511816819

Moreno, R., Mayer, R. E., Spires, H. A., and Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cogn. Instr.* 19, 177–213. doi: 10.1207/S1532690XCI1902_02

Palaigeorgiou, G., and Papadopoulou, A. (2019). Promoting self-paced learning in the elementary classroom with interactive video, an online course platform and tablets. *Educ. Inf. Technol.* 24, 805–823. doi: 10.1007/s10639-018-9804-5

Patwardhan, M., and Murthy, S. (2015). When does higher degree of interaction lead to higher learning in visualizations? Exploring the role of 'Interactivity Enriching Features'. *Comput. Educ.* 82, 292–305. doi: 10.1016/j.compedu.2014.11.018

Pi, Z., Deng, L., Wang, X., Guo, P., Xu, T., and Zhou, Y. (2022). The influences of a virtual instructor's voice and appearance on learning from video lectures. *J. Comput. Assist Learn.* 1–11. doi: 10.1111/jcal.12704 [Epub ahead of print].

Pi, Z., Hong, J., and Hu, W. (2019). Interaction of the originality of peers' ideas and students' openness to experience in predicting creativity in online collaborative groups. *Br. J. Educ. Technol.* 50, 1801–1814. doi: 10.1111/bjet.12671

Pi, Z., Tang, M., and Yang, J. (2020). Seeing others' messages on the screen during video lectures hinders transfer of learning. *Interact. Learn. Environ.* 1–14. doi: 10.1080/10494820.2020.1749671 [Epub ahead of print].

Pi, Z., Zhang, Y., Zhu, F., Chen, L., Guo, X., and Yang, J. (2021). The mutual influence of an instructor's eye gaze and facial expression in video lectures. *Interact. Learn. Environ.* 1–18. doi: 10.1080/10494820.2021.1940213 [Epub ahead of print].

Rice, K., and Redcay, E. (2016). Interaction matters: A perceived social partner alters the neural processing of human speech. *NeuroImage* 129, 480–488. doi: 10.1016/j.neuroimage.2015.11.041

Scott, P. H., Mortimer, E. F., and Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Sci. Educ.* 90, 605–631. doi: 10.1002/sce.20131

Shin, D. (2019). How does immersion work in augmented reality games? A user-centric view of immersion and engagement. *Inf. Commun. Soc.* 22, 1212–1229. doi: 10.1080/1369118X.2017.1411519

Teng, M., and Chan, B. H.-S. (2022). Collective colouring in danmu comments on Bilibili. *Discourse Context Media* 45:100577. doi: 10.1016/j.dcm.2021.100577

Van den Bulck, J. (2000). Is television bad for your health? Behavior and body image of the adolescent "couch potato". *J. Youth Adolesc.* 29, 273–288. doi: 10.1023/A:1005102523848

Vorderer, P., Klimmt, C., and Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Commun. Theory* 14, 388–408. doi: 10.1111/j.1468-2885.2004.tb00321.x

Vygotsky, L. S., and Cole, M. (1978). *Mind in Society: Development of Higher Psychological Processes*. Cambridge, MA: Harvard university press.

Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., and Konstan, J. A. (2007). "Watching together: Integrating text chat with video," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 877–886. doi: 10.1145/1240624.1240756

Wu, C., Jing, B., Gong, X., Mou, Y., and Li, J. (2021). Student's learning strategies and academic emotions: Their influence on learning satisfaction during the COVID-19 pandemic. *Front. Psychol.* 12:717683. doi: 10.3389/fpsyg.2021.717683

Yang, J. (2014). *The Effect of the Instructor on Learning Process and Effectiveness in Online Video Courses*, Ph.D thesis, Wuhan: Central China Normal University.

Yang, J., Wu, C., Liu, C., Zhang, Y., and Pi, Z. (2022). Personalization and pauses in speech: Children's learning performance via instructional videos. *Curr. Psychol.* 1–11. doi: 10.1007/s12144-022-03497-x [Epub ahead of print].

Yang, J., Wu, C., Pi, Z., and Xie, H. (2019). Facilitating Learning or Interfering Learning: A Meta-analysis of Impact of Danmaku on Learning. *e-Educ. Res.* 40, 84–90.

Yung, H. I., and Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: A cognitive load approach. *J. Educ. Technol. Soc.* 18, 153–160.

Zhang, Y., Qian, A., Pi, Z., and Yang, J. (2019). Danmaku related to video content facilitates learning. *J. Educ. Technol. Syst.* 47, 359–372. doi: 10.1177/0047239518811933

Zhao, L. (2014). *Expertise Reversal Effect in Multimedia*, Ph.D thesis, Wuhan: Central China Normal University.

Zhou, F., and Zhou, X. (2022). Across Time and Space, I Am Together with Many, Many Others": Digital Writing and Temporality on Chinese Social Media. *Soc. Media Soc.* 8:205630512211175. doi: 10.1177/20563051221117564

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fpsyg.2022.1092848

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University, China

REVIEWED BY
Yi Zhang,
Central China Normal University, China
Fangfang Zhu,
Central China Normal University, China

*CORRESPONDENCE
Jun Xiao
✉ ecnuxj2003@163.com

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

RECEIVED 08 November 2022
ACCEPTED 12 December 2022
PUBLISHED 06 January 2023

CITATION
Xiao J, Jiang Z, Wang L and Yu T (2023) What can multimodal data tell us about online synchronous training: Learning outcomes and engagement of in-service teachers.
Front. Psychol. 13:1092848.
doi: 10.3389/fpsyg.2022.1092848

What can multimodal data tell us about online synchronous training: Learning outcomes and engagement of in-service teachers

Jun Xiao[1]*, Zhujun Jiang[2], Lamei Wang[1] and Tianzhen Yu[1]

[1]Shanghai Engineering Research Center of Open Distance Education, Shanghai Open University, Shanghai, China, [2]Department of Educational Technology, School of Education, Shanghai Normal University, Shanghai, China

Teachers' engagement in online learning is a key factor in improving the effectiveness of online teacher training. This paper introduces a multimodal learning analytics approach that uses data on brain waves, eye movements and facial expressions to predict in-service teachers' engagement and learning outcomes in online synchronous training. This study analyzed to what extent the unimodal and multimodal data obtained from the in-service teachers (n = 53) predict their learning outcomes and engagement. The results show that models using facial expressions and eye movements data had the best predictive performance on learning outcomes. The performance varied on teachers' engagement: the multimodal model (integrating eye movements, facial expressions, and brain wave data) was best at predicting cognitive engagement and emotional engagement, while the one (integrating eye movements and facial expressions data) performed best at predicting behavioral engagement. At last, we applied the models to the four stages of online synchronous training and discussed changes in the level of teacher engagement. The work helps understand the value of multimodal data for predicting teachers' online learning process and promoting online teacher professional development.

KEYWORDS

engagement, eye-tracking, facial expression, EEG, in-service teacher training

Introduction

The COVID-19 pandemic has strongly boosted the development of online learning, which, however, does not prove very effective because the students are poorly motivated and engaged due to untimely feedback, lax supervision, and other factors (Nasir and Ngah, 2022). At the same time, teachers are facing great challenges in their professional development as the transition from traditional education to online learning will cause in them mental changes in relation to their identity as educators and their ideas about education (Wang et al., 2010; Teräs, 2014; Richardson and Alsup, 2015). As teachers must get prepared for online teaching in a short period and quickly grasp the methods and skills needed, providing effective training on professional development for them is of great importance. Advances in educational technology and online learning platforms and changes in educational modes (from offline to online) have made online teacher professional development (OTPD) possible and popular (Parsons et al., 2019; Nami, 2021). OTPD is defined as a format of teacher professional development (TPD) that provides teachers with continuous learning through ICT media (e.g., asynchronous, synchronous, blended or other forms of courses, seminars or learning modules provided online), without having to meet in person with their trainers/instructors and peers each time (Rogers, 2001; Ansyari et al., 2022). It offers a more flexible and personalized way of learning for teachers that overcomes geographical barriers (Chen et al., 2009; Ross, 2011; Powell and Bodur, 2019). Besides, it was found no differences between in-person TPD and OTPD in terms of teacher perceptions and learning outcomes (Ansyari et al., 2022).

Previous studies on OTPD usually focus on online learning efficiency from the perspective of technology application and management (Wang et al., 2010). But as a matter of fact, technology *per se* cannot promote TPD – that can only be realized by further clarifying the relation between technology and TPD. Engagement is an effective predictive indicator of long-term learning performance (Camacho et al., 2020). There is a growing body of research that demonstrates the importance of engagement for learning and achievement. In the field of teacher professional development, engagement is a key dimension in ensuring that teachers receive a complete training program, and some studies have shown that high levels of engagement in training help teachers apply the knowledge and skills they have learned to their practice after the training is complete (Holmes et al., 2021). According to Fredricks et al. (2004), the integration of behavior, emotion, and cognition under the concept of engagement is valuable because it can provide a richer characterization of learning than single-component studies. With the growing importance of online teacher training in recent years, researchers have turned to focus on teachers' engagement in the online learning environment. For instance, Liu and Zhang (2021) found that teachers generally think online learning is not interactive enough, giving it an average score of only 2.36 out of 5 on the interactive level, i.e., interaction with peers and instructors. Philipsen et al. (2019) found that timely support and feedback can help teachers specify their learning needs and improve training efficiency. In other words, compared with asynchronous learning, synchronous and blended learning methods can better stimulate the teachers' enthusiasm and engagement in the online training program.

There are three ways to measure engagement – self-report questionnaire, data mining based on learning logs, and sensor-based technology (Appleton et al., 2006; Sinatra et al., 2015; Camacho et al., 2020). In the past, the research on learning engagement is generally focused on the measurement of single dimensions. This includes evaluating only behavioral engagement based on postures such as hand raising, note-taking, and head propping (Liao et al., 2019; Vanneste et al., 2021), or evaluating emotional and cognitive engagement in a VR environment (Dubovi, 2022) while ignoring behavioral engagement. But in fact, engagement should not be evaluated by separate dimensions (Sharma and Giannakos, 2020). In addition, as Cleary and Zimmerman (2012) pointed out, engagement is essentially a continuous process that fluctuates in time as students become immersed in learning, so it's necessary to measure the learners' engagement in a dynamic way. In this respect, some researchers noted that the grain size of engagement measures can range from the micro level (e.g., individual engagement in the present moment, task, or learning activity) to the macro level (e.g., a group of learners in a class, course, school, or community) and suggested that at the micro level, engagement can be measured using physiological and psychological indicators such as brain imaging, eye tracking, response time, or attention allocation (Broughton et al., 2010; Sinatra et al., 2015; Dubovi, 2022). In the field of teacher professional development, we found that more and more research is focusing on teachers' learning processes, especially incorporating physiological and psychological data. For example, Chang et al. (2018) explored teachers' emotional experiences by coding the nonverbal expressions of their recorded videos. Wolff et al. (2016) investigated differences in how expert and novice teachers perceive problematic classroom scenes with eye-tracking technology. In addition, some researchers have also focused on using data such as facial expressions to evaluate the quality of teachers' teaching for their professional development (Zheng et al., 2020). It can be seen that currently in the field of TPD, while objective data channels are receiving increasing attention from researchers, multimodal data are less explored.

These days more researchers have come to analyze multimodal data because they, compared with unimodal data, can integrate subjective (e.g., self-report questionnaire) and objective data, and enable the capturing of the cognitive, emotional and behavioral learning process (Sinatra et al., 2015) from multiple perspectives. Cognitive engagement reflects the use of deep learning strategies, involving the integration of new information and existing knowledge. In measuring cognitive engagement, electroencephalogram (EEG), a neuroimaging technology, can capture the total activities of all nerve cells simultaneously oscillating in the learning process (Niedermeyer and da Silva, 2005). Studies have shown that the four patterns of EEG frequency are strongly related to emotional and cognitive states (Hassib et al., 2017). Baceviciute et al. (2020) used EEG to capture the learners' cognitive process in VR learning and found that VR learners displayed a higher level of Theta activities in the parietal lobe, which implied the possible use of long-term memory coding, searching, and other cognitive approaches. The visual attention data channel is another objective means of identifying fluctuations in cognitive engagement during learning. Bixler and D'Mello (2016) used eye-tracking sensors to record general eye gaze indicators, such as the number of fixations, fixation durations, variability in fixation durations, and saccade lengths, to measure wandering during computerized reading. Moreover, galvanic skin response (GSR), heart rate (HR) are also used to measure cognitive load (Cranford et al., 2014; Larmuseau et al., 2020) and concentration (Cooper et al., 2006; Sharma et al., 2020).

Regarding emotional engagement, it refers to the learner's emotion-related states during learning activities, such as happiness, enjoyment, boredom and frustration (D'Mello et al., 2017). Facial expressions are mostly used to measure and predict emotional engagement, for instance, when learners interacted with a game-based learning environment, Taub et al. (2019) captured seven facial expressions (i.e., joy, sadness, disgust, contempt, surprise, fear, anger) of learners and combined them with a traditional self-report questionnaire to portray the dynamics of learners' emotional engagement. Behavioral engagement refers to a person's behaviors of efforts and contributions during a learning activity (Fredricks et al., 2004). Since the mind–body connection suggests that observable physical responses can be used to infer unobservable mental states (D'Mello et al., 2017), some researchers have collected data such as human-computer interactions based on gamified learning environments to measure behavioral engagement (Psaltis et al., 2018). In a synchronous learning environment, non-verbal cues such as facial expressions, gestures and body postures captured from video image frames of classroom data can be used to effectively identify unobtrusive behavioral engagement (Whitehill et al., 2014; Ashwin and Guddeti, 2019).

According to Ning and Downing (2011) and Liu and Zhang (2021), learning experience can be interpreted as the learner's interaction with the teaching and learning environment, leading to the acquisition of subject-related knowledge or the development of personal/professional skills. The previous studies of learning prediction focused on identifying risk learners by using online learning data to predict dropout rate (Costa et al., 2017; Moreno-Marcos et al., 2020) and paid little attention to the learning experience. Nowadays, more researchers and scholars are paying attention to learners' interests, motivation, engagement, and other indicators, the development of which is greatly beneficial for improving self-directed learning in the learners and improving the teaching process (Wang et al., 2022). It is equally important because it helps improve the learning experience. Furthermore, it is possible to precisely predict a series of learning indicators with sensor-captured data, including data on emotions, eye movements, brain waves, GSR, or various combinations of them (Emerson et al., 2020; Olsen et al., 2020; Sharma et al., 2020). However, these studies are mostly focused on human-computer interaction learning environments, such as gamification environments (Giannakos et al., 2019; Emerson et al., 2020) and human-robot interaction (Cui et al., 2022) with little attention to computer-assisted collaborative learning scenarios. Olsen et al. (2020) divided the students into groups of two and used multimodal data to predict collaborative learning outcomes. This is an innovative study that broadens the scope of the application of Multimodal Learning Analytics (MMLA) in collaborative learning. As a matter of fact, a key element of OTPD is collaborative and interactive learning among teachers, which also holds the key to adult learning (Powell and Bodur, 2019).

Generating data on teachers' behavioral patterns, cognitive processes, as well as emotional experiences, has the potential to help develop and refine more effective pedagogy and support tools for use in informal and formal teacher professional development opportunities. At present, although researchers have explored many data stream combinations, few studies in the field of OTPD have ever examined the relation between unimodal and multimodal data to understand their synergetic effects and ability to explain the teachers' performance in the test and other critical indicators (e.g., engagement). But that is what's vitally important because as adult learners, the teachers' online learning outcomes are also affected by multiple factors. For instance, "time sequence" plays an important role in interactions and communication during online synchronous learning, and it has been used to analyze interactions among fellow learners (Chen et al., 2009). This is to say that when online course designers guide the trainee teachers to study by themselves, discuss or make reports, the sequence of doing all that will affect the learners' degree of concentration and other aspects. Therefore, predictive analytics can help the designers understand the teachers' engagement and other experiences in online learning, which is of great importance for promoting self-regulated learning (SRL; Sharma et al., 2020). Although multimodal data has shown great potential in the field of education, its ability to serve as a means of understanding and improving teachers' learning processes remain largely unexplored. To better leverage the design capabilities of multimodal data, we need to evaluate the effectiveness of multimodal data. This paper systematically assesses how different data streams can benefit predictive analytics. Our findings quantify the expected benefits of using various multimodal data from physiological sensing and help advance research in the area of learning technologies.

Research objectives and research questions

In this paper, we build predictive models on the learners' eye gaze, facial expressions of emotions, brain waves, self-report engagement, and test of knowledge points in an effort to make up for the scarcity of literature in OTPD.

We aim to (a) build predictive models of different modal combinations and examine the precision of unimodal and multimodal models, including data acquisition, data preprocessing and model training. Specifically, data acquisition refers to the acquisition of learner brainwave timing data, eye movement timing data, and facial timing data. The data acquisition and pre-processing module are used to acquire the temporal data of brainwave, emotion, and eye movement over time as well as the questionnaire data. The preprocessing part completes the process of data cleaning, data purification, and time calibration to obtain the unified standard online learning temporal data under multimodality. The training model refers to the multimodal analysis system, which takes raw brainwave, eye movement and expression data as multimodal input data and questionnaire data as indicators to extract the features of multimodal input data and trains them to generate prediction models of different indicators. Among them, the feature engineering module adopts the form of automatic machine mining to realize dynamic feature extraction, feature filtering, feature correlation analysis with questionnaire big data, and feature principal component analysis for online learning temporal data.

(b) use the models to predict in-service teachers' changing engagement in the learning process. The feature engineering segments the temporal data according to the teaching design, and further feature extraction is performed for each segment of data. The data modeling and analytical inference module models the time-series data according to its features with participation and knowledge tests, which can be used to infer the indicators within each period. The stage prediction refers to using the indicator model to make predictions for the input data in different periods to get each indicator within different periods. This includes metrics for groups and metrics for individuals.

For these goals, we have three research questions.

1. Does multimodal data provide more precise predictions than those gained by unimodal data for engagement?
2. How well do combinations of brain waves, facial expressions and eye gaze predict the engagement of in-service teachers?
3. What are the features of learner engagement according to the prediction model?

Materials and methods

Participants, experimental design, and procedure

The participants in this study included 56 in-service teachers who were enrolled in a teacher training program in Shanghai, China. Participants had not previously attended a training program related to ClassIn. During the experiment, data about three teachers were invalid because of falling headbands or other reasons, which gave us valid data on 53 teachers for further analysis. There were 28 males and 25 females; 80.4% of them were aged 20–40 and 19.6% were 40–50. The participant's personal information will be kept confidential, and only their ID, testing score, and the captured data will be maintained. They will be notified of the data collection and asked to sign the Informed Consent Form (ICF).

The training course – Online teaching based on ClassIn – is selected for this experiment. ClassIn is a useful online class system that has been used in the schools of many localities across China. It has a rich pool of functions, but many teachers do not know how to use it to facilitate or improve their instruction, which is why the researchers decide to provide training on this subject. This course is focused on how to do online teaching through ClassIn and contains four main aspects:

(1) Critical view of online teaching, (2) Instruction on online teaching and ClassIn, (3) Experience with ClassIn, (4) Feedback and Reflection.

In light of the features of adult learning (Ke and Xie, 2009; Abedini et al., 2021), the course centers on collaborative tasks and involves four stages (Figure 1). The first stage is an introduction, in which the instructor, by sharing real cases, introduces what will be taught and urges the teachers to think and share their views on online teaching during the pandemic. In the second stage, the instructor will introduce the functions of ClassIn, such as the *Group Discussion Function*. The third stage features collaborative learning, in which teachers are divided into several groups to practice with ClassIn, e.g., preparation before class, interaction and feedback during class. The fourth stage is for feedback and reflection, in which the instructor takes the teachers to review and reflect on what they have learned.

The research design of our study is a single-group time series design that involves repeated measurement of a group (Ross and Morrison, 2004). The experimental protocol consists of three sessions and took a total of 75 min. 10 min before the experiment started help participants calibrate an eye tracker, EEG device, and facial expressions of emotions software. This study used BrainCo headbands called Focus 1 (Focus, BrainCo, 2022), a wearable EEG device with 3 hydrogel electrodes, to collect and analyze the EEG data at 160 Hz *via* Wi-Fi (Kosmyna and Maes, 2019). As to facial expressions, we first turned on cameras to record the facial expressions, then Facereader software was used to analyze the data (Terzis et al., 2013). Eye tracker (Tobii T120) was used to capture gaze data. Once calibration was completed, participants began the training course session which took 40 min. After the course, participants filled in a questionnaire and knowledge test that took 25 min.

Measures

The questionnaire, adapted from the ones developed by Deng et al. (2020) and Liu et al. (2010), concerns engagement – cognitive, emotional and behavioral (Fredricks et al., 2004). A Cronbach's alpha of 0.801 indicates good reliability of the questionnaire. Of the three indicators, cognitive engagement includes seven items (e.g., "I think about the relation among different knowledge during online learning") (0.833), emotional engagement includes six items (e.g., "I enjoy the atmosphere of online synchronous training" and "I like online synchronous training"), and behavioral engagement includes seven items (e.g., "When I have a question, I'd ask the instructor and fellow teachers through the chat box of the live streaming platform"). The questionnaire uses a five-point Likert scale, from "strongly agree" to "strongly disagree," a higher score indicates a higher level of reported engagement.

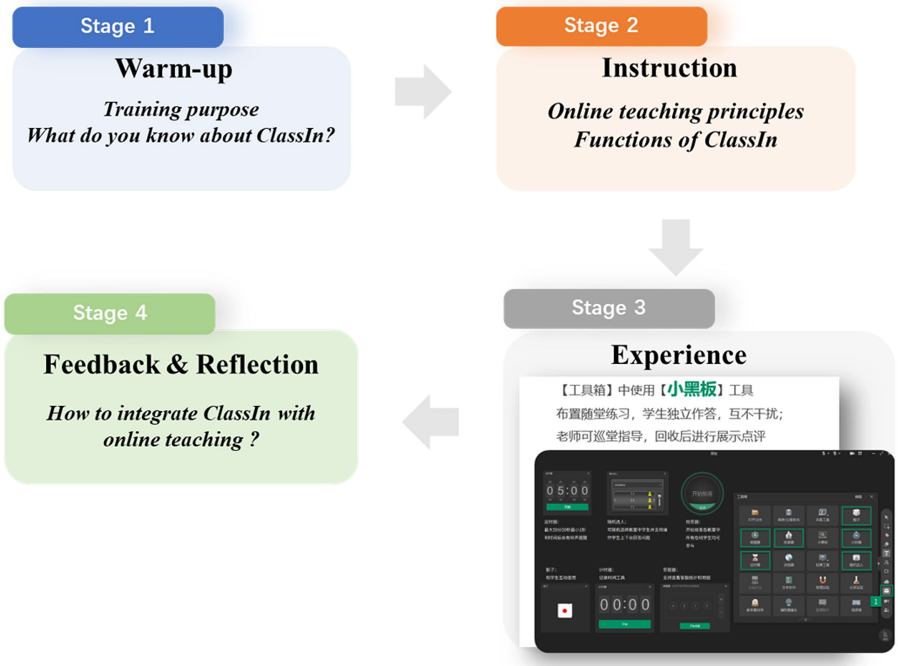


FIGURE 1
The procedure of the training program.

The knowledge test is designed by the research team (research members have long years of experience in teacher training) according to the contents of the course to understand how well the teachers have grasped what they are taught. There are single-choice questions (six items, 36 points), and multi-choice questions (eight items, 64 points) with a total score of 100. The test measured retention and also the comprehension and transfer of knowledge. For example, one of the questions asked the basic functions of ClassIn "What are the forms of assignments that can be submitted online in ClassIn." Another example is about the deep understanding of online teaching "What principles to keep in mind when students study in groups online." Analysis of the knowledge results indicated a good internal consistency score, with a Cronbach's alpha of 0.76.

Multimodal data collection and pre-processing

During the study, we captured participants' knowledge test scores. In addition, we collected sensor data from three different sources: EEG, facial expression, and eye-tracking.

EEG

To study brain wave values, we based our research on prior research that reported using consumer EEG headbands with 1–6 channels (Andujar and Gilbert, 2013). This study used the Focus band (Focus 1, BrainCo, 2022), co-developed by scientists from Harvard's center of brain science, to collect EEG features from the participants, which has also been used to detect engagement in previous studies (Kosmyna and Maes, 2019). According to the International 10–20 electrode placement system, one electrode is located at the FPz position, as well as the reference and ground electrodes of TP9. Neural oscillations α (7–11 Hz), β (11–20 Hz), and θ (4–7 Hz) were collected and normalized to EEG values between 0 and 100, and a higher value indicates higher attention. We use one data extracted every 10 s for analysis. After pre-processing, the EEG data of each learner was output with Time and EEG values.

Facial expression

To study the participants' facial expressions, we used FaceReader, a video-based facial expression tracking system (Noldus, 2019), to analyze the facial expressions of the teachers or learners. The analytical system is a reliable, professional software used for automatic facial expression analysis that can tell seven basic emotions: neutral, delight, surprise, sorrow, anger, fear, and disgust. Its working principles are as follows:

(a) Face finding: an algorithm based on deep learning is used to find human faces;
(b) Face modeling: nearly 500 key points are used to produce precise artificial face models;

(c) Face classification: an artificial neural network is used to classify the expressions.

Seven expressions are identified every time, each scored with a floating-point number ranging from −1 to 1. First, we removed empty or failed results. Second, we screened off those not obtained during the experiment (according to the official start and finish time). Then we selected the highest of the seven numbers as the facial expression of that very moment and classified it with a number from one to seven. The results of facial expression identification came out every 0.2 s, and we sampled them every second considering the huge amounts of data. We also noticed that the software did not export anything when the expression remained unchanged, and only exported a record with a corresponding time when it changed. Therefore, after we sampled the expressions by second, we filled up the lost values to make sure there was a facial expression score for every second. The expression data of all teachers ($n = 53$) put together constituted 3,901 pieces of time-sequenced records. Then we used tsfresh, a Python package for systematic feature engineering from time series and other sequential data, to extract 779 static features, which were imported into the models for training.

Eye tracking

To capture the trajectory of eye movements for analysis, we used Tobii T120 eye tracker to record where every participant looked on the screen during the course and how far their eyes were from the screen. Before the training, each participant was required to make adjustments for sitting position and distance. First, the researcher will turn on the eye-tracking test function of the eye-tracking device, and two dots will appear on the screen to indicate the gaze points of the left and right eyes. By micro-adjusting the seat distance to ensure that each participant's gaze point is at a close uniform level, the gaze range is just the entire learning material. The software recorded the position and distance-to-screen of both the left and right eyes and formed six-dimensioned data (Table 1).

We compared the time-sequenced data of eye movement trajectory with videos of the online course to match where the eyes gazed at and the duration of gaze with each frame. Then we designed polygonal interested areas that involved about four to six interested targets, and determined whether the captured eye movement trajectory (only horizontal and vertical position was considered) was in that area as a high-level description of the teachers' eye gaze. The proportion of how long they gazed at each interested area was recorded as an important indicator of interest or engagement.

TABLE 1 Set of GAZE features.

Feature	
EyePosLeftX	Horizontal offset from screen, left eye gaze
EyePosLeftY	Vertical offset from screen, left eye gaze
EyePosLeftZ	Distance from screen, left eye
EyePosRightX	Horizontal offset from screen, right eye gaze
EyePosRightY	Vertical offset from screen, right eye gaze
EyePosRightZ	Distance from screen, right eye

At the same time, the six-dimensioned data were imported into tsfresh as representing the eye movement trajectory to extract features, and altogether 779 static features were obtained, which, when applied to follow-up model training, led to six major features. The six major features were then combined with features describing the gaze at the interested areas, which gave us a collection of features reflecting the eye movement trajectory, and that was used for the next step of multimodal model training.

Multimodal predictive modeling

To study how well the multimodal data can predict the teachers' learning feedback, we developed a multimodal learning prediction model based on brain waves, facial expressions, and eye-gaze trajectory, and designed the questionnaire surrounding four analytical targets: cognitive engagement, emotional engagement, behavioral engagement, and learning outcomes. For constructing the model, this study uses an analysis method based on time series data features (Emerson et al., 2020; Olsen et al., 2020), by extracting features of time series data, filtering features, and analyzing the main features to obtain multiple sets of features. These feature values can be used as the input sample independent variables X of the model, i.e., brain wave feature, emotion feature, eye move feature. The dependent variables are derived from the questionnaire. Specifically, reported cognitive engagement, affective engagement, behavioral engagement, and test-based learning outcomes were used as the target Y for the study analysis, and a model from X to Y was constructed and trained to predict the target value Y for each stage of analysis (Figure 2). Facial expressions were labeled in seven different categories to create a time-sequenced series, from which we extracted 779 general features. The movement trajectory of the left and right eye (X, Y, Z) went through feature engineering to generate 779 general features too, so did the time-sequenced brain wave data. The three sets of features – totaling 2,337 – were imported into the decision tree as independent variables, while the scores of every target in the questionnaire were the dependent variables. To prevent over-fitting, they were imported into the training model, with 70% as the training set and 30% as the test set (Kang and Oh, 2020). The allocation ratio was set after several tests based on the best training fit. Each analytical target corresponded with a model, so four targets, and four multimodal predictive models. At that time, all data on the entire course were divided according to the four stages, and the four models were used to, respectively, predict each of the four targets in each stage.

During the experiment, we made the hypothesis that there must be some major features among the general ones that were strongly relevant to brain waves, facial expressions, and eye movement trajectory. So we extracted major features related to the predictive targets during model training, and imported them, as

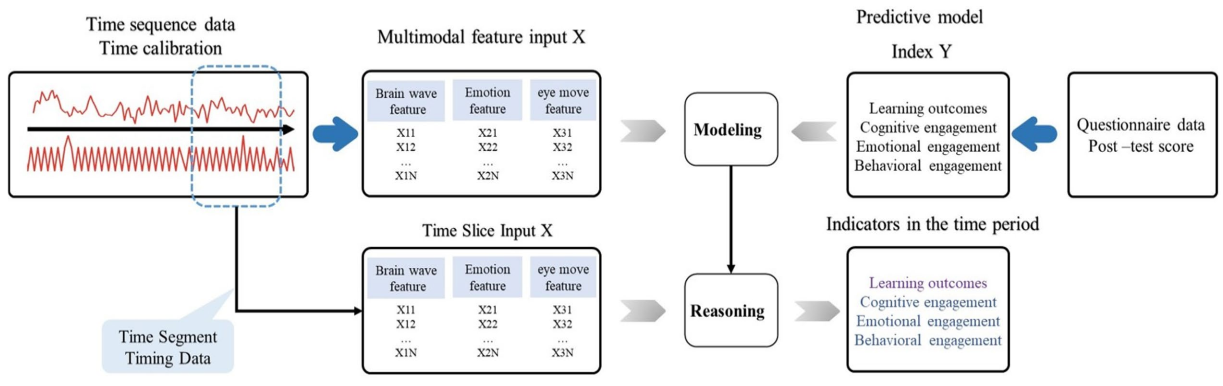


FIGURE 2
Multimodal data streams and predictive modeling approach.

descriptive of brain waves, facial expressions, and eye movement trajectory, into the models for training again, which produced multimodal predictive models based on major features. Different multimodal experiments generated different multimodal major features (about 4–10 of them), which meant the latter was data-sensitive.

Results

We use tsfresh to extract the features of temporal data, and import the obtained features to Decision Tree Classifier for training, Decision Tree Classifier comes from sklearn (an open source python language-based machine learning library), then sklearn's classification report function was used to automatically calculate precision, recall, F1. Precision is a measure of result relevancy while recall is a measure of how many truly relevant results are returned. The F1 score is reported as an agglomerative measure between precision and recall. The focus of this study is on the precision of the prediction model, so we choose it as the key indicator (Sharma et al., 2020).

To answer research question 1 and 2, we investigated how well unimodal and multimodal models (e.g., EGG, facial expression, and eye gaze) could predict the teachers' engagement. As shown in Table 2, the predictive model integrating data on eye movements, facial expressions and brain waves is the most precise (0.65) in predicting cognitive engagement, with the highest Recall (0.67) and F1 score (0.64) as well. That is higher than the scores of unimodal prediction and bimodal prediction.

The researchers also found that the multimodal predictive model integrating data on eye movements, facial expressions and brain waves had the highest precision (0.61), Recall (0.67) and F1 score (0.52) in predicting emotional engagement too, but that does not mean more modal data would naturally lead to higher predictive precision. For instance, the unimodal model using only brain wave data has a precision of 0.47, higher than any bimodal data combinations (see Table 3).

TABLE 2 Cognitive engagement prediction results.

Data used	Precision	Recall	F1
Gaze	0.36	0.43	0.40
EEG	0.26	0.48	0.33
Face	0.40	0.57	0.46
Gaze + EEG	0.43	0.48	0.45
EEG + Face	0.44	0.39	0.39
Gaze + Face	0.47	0.43	0.45
Gaze + Face + EEG	**0.65**	**0.67**	**0.64**

Bold values represent best performance.

Table 4 shows that multimodal predictive models are more precise than unimodal models in predicting behavioral engagement, with the model combining data on facial expressions of emotions and eye-gaze being most predictive with a precision of 0.75. Of unimodal models, the one using facial expressions data is most predictive with a precision of 0.43, while that using EEG data performs worst with a precision of only 0.17.

Table 5 shows that as far as learning outcomes are concerned, the predictive model combining data on eye movements and facial expressions has the highest precision of 0.66 – higher than the model integrating data on eye movements, facial expressions, and brain waves. We also found that of unimodal models, the one using eye movements data has the highest precision of 0.52 whereas that using brain wave data has the lowest precision of 0.29.

To answer question 3, we exported the scores of the learners' emotional, cognitive and behavioral engagement in the four learning stages. As shown in Figure 3, a thermal distribution map of multi-modal fusion characteristic data was used to evaluate the engagement of each stage, a darker color means a predicted score higher and closer to 5, and a lighter color means a score lower and closer to 0. Generally speaking, learners have the highest score on behavioral engagement and the lowest on emotional engagement. As to the change of engagement through four stages, cognitive

TABLE 3 Emotional engagement prediction results.

Data used	Precision	Recall	F1
Gaze	0.27	0.35	0.47
EEG	0.47	0.48	0.48
Face	0.24	0.23	0.28
Gaze + EEG	0.25	0.33	0.24
EEG + Face	0.35	0.30	0.43
Gaze + Face	0.31	0.30	0.42
Gaze + Face + EEG	**0.61**	**0.67**	**0.52**

Bold values represent best performance.

TABLE 4 Behavioral engagement prediction results.

Data used	Precision	Recall	F1
Gaze	0.43	0.38	0.39
EEG	0.17	0.38	0.24
Face	0.49	0.52	0.60
Gaze + EEG	0.28	0.24	0.25
EGG + Face	0.56	0.62	0.59
Gaze + Face	**0.75**	**0.52**	**0.60**
Gaze + Face + EGG	0.70	0.62	0.62

Bold values represent best performance.

TABLE 5 Learning outcomes prediction results.

Data used	Precision	Recall	F1
Gaze	0.52	0.48	0.44
EEG	0.29	0.24	0.26
Face	0.35	0.33	0.32
Gaze + EEG	0.56	0.43	0.41
EGG + Face	0.45	0.36	0.39
Gaze + Face	**0.66**	**0.43**	**0.41**
Gaze + Face + EEG	0.52	0.55	0.53

Bold values represent best performance.

engagement wanes first and waxes later. In the first stage, for example, the instructor aroused the teachers' interests by presenting a research report on the current status of online teaching and sharing real cases, and urged them to reflect and contemplate on the common problems occurring in their classes. Behavioral engagement waxes first and wanes later. In the second and third stages, the teachers discussed specific topics and solved problems collaboratively, including research, sharing of views, and group report, which stimulated their learning enthusiasm. Emotional engagement wanes first and waxes later. The highest score in the third stage indicates the highest emotional engagement during collaboration and interaction, which is consistent with the questionnaire results – teachers are generally more interested in collaborative learning.

As shown in Figure 4, researchers have developed 3D coordinates for engagement based on the predictive models. The X axis represents the learners' serial number, Y axis the four learning stages, and Z axis the predictions on cognitive, emotional and behavioral engagement. The coordinates can reflect how each learner's engagement changes through the four stages. We found that most learners maintain a high level of behavioral engagement through the stages with little change. Predictions on their emotional engagement show that most of them have a low level of emotional engagement at first, but some see it increasing over time. Their cognitive engagement changes rather drastically, and it drops significantly in the third and fourth stage for a few learners.

Discussion

Improved quality of instruction contributes to better student learning achievement (Ansyari et al., 2022). Teachers, as adult learners, must seek self-improvement constantly to promote professional development and embrace changes. That's why designing and planning high-quality teaching training for teachers is highly important (Creemers et al., 2012; Carrillo and Flores, 2020). Learning analytics is a key approach to refining the teaching process. Although the teachers' learning indicators can be explained with different data streams, one important question is how to merge the data obtained from various channels to provide a better, more comprehensive picture of the learning process (Chango et al., 2021). With the rapid development of artificial intelligence such as sensor technology and machine learning, it is possible to capture the participants' subconscious emotions (Vanneste et al., 2021) and use multimodal data to predict online learning process. In this paper, we extracted the features of multimodal data for training and generated predictive models for different indicators. To be specific, using an analytic method suited to the features of time-sequenced data, we extracted and filtered the features of time-sequenced data on brain waves, facial expressions, and eye movements analyzed the major features, and obtained multiple feature sets, which can be imported into the models as sample X. We took the indicators in the engagement questionnaire and knowledge test as target Y, and developed models matching X with Y for training to predict the target value of Y at each time period.

Addressing RQ1 (Does multimodal data provide more precise predictions from those gained by unimodal data for engagement?), we see that multimodal models are generally more precise than unimodal models on predicting engagement and learning outcomes. However, there are some differences in the predictive results for the three sub-dimensions of engagement. On the one hand, we found that the trimodal prediction model integrating data on facial expressions, eye movements, and brain wave is most precise regarding cognitive engagement and emotional engagement, while the bimodal prediction model that combines facial expression and eye movement data has the best predictive performance in terms of behavioral engagement and learning outcomes. On the other hand, we found that the predictive model

10.3389/fpsyg.2022.1092848

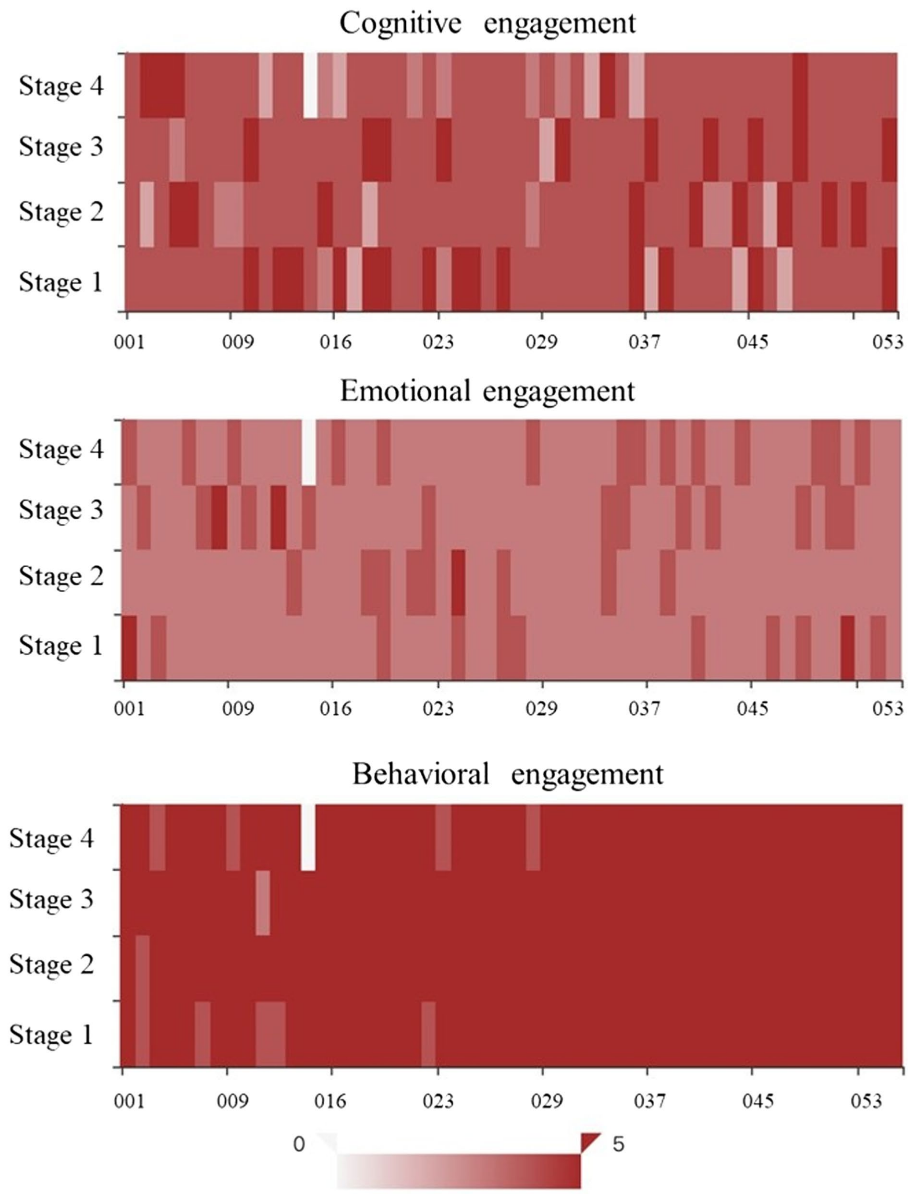


FIGURE 3
Thermal distribution map of multi-modal fusion characteristic data.

integrating Omni-modal data does not always produce the best predictions, which is consistent with the conclusions of previous studies (Emerson et al., 2020). One possible reason is that the excessive noises have undermined the model's robustness.

That brings us to our second research question, RQ2 (How well do combinations of brain waves, facial expressions and eye gaze predict the engagement of in-service teachers?). First, from the predictive results of cognitive engagement, we found that multimodal predictive models perform better than any unimodal model in prediction, and specifically, we found that the model using EEG data alone is least satisfactory. Cognitive engagement includes psychological positioning, cognitive efforts, and the thinking or attention aroused during the learning activity (Greene et al., 2015). In fact, an imbalance in cognitive understanding, if not properly addressed, may lead to emotional frustration. This may explain why the unimodal model using facial expression data does better than that using eye movement data or using EEG data in predicting cognitive engagement. Second, in terms of emotional engagement, although the most frequently used method to measure emotional engagement without disrupting the learners is analyzing their facial expressions, which helps capture their subconsciously fast-changing emotions (Taub et al., 2019; Vanneste et al., 2021), this study found that the predictive model using only data on facial expressions does not perform well, whereas the trimodal model has the best predictive performance. We also found that the unimodal model using EEG data only is

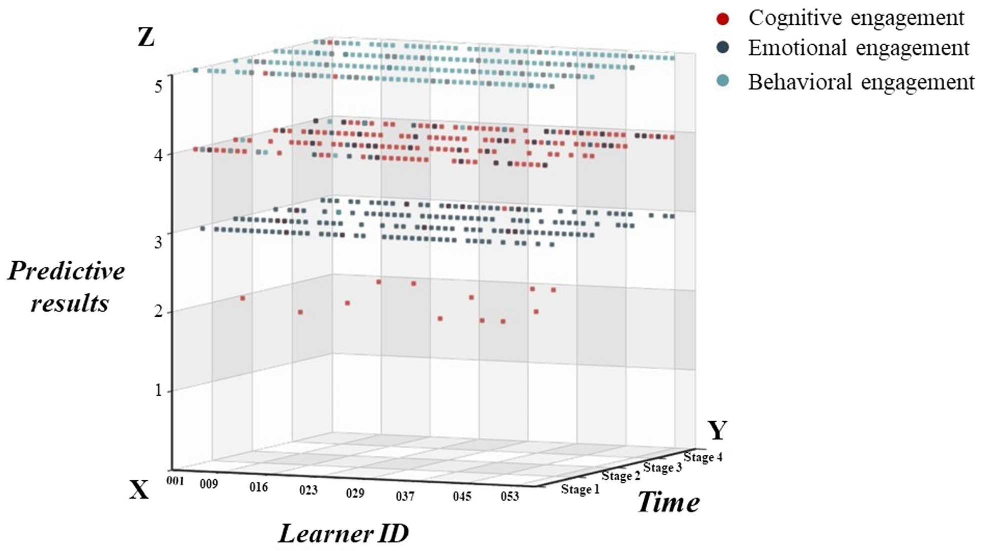


FIGURE 4
Spatial and temporal distribution coordinates of engagement.

the second most precise in predicting emotional engagement, better than other bimodal or unimodal models, which is inconsistent with a previous study by Soleymani et al. (2016), whose data results demonstrated that facial expression performs better than EEG data. The possible reason may be that our study was based on a real online learning environment. We did not provide learners with videos specifically selected as emotional arousal stimuli to cover the entire emotional range as previous studies have done. Third, this paper found that the model combining data on eye movements and facial expressions is the most precise in predicting behavioral engagement, but adding EEG data into the model would lower its precision. The facial expressions gave better prediction performance, which supports the findings of previous studies (Ashwin and Guddeti, 2019). In sum, the results add to a gap in the field of related research in the past, where many studies have confirmed that facial expression and gaze contribute to identifying, monitoring, and classifying behavioral engagement (D'Mello et al., 2017; Alkabbany et al., 2019), but few studies have focused on how well the predictions work in the combined model.

Addressing RQ3 (What are the features of learner engagement according to the prediction model?), this study showed that the highest score of teachers' cognitive engagement in the third learning stage. This consists with previous researchers' conclusions that cognitive engagement is essentially a process of continuous fluctuation that occurs when the person interacts with a specific scenario. When teachers interact with specific learning tasks or environments, cognitive engagement happens (Li and Lajoie, 2021). Helme and Clarke (2001) identified three interacting factors that influence cognitive engagement: the individual, the learning environment, and the task. In the third stage, the online synchronous training environment provides a platform for cognitive engagement where teachers work in small groups to collaborate around specific tasks, which helps stimulate deeper strategies and efforts. However, it is noteworthy that a few learners in this study had significantly decreased cognitive engagement in the third and fourth learning stage. The third stage of training in this study is the collaborative learning stage, where the learners' emotional and cognitive processes become more complex in an online collaborative learning environment because each group member's reaction affects the overall emotional climate and learning process (Törmänen et al., 2021; Ye and Zhou, 2022). One possible reason for this is that in this study, we allowed teachers to choose their topics for collaborative inquiry, but lacked scaffolding to facilitate deep reflection and cognitive processing and due to time constraints, some teachers exhibited relatively more low-level cognitive processes (e.g., understand; Lin et al., 2014). In other words, instructors can also appropriately clarify collaboration requirements and evaluation criteria to help learners with self-regulation and self-control (Dabbagh and Kitsantas, 2004).

As to emotional engagement, the predictions also show that of the three dimensions of engagement, the score of teachers' emotional engagement is the lowest but it increases gradually. This means as the learning activity proceeds, especially after the teachers are divided into groups, they get a stronger sense of belonging and consequently display a higher emotional engagement (Ulmanen et al., 2016). However, the overall score on this dimension remains low may because they are not familiar with each other or may not all be interested in the training theme. Previous research has found a significant relationship between the perceived value of feedback and the emotional engagement with feedback during online learning. Therefore it is not enough to provide feedback during an activity, it is also important to understand how teachers perceive the feedback they receive (Mayordomo et al., 2022).

The results also show that of the three dimensions of engagement, teachers' behavioral engagement has the highest score as well as the highest prediction in the second and third stages. The Expectancy-Value-Cost Model of Motivation suggests that perceived task value directly influences choice, persistence, and performance, and that engagement translates motivation into action (Barron and Hulleman, 2015), therefore, it is important to support teachers in finding value and relevance in their training (Wigfield et al., 2015). This implies that collaborative learning may be an important way of raising their behavioral engagement in online learning, as teachers may display various interactive behaviors with the contents, materials, and fellow teachers, such as research, communication, and division of work. In addition, we found that higher behavioral engagement does not necessarily represent a higher cognitive process. In other words, higher behavioral engagement may be predominantly low-level cognitive processes (e.g., memorization and comprehension; Ye and Zhou, 2022).

The findings of the study may contribute to the empirical and theoretical development of online teacher professional development. First, many studies have emphasized that online teacher training is beneficial to promote teachers' professional development. Our study quantifies the predictive and explanatory ability of multimodal data on teachers' online learning process, which can help advance online learning platforms to design and optimize online courses in the future. Second, this study focuses on an important indicator of teachers' online learning, namely engagement. In particular, our findings reaffirm that engagement is a fluctuating variable, and we find large differences in teachers' engagement in training across cognitive, emotional, and behavioral dimensions, as revealed by multimodal data, rather than the traditional use of questionnaires at the post-test. Third, the fact that this study found differences in the variation of teachers' engagement across instructional activity designs promotes our thinking about how to design sequences of instructional activities to improve the effectiveness of teacher training, especially regarding collaborative learning among teachers.

There are some limitations to this research which also can be considered for future research. First of all, the questionnaire is designed in such a way that the participants, out of habit, would prefer moderate answers to radical ones such as "strongly agree" or "strongly disagree." As a result, the models have no access to fringe scenarios and are therefore not good at predicting them. Wider samples should be considered in the future to enrich our findings, and the "think aloud" approach can also be adopted to examine and improve the validity of inferring data on the behavioral trajectory. Secondly, to not disturb the teachers during learning, we mainly used the usual data on brain waves, eye movements, and facial expressions for this experiment, but data on more dimensions can be incorporated in the future to expand and enrich the predictive models. Thirdly, as our findings indicate an inclination among the teachers to choose longer-term TPD (Philipsen et al., 2019), follow-up studies can be conducted going forward at greater depth by, for instance, collecting multimodal data on the teachers when they sign up for weeks-long, months-long or even year-long online training. Finally, this study found differences in sub-dimensions of engagement through a predictive perspective, the next step is to conduct a more in-depth analysis of the interplay between cognitive process, emotion, and behavioral engagement in conjunction with the predictive model. Besides, It is critical to help improve teacher training programs based on predicted effects, so that in the future, training course content and processes can be optimized in conjunction with design-based research methods.

Conclusion

The development of artificial intelligence, including sensor technology, has provided the means to collect and analyze learning data from various channels and to make the predictive models on learners' engagement and test performance more precise. This information has shed light on how to improve the approach to online teacher training and develop self-adaptive tools. Previous studies have shown the prospects of multimodal data in predicting learners' learning performance in human-computer interaction, but in the field of OTPD, hardly any researcher has ever noticed the synergizing potential of multimodal data for online synchronous learning.

It is against such a background that this paper created predictive models using various data combinations to examine and evaluate how precise the predictions on learners' engagement and test performance are. Unlike previous studies that only focused on one or two dimensions of engagement, we developed predictive models for all three dimensions – cognitive engagement, emotional engagement, and behavioral engagement – separately. The results show that by and large, models using bimodal or multimodal data are more precise in predicting engagement, but more modal data does not necessarily result in higher predictive precision. This study tries to make a predictive analysis of the learners' learning process based on the predictive models, which can reflect the real-time change of their engagement, as we found that the learners' cognitive engagement, emotional engagement, and behavioral engagement all displayed different features in different learning stages.

Data availability statement

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

Ethics statement

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

JX and ZJ: conceptualization, writing—review and editing, writing—original draft preparation, and methodology. JX: project administration and funding acquisition. LW: investigation process and data collection. TY: formal analysis. All authors contributed to the article and approved the submitted version.

Funding

This study was supported both by Shanghai Science and Technology Innovation Action Plan International Cooperation project "Research on international multi language online learning platform and key technologies (no. 20510780100)" and Science and Technology Commission of Shanghai Municipality research project "Shanghai Engineering Research Centre of Open Distance Education (no. 13DZ2252200)."

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

Abedini, A., Abedin, B., and Zowghi, D. (2021). Adult learning in online communities of practice: a systematic review. *Br. J. Educ. Technol.* 52, 1663–1694. doi: 10.1111/bjet.13120

Alkabbany, I., Ali, A., Farag, A., Bennett, I., Ghanoum, M., and Farag, A. (2019). Measuring student engagement level using facial information. in *IEEE International Conference on Image Processing (ICIP)*, 3337–3341.

Andujar, M., and Gilbert, J. E. (2013). "Let's learn! Enhancing user's engagement levels through passive brain-computer interfaces," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. eds. W. E. Mackay, S. A. Brewster and S. Bødker (New York, NY, USA: Association for Computing Machinery), 703–708.

Ansyari, M. F., Groot, W., and De Witte, K. (2022). Teachers' preferences for online professional development: evidence from a discrete choice experiment. *Teach. Teach. Educ.* 119:103870. doi: 10.1016/j.tate.2022.103870

Appleton, J. J., Christenson, S. L., Kim, D., and Reschly, A. L. (2006). Measuring cognitive and psychological engagement: validation of the student engagement instrument. *J. Sch. Psychol.* 44, 427–445. doi: 10.1016/j.jsp.2006.04.002

Ashwin, T. S., and Guddeti, R. M. R. (2019). Unobtrusive behavioral analysis of students in classroom environment using non-verbal cues. *IEEE Access* 7, 150693–150709. doi: 10.1109/ACCESS.2019.2947519

Baceviciute, S., Mottelson, A., Terkildsen, T., and Makransky, G. (2020). "Investigating representation of text and audio in educational VR using learning outcomes and EEG," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20)*. New York, NY, USA: Association for Computing Machinery, 1–13.

Barron, K. E., and Hulleman, C. S. (2015). "Expectancy-value-cost model of motivation," in *International Encyclopedia of the Social & Behavioral Sciences*. ed. J. D. Wright (Oxford), 503–509.

Bixler, R., and D'Mello, S. (2016). Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Model. User-Adap. Inter.* 26, 33–68. doi: 10.1007/s11257-015-9167-1

BrainCo (2022). Available at: https://www.brainco.tech (Accessed December 1, 2022).

Broughton, S. H., Sinatra, G. M., and Reynolds, R. E. (2010). The nature of the refutation text effect: an investigation of attention allocation. *J. Educ. Res.* 103, 407–423. doi: 10.1080/00220670903383101

Camacho, V. L., Guía, E. D. L., Olivares, T., Flores, M. J., and Orozco-Barbosa, L. (2020). Data capture and multimodal learning analytics focused on engagement with a new wearable IoT approach. *IEEE Trans. Learn. Technol.* 13, 704–717. doi: 10.1109/TLT.2020.2999787

Carrillo, C., and Flores, M. A. (2020). COVID-19 and teacher education: a literature review of online teaching and learning practices. *Eur. J. Teach. Educ.* 43, 466–487. doi: 10.1080/02619768.2020.1821184

Chang, C.-F., Gröschner, A., Hall, N. C., Alles, M., and Seidel, T. (2018). Exploring teachers' emotions via nonverbal behavior during video-based teacher professional development. *AERA Open* 4:233285841881985. doi: 10.1177/2332858418819851

Chango, W., Cerezo, R., Sanchez-Santillan, M., Azevedo, R., and Romero, C. (2021). Improving prediction of students' performance in intelligent tutoring systems using attribute selection and ensembles of different multimodal data sources. *J. Comput. High. Educ.* 33, 614–634. doi: 10.1007/s12528-021-09298-8

Chen, Y., Chen, N.-S., and Tsai, C.-C. (2009). The use of online synchronous discussion for web-based professional development for teachers. *Comput. Educ.* 53, 1155–1166. doi: 10.1016/j.compedu.2009.05.026

Cleary, T. J., and Zimmerman, B. J. (2012). "A cyclical self-regulatory account of student engagement-theoretical foundations and applications" in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. Wylie and C. Reschly (Boston, MA: Springer Science+Business Media), 237–257.

Cooper, N., Burgess, A., Croft, R., and Gruzelier, J. (2006). Investigating evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task. *Neuroreport* 17, 205–208. doi: 10.1097/01.wnr.0000198433.29389.54

Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., and Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Comput. Hum. Behav.* 73, 247–256. doi: 10.1016/j.chb.2017.01.047

Cranford, K. N., Tiettmeyer, J. M., Chuprinko, B. C., Jordan, S., and Grove, N. P. (2014). Measuring load on working memory: the use of heart rate as a means of measuring chemistry students' cognitive load. *J. Chem. Educ.* 91, 641–647. doi: 10.1021/ed400576n

Creemers, B., Kyriakides, L., and Panayiotis, A. (2012). Teacher Professional Development for Improving Quality of Teaching. New York, USA: Springer Publishing.

Cui, Y., Song, X., Hu, Q., Li, Y., Sharma, P., and Khapre, S. (2022). Human-robot interaction in higher education for predicting student engagement. *Comput. Electr. Eng.* 99:107827. doi: 10.1016/j.compeleceng.2022.107827

Dabbagh, N., and Kitsantas, A. (2004). Supporting self-regulation in student-centered web-based learning environments. *EdMedia + Innovate Learn. Online 2022* 3, 40–47. Available at: https://www.learntechlib.org/primary/p/4104/

Deng, R., Benckendorff, P., and Gannaway, D. (2020). Learner engagement in MOOCs: Scale development and validation. *Br. J. Educ. Technol.* 51, 245–262. doi: 10.1111/bjet.12810

D'Mello, S., Dieterle, E., and Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educ. Psychol.* 52, 104–123. doi: 10.1080/00461520.2017.1281747

Dubovi, I. (2022). Cognitive and emotional engagement while learning with VR: the perspective of multimodal methodology. *Comput. Educ.* 183:104495. doi: 10.1016/j.compedu.2022.104495

Emerson, A., Cloude, E. B., Azevedo, R., and Lester, J. (2020). Multimodal learning analytics for game-based learning. *Br. J. Educ. Technol.* 51, 1505–1526. doi: 10.1111/bjet.12992

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

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., and Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *Int. J. Inf. Manag.* 48, 108–119. doi: 10.1016/j.ijinfomgt.2019.02.003

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: reflections from over 20 years of research. *Educ. Psychol.* 50, 14–30. doi: 10.1080/00461520.2014.989230

Hassib, M., Khamis, M., Friedl, S., Schneegaß, S., and Alt, F. (2017). "Brainatwork: logging cognitive engagement and tasks in the workplace using electroencephalography," in *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM'17)*. New York, NY, USA: Association for Computing Machinery, 305–310.

Helme, S., and Clarke, D. (2001). Identifying cognitive engagement in the mathematics classroom. *Math. Educ. Res. J.* 13, 133–153. doi: 10.1007/BF03217103

Holmes, S. R., Reinke, W. M., Herman, K. C., and David, K. (2021). An examination of teacher engagement in intervention training and sustained intervention implementation. *Sch. Ment. Heal.* 14, 63–72. doi: 10.1007/s12310-021-09457-3

Kang, D., and Oh, S. (2020). Balanced training/test set sampling for proper evaluation of classification models. *Intell. Data Anal.* 24, 5–18. doi: 10.3233/IDA-194477

Ke, F., and Xie, K. (2009). Toward deep learning for adult students in online courses. *Internet High. Educ.* 12, 136–145. doi: 10.1016/j.iheduc.2009.08.001

Kosmyna, N., and Maes, P. (2019). AttentivU: an EEG-based closed-loop biofeedback system for real-time monitoring and improvement of engagement for personalized learning. *Sensors (Basel)* 19:5200. doi: 10.3390/s19235200

Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., and Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. *Br. J. Educ. Technol.* 51, 1548–1562. doi: 10.1111/bjet.12958

Li, S., and Lajoie, S. P. (2021). Cognitive engagement in self-regulated learning: an integrative model. *Eur. J. Psychol. Educ.* 37, 833–852. doi: 10.1007/s10212-021-00565-x

Liao, W., Xu, W., Kong, S., Ahmad, F., and Liu, W. (2019). "A two-stage method for hand-raising gesture recognition in classroom," in *Proceedings of the 2019 8th International Conference on Educational and Information Technology*. 38–44.

Lin, P.-C., Hou, H.-T., Wu, S.-Y., and Chang, K.-E. (2014). Exploring college students' cognitive processing patterns during a collaborative problem-solving teaching activity integrating Facebook discussion and simulation tools. *Internet High. Educ.* 22, 51–56. doi: 10.1016/j.iheduc.2014.05.001

Liu, I. F., Chen, M. C., Sun, Y. S., Wible, D., and Kuo, C.-H. (2010). Extending the TAM model to explore the factors that affect intention to use an online learning community. *Comput. Educ.* 54, 600–610. doi: 10.1016/j.compedu.2009.09.009

Liu, D., and Zhang, H. (2021). Developing a new model for understanding teacher satisfaction with online learning. *SAGE Open* 11:215824402110364. doi: 10.1177/21582440211036440

Mayordomo, R. M., Espasa, A., Guasch, T., and Martínez-Melo, M. (2022). Perception of online feedback and its impact on cognitive and emotional engagement with feedback. *Educ. Inf. Technol.* 27, 7947–7971. doi: 10.1007/s10639-022-10948-2

Moreno-Marcos, P. M., Muñoz-Merino, P. J., Maldonado-Mahauad, J., Pérez-Sanagustín, M., Alario-Hoyos, C., and Delgado Kloos, C. (2020). Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Comput. Educ.* 145:103728. doi: 10.1016/j.compedu.2019.103728

Nami, F. (2021). Developing in-service teachers' pedagogical knowledge of CALL through project-oriented tasks: the case of an online professional development course. *ReCALL* 34, 110–125. doi: 10.1017/s0958344021000148

Nasir, M. K. M., and Ngah, A. H. (2022). The sustainability of a Community of Inquiry in online course satisfaction in virtual learning environments in higher education. *Sustainability* 14:9633. doi: 10.3390/su14159633

Niedermeyer, E., and da Silva, F. L. (2005). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, vol. *99*. Philadelphia, USA: Lippincott Williams & Wilkins, 107827.

Ning, H. K., and Downing, K. (2011). The interrelationship between student learning experience and study behaviour. *Higher Educ. Res. Dev.* 30, 765–778. doi: 10.1080/07294360.2010.539598

Noldus (2019). Available at: https://www.noldus.com/applications/emotion-analysis (Accessed December 1, 2022).

Olsen, J. K., Sharma, K., Rummel, N., and Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. *Br. J. Educ. Technol.* 51, 1527–1547. doi: 10.1111/bjet.12982

Parsons, S., Hutchison, A., Hall, L., Parsons, A., Ives, S., and Leggett, A. (2019). U.S. teachers' perceptions of online professional development. *Teach. Teach. Educ.* 82, 33–42. doi: 10.1016/j.tate.2019.03.006

Philipsen, B., Tondeur, J., Pareja Roblin, N., Vanslambrouck, S., and Zhu, C. (2019). Improving teacher professional development for online and blended learning: a systematic meta-aggregative review. *Educ. Technol. Res. Dev.* 67, 1145–1174. doi: 10.1007/s11423-019-09645-8

Powell, C. G., and Bodur, Y. (2019). Teachers' perceptions of an online professional development experience: implications for a design and implementation framework. *Teach. Teach. Educ.* 77, 19–30. doi: 10.1016/j.tate.2018.09.004

Psaltis, A., Apostolakis, K. C., Dimitropoulos, K., and Daras, P. (2018). Multimodal student engagement recognition in prosocial games. *IEEE Trans. Games* 10, 292–303. doi: 10.1109/TCIAIG.2017.2743341

Richardson, J. C., and Alsup, J. (2015). From the classroom to the keyboard: how seven teachers created their online teacher identities. *Int. Rev. Res. Open Distrib. Learn.* 16, 142–167. doi: 10.19173/irrodl.v16i1.1814

Rogers, P. (2001). Traditions to transformations: the forced evolution of higher education. *AACE J.* 9, 47–60.

Ross, J. D. (2011). *Online Professional Development: Design, Deliver, Succeed!* Thousand Oaks, California: Corwin Press.

Ross, S. M., and Morrison, G. R. (2004). "Experimental research methods," in *Handbook of Research on Educational Communications and Technology*. Mahwah: Lawrence Erlbaum Associates, *Vol. 2*. 1021–1043.

Sharma, K., and Giannakos, M. (2020). Multimodal data capabilities for learning: what can multimodal data tell us about learning? *Br. J. Educ. Technol.* 51, 1450–1484. doi: 10.1111/bjet.12993

Sharma, K., Papamitsiou, Z., Olsen, J. K., and Giannakos, M. (2020). "Predicting learners' effortful behaviour in adaptive assessment using multimodal data," in *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*.

Sinatra, G. M., Heddy, B. C., and Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educ. Psychol.* 50, 1–13. doi: 10.1080/00461520.2014.1002924

Soleymani, M., Asghari-Esfeden, S., Fu, Y., and Pantic, M. (2016). Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Trans. Affect. Comput.* 7, 17–28. doi: 10.1109/TAFFC.2015.2436926

Taub, M., Sawyer, R., Lester, J., and Azevedo, R. (2019). The impact of contextualized emotions on self-regulated learning and scientific reasoning during learning with a game-based learning environment. *Int. J. Artif. Intell. Educ.* 30, 97–120. doi: 10.1007/s40593-019-00191-1

Teräs, H. (2014). Collaborative online professional development for teachers in higher education. *Prof. Dev. Educ.* 42, 258–275. doi: 10.1080/19415257.2014.961094

Terzis, V., Moridis, C., and Economides, A. (2013). Measuring instant emotions based on facial expressions during computer-based assessment. *Pers. Ubiquit. Comput.* 17, 43–52. doi: 10.1007/s00779-011-0477-y

Törmänen, T., Järvenoja, H., and Mänty, K. (2021). Exploring groups' affective states during collaborative learning: what triggers activating affect on a group level? *Educ. Technol. Res. Dev.* 69, 2523–2545. doi: 10.1007/s11423-021-10037-0

Ulmanen, S., Soini, T., Pietarinen, J., and Pyhältö, K. (2016). Students' experiences of the development of emotional engagement. *Int. J. Educ. Res.* 79, 86–96. doi: 10.1016/j.ijer.2016.06.003

Vanneste, P., Oramas, J., Verelst, T., Tuytelaars, T., Raes, A., Depaepe, F., et al. (2021). Computer vision and human behaviour, emotion and cognition detection: a use case on student engagement. *Mathematics* 9:287. doi: 10.3390/math9030287

Wang, Y., Cao, Y., Gong, S., Wang, Z., Li, N., and Ai, L. (2022). Interaction and learning engagement in online learning: the mediating roles of online learning self-efficacy and academic emotions. *Learn. Individ. Differ.* 94:102128. doi: 10.1016/j.lindif.2022.102128

Wang, Y., Chen, N.-S., and Levy, M. (2010). Teacher training in a synchronous cyber face-to-face classroom: characterizing and supporting the online teachers' learning process. *Comput. Assist. Lang. Learn.* 23, 277–293. doi: 10.1080/09588221.2010.493523

Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., and Movellan, J. R. (2014). The faces of engagement: automatic recognition of student engagement from facial expressions. *IEEE Trans. Affect. Comput.* 5, 86–98. doi: 10.1109/TAFFC.2014.2316163

Wigfield, A., Eccles, J., Fredricks, J., Simpkins, S., Roeser, R., and Schiefele, U. (2015). "Development of achievement motivation and engagement," in *Handbook of child psychology and developmental cience:socioemotional processes*. eds. M. E. Lamb and R. M. Lerner (John Wiley & Sons, Inc.) 1–44.

Wolff, C. E., Jarodzka, H., van den Bogert, N., and Boshuizen, H. P. A. (2016). Teacher vision: expert and novice teachers' perception of problematic classroom management scenes. *Instr. Sci.* 44, 243–265. doi: 10.1007/s11251-016-9367-z

Ye, J.-M., and Zhou, J. (2022). Exploring the relationship between learning sentiments and cognitive processing in online collaborative learning: a network analytic approach. *Internet High. Educ.* 55:100875. doi: 10.1016/j.iheduc.2022.100875

Zheng, K., Yang, D., Liu, J., and Cui, J. (2020). Recognition of teachers' facial expression intensity based on convolutional neural network and attention mechanism. *IEEE Access* 8, 226437–226444. doi: 10.1109/ACCESS.2020.3046225

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 January 2023
DOI 10.3389/fpsyg.2022.1042970

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University, China

REVIEWED BY
Yizhou Fan,
The University of Edinburgh,
United Kingdom
Peiyu Wang,
Central China Normal University, China

*CORRESPONDENCE
Yanyan Li
liyy@bnu.edu.cn

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

RECEIVED 13 September 2022
ACCEPTED 15 December 2022
PUBLISHED 17 January 2023

CITATION
Li XR, Hu WQ, Li YY and Mao ZQ (2023) Exploring what synchronized physiological arousal can reveal about the social regulatory process in a collaborative argumentation activity. *Front. Psychol.* 13:1042970.
doi: 10.3389/fpsyg.2022.1042970

Exploring what synchronized physiological arousal can reveal about the social regulatory process in a collaborative argumentation activity

Xiaoran Li[1,2], Wanqing Hu[1], Yanyan Li[1]* and Ziqi Mao[1]

[1]Faculty of Education, Beijing Normal University, Beijing, China, [2]Beijing Advanced Innovation Center for Language Resources, Beijing Language and Culture University, Beijing, China

Combining physiological measures with observational data (e.g., video or self-reports) to further capture and understand the temporal and cyclical process of social regulation has become a trend in the field. Synchronized physiological arousal is a particularly meaningful situation in collaboration. However, little attention has been given to synchronized physiological arousal episodes and their relationship with the social regulatory process. In addition, only a few research utilized heart rate (HR) as a physiological measure in the current collaboration literature. More research is necessary to reveal the potential of HR to expand the diversity of physiological indicators in the field. Therefore, the current study aimed to explore what synchronized physiological arousal can further reveal about the social regulatory process. To achieve this goal, this study designed a collaborative argumentation (CA) activity for undergraduates (mean age 20.3). It developed an arousal-regulation analysis platform, which could automatically detect synchronized physiological arousal in HR and align them with coding challenges and social regulation based on the timeline. In total, 14 four-member groups were recruited. After analyzing both videos and HR data, several findings were obtained. First, only one-third of episodes were synchronized physiological arousal episodes, and the situations where four members were all in arousal states were rare during CA. Second, synchronized physiological arousal was more sensitive to socio-emotional aspects of collaboration as the shared physiological arousal more frequently co-occurred with socio-emotional challenges and socio-emotional regulation, while it happened the least under motivational challenges. Third, synchronized physiological arousal has also been found to be associated with the challenges being regulated. Finally, pedagogical implications were suggested.

KEYWORDS

synchronized physiological arousal, social regulatory process, heart rate, analysis platform, physiological synchrony

1. Introduction

While the expected collaborative learning is a prolific form of learning, it is often the case that learning teams encounter diverse cognitive, socio-emotional, and motivational challenges and function poorly (O'Donnell and Hmelo-Silver, 2013). To succeed in collaboration, there is a need for group members to jointly engage in a quantity of social regulation, which refers to a group deliberately and strategically taking control of tasks through shared and negotiated regulation of cognitive, behavioral, motivational, and emotional conditions (Hadwin et al., 2018a). Understanding how social regulation occurs is meaningful for conducting high-quality collaboration. Existing investigations have been successful in identifying the important patterns of social regulation from single-channel data, for example, videos (Järvenoja et al., 2019) and chat logs (Su et al., 2018). While considering the multidimensional and complex nature of social regulation, there is an increased focus on combining physiological measures with observational data (e.g., video or self-reports) to further capture and understand the temporal and cyclical process of social regulation (Järvelä et al., 2019a; Malmberg et al., 2019a).

Recent studies utilized physiological synchrony (PS) to explore how group members' similar simultaneous changes in physiological signals [e.g., electrodermal activity (EDA), heart rate (HR)] are related to the social regulatory process (Järvelä et al., 2019a). For example, Dindar et al. (2020) unpacked that the groups' PS index derived from EDA is only positively related to their self-report cognitive regulation rather than behavioral, motivation, and emotional regulation. Haataja et al. (2018) indicated that the high values of PS (derived from EDA) co-occurred with frequent monitoring behavior (coded from videos). From that, these investigations attempted to reveal the sharedness and the invisible social process reflected by physiological measures during the social regulatory process (Palumbo et al., 2017; Dindar et al., 2022). However, the prominent PS measures in these studies combined the correspondence between the signals of interacting individuals during both physiological arousal and non-arousal episodes (Dindar et al., 2022). As proposed by many researchers, physiological arousal is a particularly desirable goal in learning as it accounts for learners' cognitive and/or affective activation and is more associated with active participation (Pijeira-Díaz et al., 2018). The PS measures that do not distinguish between physiological arousal and non-arousal episodes may conceal the more meaningful moments of collaboration (Haataja et al., 2018; Malmberg et al., 2019b). Therefore, researchers call for more attention to synchronized physiological arousal episodes and their relationship with social regulation (Malmberg et al., 2019a; Pijeira-Díaz et al., 2019; Dindar et al., 2022). In addition, only a few research utilized HR to profile PS in the current literature. More research is necessary to reveal the potential of HR to measure interpersonal physiology so as to expand the diversity of physiological indicators in the field (Järvelä et al., 2019a).

The current study attempts to fill in the gaps by utilizing HR to profile the synchronized physiological arousal and exploring how the synchronized physiological arousal is related to the social regulatory process in a collaborative argumentation (CA) activity. As for studying the social regulatory process, prior research has identified two important perspectives that can be drawn upon for this study, that is, encountering challenges and social regulation focuses (Su et al., 2018; Järvenoja et al., 2019; Li et al., 2022). Social regulation is multidimensional, including the regulation of cognitive, behavioral, motivational, and emotional conditions (Dindar et al., 2020), and encountering challenges are critical moments to provoke social regulation (Järvenoja et al., 2019). Preliminary research has indicated that PS is positively related to cognitive regulation (Dindar et al., 2020) and monitoring behavior (Haataja et al., 2018). Little is still known about the relationship between interpersonal physiology and encountering challenges. Further investigation combining physiological measures with observational data on the encountering challenges and social regulation focuses could shed light on this multifaced process. Therefore, this study uses HR to profile synchronized physiological arousal and explores the differences in synchronized physiological arousal when groups (a) face different types of challenges (e.g., cognitive, socio-emotional, and motivational challenges) and (b) engage in different social regulation focuses (e.g., metacognitive vs. socio-emotional regulation). In addition, the association between the challenge being regulated and synchronized physiological arousal is also investigated. From that, the study aims to unpack what synchronized physiological arousal can reveal about the social regulatory process. To achieve the goal, this study designed a CA activity titled "is the fast diet a healthy way to lose weight?" for undergraduates and developed an arousal-regulation analysis platform in which synchronized physiological arousal is detected through HR and is aligned with challenge and social regulation behavior coding from videos based on the timeline.

2. Literature review

2.1. Encountering challenges and social regulation focus on collaborative learning

High-quality collaborative learning is hard to achieve for the reason that learning teams could encounter diverse cognitive, socio-emotional, and motivational challenges

(Näykki et al., 2014). Researchers have strived to uncover what challenges students encounter during the collaboration for better understanding and supporting this process (Järvenoja et al., 2019). For example, Koivuniemi et al. (2017) interviewed 107 first-year higher education students about their challenge experience in collaborative learning situations. They found that students encountered diverse challenges in collaboration, for example, concentration (motivational challenge), lacking prior knowledge (cognitive challenge), tiredness (wellbeing challenge), and frustration (emotion challenge). Recently, encountering challenges has become an important perspective in studying the social regulatory process because they are regarded as critical moments to provoke social regulation behavior (Järvelä et al., 2019b). Researchers focus on what social regulation is triggered by encountering challenges. Järvenoja et al. (2019) videotaped a 6-week mathematics course for 62 higher education students and explored how students activate group-level socio-emotional regulation in the face of diverse challenges. Through building a process model, they found that socio-emotional regulation only occurred after emotional and motivational or social context challenges. In a 2-month multimedia course for 103 teacher education students, Malmberg et al. (2015) collected the self-report data from a Virtual Collaborative Research Institute (VCRI) learning environment. They investigated how groups' social regulatory processes progressed as collaboration developed. In the process model, they revealed that the strategies adopted by high-performing groups to regulate challenges shifted from cognitive regulation to emotional regulation, while the low-performing groups stagnated at cognitive regulation.

Apart from encountering challenges, the social regulation focus is another prominent perspective to examine the social regulatory process. Some investigations trace the process according to social regulation focus, such as cognitive, emotional and motivational regulation, and unpack important patterns contributing to a successful collaboration (Järvelä et al., 2019b). Su et al. (2018) analyzed the chat logs from Tencent QQ generated in a semester-long online collaborative language learning course. They conducted sequential analysis and found that the high-performing undergraduate groups demonstrated more significant sequential links between socio-emotional regulations than the low-performing groups. In Ucan and Webb's (2015) investigation, they examined primary students' utterances from a 7-week science course, and by comparing the frequency, they indicated that students not only engaged in metacognitive regulation but can also actively regulate emotional and motivational states to maintain effective group functioning. Overall, the two important perspectives identified in previous studies, encountering challenges and social regulation focus, provide the basis for this study to explore the social regulatory process.

2.2. Physiological synchrony and physiological arousal in the social regulatory process

Although, as shown in previous studies, data from one channel (e.g., video or chat logs) can provide valuable information about the social regulatory process in terms of strategic adaptation in the face of challenges and social regulation focus, the advances in technology and new data-capturing devices offer novel ways to examine and understand the role of these processes across learning contexts, age groups, and tasks (Järvelä et al., 2019a). Studies indicate that in interactive and collaborative learning, an individual learners' physiological activity could be dependent upon group members (Dindar et al., 2020). This interpersonal physiology, which is defined as any interdependent or associated activity identified in the physiological processes of two or more individuals (Palumbo et al., 2017), can reflect invisible social processes co-occurring with observable interactions (Malmberg et al., 2019a). Therefore, researchers have explored how PS, one prominent indicator of interpersonal physiology, is related to the social regulatory process, and preliminary findings were obtained. For example, Dindar et al. (2019) investigated the relationship between monitoring behavior and PS between the collaborating group members. They found that the relationship between PS and shared monitoring might be dependent on task type. That is, a significant relationship was observed in one session, whereas no significant relationship was observed in the other session. In Sobocinski et al.'s (2020) research, they explored the group-level physiological state transitions during collaboration. They found that the group-level physiological state transitions were positively correlated with on-track sequences.

However, some researchers proposed that these PS measures that do not distinguish between physiological arousal and non-arousal episodes may hide more meaningful moments of collaboration (Malmberg et al., 2019a; Pijeira-Díaz et al., 2019; Dindar et al., 2022). In general, arousal can be described as a state of physiological wakefulness with emotional reactivity, enhanced cognitive processing, and increased motor activation (De Lecea et al., 2012; Critchley et al., 2013). Previous empirical research has proven that physiological arousal is positively related to a learner's achievement (Pijeira-Díaz et al., 2018) and increased mental effort (Malmberg et al., 2019b). Therefore, the synchronized arousal of two or more collaborating members, referring to synchronized physiological arousal, is a particularly meaningful situation and is informative of collaboration (Malmberg et al., 2019a; Pijeira-Díaz et al., 2019; Dindar et al., 2022). The scant research that has been conducted only recently has begun to explore synchronized physiological arousal during the social regulatory process. In Malmberg et al.'s (2019a) exploratory research, they investigated how the simultaneous arousal between two or more group members revealed social regulation. Their findings were that most of the collaborative interaction during simultaneous arousal was low level, and social regulation was not observed. However, when the interaction was high level, and social regulation was present; when the interaction was confused, it included monitoring behavior. This investigation provides preliminary insight into the meaningfulness of synchronized physiological arousal in social regulation. More explorations are still needed to reveal what additional valuable information the synchronized physiological arousal can provide about the social regulatory process. Furthermore, previous research mainly utilized EDA to profile learners' PS, and only a few research adopted HR (e.g., Sobocinski et al., 2020, 2022). More research is necessary to reveal the potential of HR to measure interpersonal physiology so as to expand the diversity of physiological indicators in the field (Järvelä et al., 2019a).

2.3. Using heart rate to profile synchronized physiological arousal

Although previous research has successfully utilized HR to profile PS in collaborative learning (e.g., Sobocinski et al., 2020, 2022), little study has adopted HR to profile synchronized physiological arousal in a collaborative learning context. A major approach to detecting physiological arousal is measuring responses of the autonomic nervous system (ANS), which consists of sympathetic and parasympathetic branches, and primarily serves a regulatory function by helping the body adapt to internal and environmental demands (Kreibig, 2010; Roos et al., 2021). Heart rate (HR) and electrodermal activity (EDA) are popular measures derived from the ANS (American Psychiatric Association, 2000). Several advantages of using ANS response in educational settings have been established. First, it is difficult for individuals to mask or control the ANS reactions, thereby creating the possibility of more objectively gauging learner's arousal compared to self-reports (Pijeira-Díaz et al., 2019). In addition, the continuous data resulting from ANS response allow for temporal and dynamic analysis of emotional and cognitive processes of learning, some of which are even executed outside of learners' awareness (Mendes, 2009).

Heart rate (HR) is a primary indicator of ANS response (Kreibig, 2010; Roos et al., 2021). Research in the psychophysiology field has confirmed that increases in HR reactivity may reflect arousal and have been used as popular measures for detecting psychological alertness in facing stimulus (Kreibig, 2010; Kreibig et al., 2012; Critchley et al., 2013). Mason et al. (2018) traced learners' HRs when they read web pages about the health risks of mobile phone use with different reliability, for example, personal blog, online magazine, and academic journal (all including 420 words). They treated the increased HR as a state of arousal. Then they explored how learners' arousal in the HR, while reading was correlated with post-reading comprehension scores. Likewise, in the investigation of Shalom et al. (2015), they attempted to explore differences in physiological arousal between high- and low-social anxiety participants when they engaged in different communication situations (computer-mediated vs. face-to-face communication). Participants' HRs under the two conditions were collected. Arousal was identified when participants' HR was higher than baseline, and Shalom et al. (2015) found that participants with different levels of social anxiety experienced different arousals in two situations. Previous studies that use HR to profile physiological arousal at the individual level hereby provide a basis for further exploring synchronized arousal in groups.

Therefore, the current research attempts to fill in the gaps by utilizing HR to profile the synchronized physiological arousal and exploring how the synchronized physiological arousal is related to the social regulatory process in terms of two important perspectives identified in previous research, encountering challenges and social regulation focus. The following four specific research questions (RQs) guide this study:

RQ1: How often do group member's synchronized physiological arousal happen during the CA?

RQ2: What are the differences in synchronized physiological arousal episodes when groups face different types of challenges (e.g., cognitive, socio-emotional, and motivational challenges)?

RQ3: What are the differences in synchronized physiological arousal episodes when groups engage in different social regulation focuses (e.g., metacognitive vs. socio-emotional regulation)?

RQ4: How does synchronized physiological arousal relate to the challenge being regulated?

3. Materials and methods

3.1. Participants

A total of 56 (mean age 20.3 years, 51 females) undergraduates from a normal university in Beijing, China were recruited to involve in a CA activity titled "Is the fast diet a healthy way to lose weight?" To recruit the participants, the researchers posted announcements on the school's online forum and offered a free lunch ticket for participation. Participants primarily majored in education and foreign language, and their standpoints on the fast diet (whether the fast diet is a healthy way to lose weight? Yes or no) were investigated through an online questionnaire before the activity. Then, the participants were randomly assigned into 14 four-member groups according to their standpoints, which consisted of two supporters and two opponents of the fast diet in each. All participants signed the Ethical Consent Form before the activity, representing that they were informed of the research purpose, confidentiality, and right to withdraw from the study.

3.2. Collaborative argumentation activity

This study was carried out in CA context. CA is a productive form of collaborative learning, which has been proven to deepen students' understanding of complex subject concepts (Wecker and Fischer, 2014) and improve critical thinking (Ngajie et al., 2020) through a set of interactions to convince others of their arguments' validity (Krummheuer, 1995). As the task required, students with contrary standpoints on the fast diet participated in the activity titled "Is the fast diet a healthy way to lose weight?," aiming to reach a consensus through CA and complete a group argument diagram (**Figure 1**), which includes claim (final claim, supporting claim, and counterclaim), evidence, and rebuttal elements (Toulmin, 1958). This face-to-face activity was supported by an online collaborative diagramming platform.[1] Group member used their computers to log in to the platform and synchronously edited the shared group argument diagram. Before CA, participants were trained in drawing argument diagrams and given 15 min to draw an individual argument diagram based on their existing standpoints in which students could read providing material about the nutrients and prepare for CA. Afterward, a 45-min CA was conducted in which students persuaded opponents, reached a consensus, and finished the group argument diagram (for the procedure, see **Figure 2**).

3.3. Data collection

Videos and participants' HRs were collected in this study. The activity took place in a classroom-like research space with 360-degree cameras; therefore, the process of CA was videotaped for each group. Scosche Rhythm + armband HR monitor (Valencell, Inc., Raleigh, NC, USA) was used to record each participant's HR continuously and unobtrusively during the whole activity. The Scosche Rhythm + is a precision biometrics apparatus to monitor HR and has a dual-mode processor in which the HR data can simultaneously transmit to multiple ANT + displays or Bluetooth-enabled devices. The former was chosen as more stable for offline processing, given our exploratory intentions. Before the activity, the Scosche Rhythm + armband HR monitor has installed on each participant's non-dominant arm, and the sampling frequency of the sensor is 1 Hz (one sample per second). After excluding the groups with incomplete HR data, nine groups with 36 participants were finally included in this study. The total duration of valid CA was approximately 7.2 h (mean 48.6 min, SD 4.1 min).

1 https://www.processon.com

3.4. Data analysis

3.4.1. Video analysis

The purpose of video analysis was to identify the challenges and social regulatory behavior during CA. Videos of nine groups were analyzed by Nvivo 12 software, and the analysis unit was 30-s episodes. This time-based segmentation allowed a temporally unfolding overview of the group situations and provided a manageable and consistent unit of analysis (Järvenoja et al., 2019). There were two steps in coding the videos. The first step was to code the challenges and social regulatory behavior separately. In the second step, if a challenge is regulated by certain social regulatory behavior, an additional code as "challenge being regulated" will be given to the challenge; otherwise, a code as "challenge not being regulated" will be given.

Specifically, two coding schemes were utilized to code challenges and social regulatory behavior. For challenges, the coding scheme was adopted from Hadwin et al. (2018b), in which they identify five broad types of challenges that students encounter in various collaboration settings, namely, motivational, socio-emotional, cognitive, metacognitive, and environmental challenges. The current study integrated the metacognitive challenge into the cognitive one because it occurred less frequently (**Table 1**). For social regulation, the coding scheme referred to previous research by Ucan and Webb (2015), in which they have defined and conceptualized two main social regulation categories as metacognitive regulation and emotional and motivational regulation. The current study combined emotional and motivational regulation into socio-emotional regulation for the reason that they served a similar function in maintaining a productive socio-emotional climate in this investigation (**Table 2**). Notably, 25% of data from the CA videos were randomly chosen for inter-coding. Two researchers of this study coded the video independently, and Cohen's kappa coefficient was calculated to judge the inter-rater reliability of the coded variables. The kappa values of challenges, social regulation, and challenges whether being regulated were 0.931, 0.864, and 0.917, respectively.

3.4.2. Heart rate analysis

To analyze HR data and align it with video data on the timeline, the current study developed an arousal-regulation analysis platform with two main functions, namely, identifying synchronized physiological arousal episodes in HR and aligning it with coding challenge or social regulation based on the timeline. Using the platform, two types of files needed to be prepared, separate HR files (downloaded from the HR receiver) of each member in a group and the group's video coding file. The timeline of the two types of files should be aligned. After uploading the two types of files, the platform can automatically identify the synchronized physiological arousal episodes within a group and correspond them with the group's coding challenge

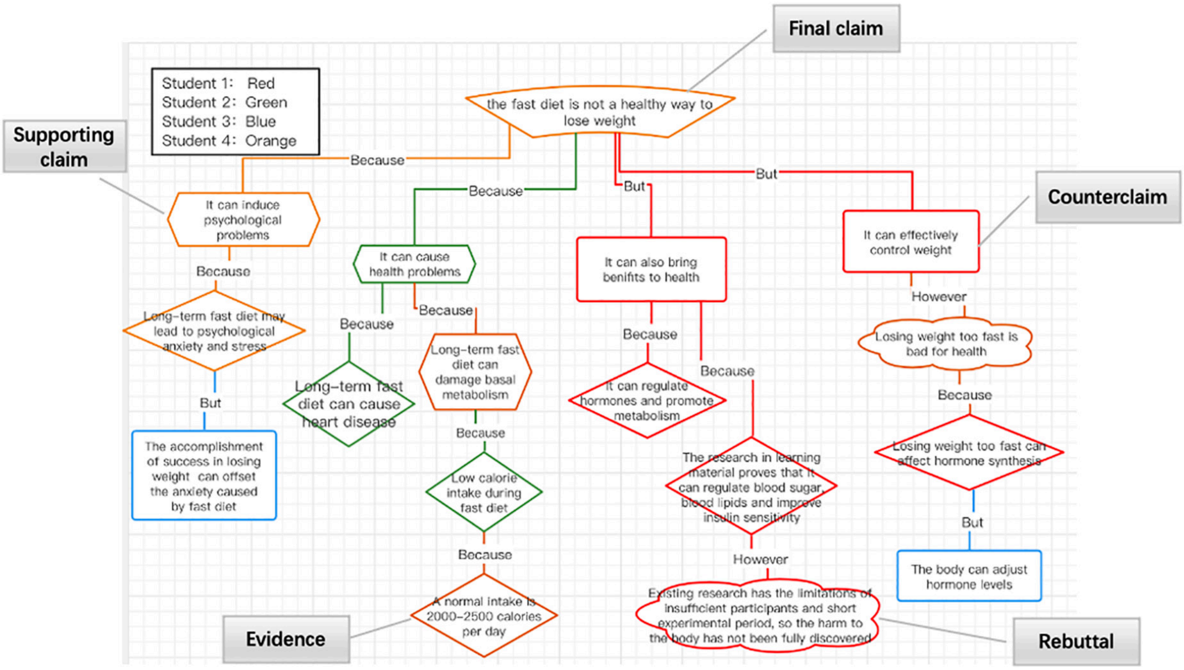


FIGURE 1
Group argument diagram.

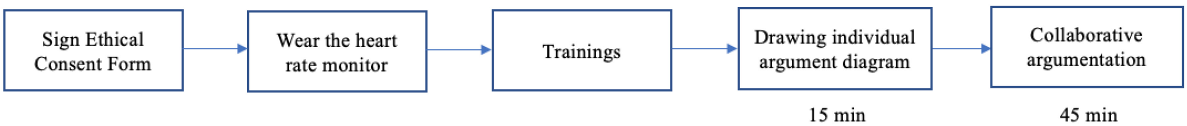


FIGURE 2
Experimental procedure.

TABLE 1 Coding scheme for challenges.

Coding category	Description	Examples
Cognitive challenge	Cognitive challenges refer to difficulties in achieving a shared understanding of the task and domain, or in choosing effective solution paths and strategies.	I lack the knowledge to judge which arguments to follow. I highly doubt the credibility of this evidence. I think our discussion just now was meaningless because you guys didn't get our points. We're totally stuck, and no one can convince anyone.
Motivational challenge	Motivational challenges revolve different personal priorities, self-efficacy or different participation levels. Typically, these challenges result in declines in effort, engagement or participation.	Some students played with mobile and didn't participate in the activity for a while. A: I still don't understand the differences between the counterclaim and rebuttal. B: Whatever, it doesn't matter.
Socio-emotional challenge	Socio-emotional challenges refer to challenges in creating and maintaining a positive climate, such as relational problems associated with achieving psychological safety, communicating effectively, and navigating power relationships.	I'm really speechless. We are totally not in the same mind. . .Oh my, I'm out of ideas. I quit. I think you are very ridiculous to require me to provide all the literature? I'm not here to teach you how to do a literature review.
Environmental challenge	Environmental challenges related to external conditions surrounding collaborative work such as technology or physical discomfort caused by environments.	I tried many times, but I don't know how to add "because" on the edge. The room is too hot to breathe.

TABLE 2 Coding scheme for social regulation.

Coding category	Description	Examples
Metacognitive regulation	Group members jointly enact various behavior pertaining to planning, monitoring, task-specific strategies using and evaluation to regulate the cognitive aspects of the learning process.	How about you two write the supporting ideas and we conclude the opposing ones. How much time is left? It seems that there are only 15 min. We need to draw the diagram immediately. Hey, let's discuss this issue from another angle. The content of our argument diagram is quite rich.
Socio-emotional regulation	Group members jointly enacted various behaviors to manage emotional states and promote motivation, alternatively, to maintain a productive socio-emotional climate.	From my body shape, you guys could tell that I have no right to talk about losing weight. Could you talk more about your opinion? I agree with your points. They are very reasonable.

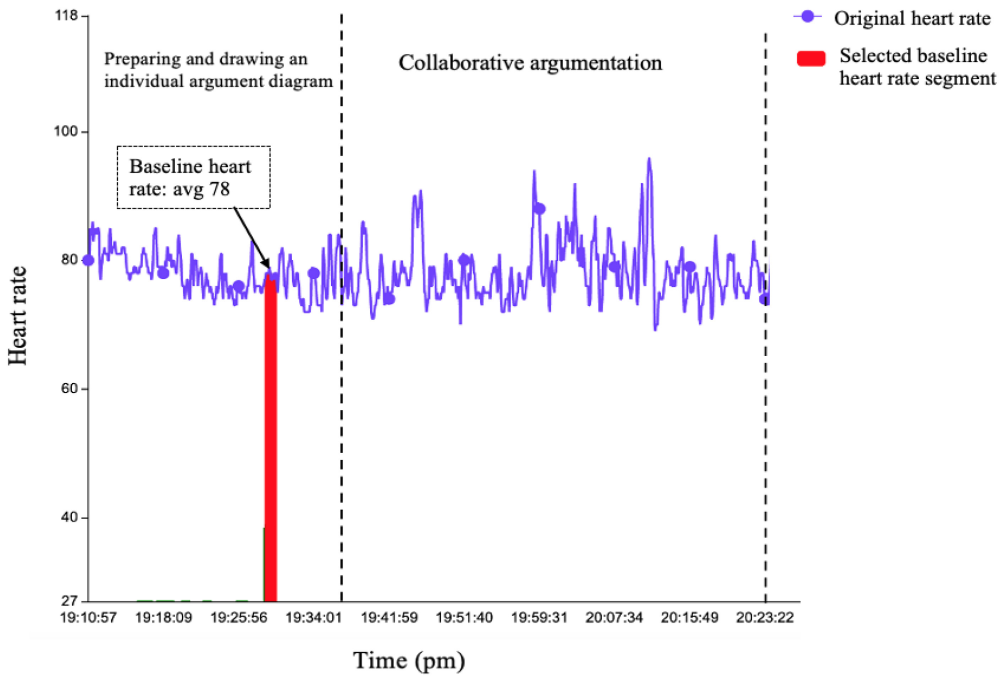


FIGURE 3
Computing individual student's baseline heart rate (HR) [from 19:10:57 to 19:36:52 p.m., one student in group 4 was at the first stage as reading material and drawing the individual argument diagram, and from 19:36:52 to 20:23:22 p.m., she engaged in collaborative argumentation (CA). The baseline HR is 78, the average score of the selected baseline HR segment].

or social regulation based on the timeline. While providing visual presentations, this platform can also output result files for further statistical analysis. The realization of the key steps in the platform will be elaborated on next.

3.4.2.1. Identifying synchronized physiological arousal episodes

Before acquiring synchronized physiological arousal episodes in a group, each student's arousal episodes needed to be identified first. Based on previous research in the psychophysiology field (Mendes, 2009; Shalom et al., 2015; Mason et al., 2018), the key point to identifying individuals' arousal was to define baseline HR and find the episodes in which the HR was higher than baseline during the CA. Therefore, to automatically obtain an individual's baseline HR, this study regarded the first stage of activity, that is, 15 min of reading material and drawing an individual argument diagram, as the data set for finding baseline HR because participants completed the task calmly and independently without any interference at this stage. The HR data were calculated using a moving window approach, with a window width of 1 min and a moving step of 1 s, the sampling interval of the sensor, to find the segment with the smallest standard deviation and longest duration, which represented a more stable HR and was closer to the resting HR. The average HR within the selected segment was regarded as the baseline HR. One-third of the participants were randomly chosen to verify the selected baseline HR with their actual resting HR, and the accuracy rate was 93.8% (**Figure 3**).

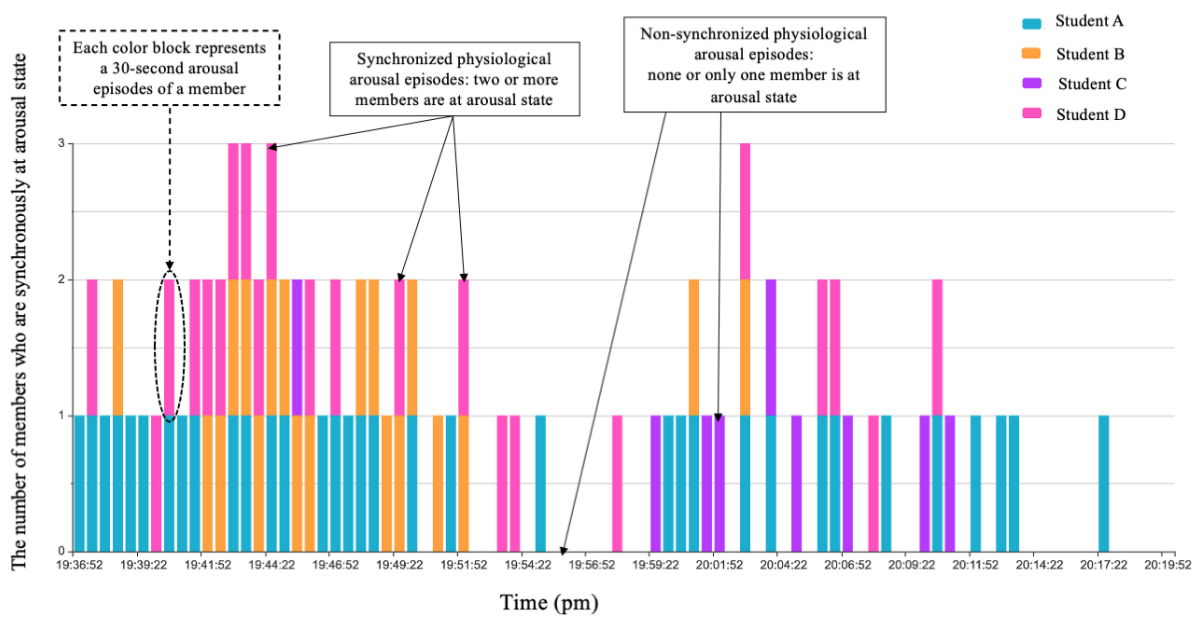


FIGURE 4
Identifying synchronized physiological arousal episodes within a group based on timeline [group 4 had four members and engaged in collaborative argumentation (CA) from 19:36:52 to 20:23:22 p.m. The *x*-axis represents time, and the *y*-axis represents the number of members who are synchronously in an arousal state. The episodes with superimposed color blocks are the synchronized physiological arousal episodes].

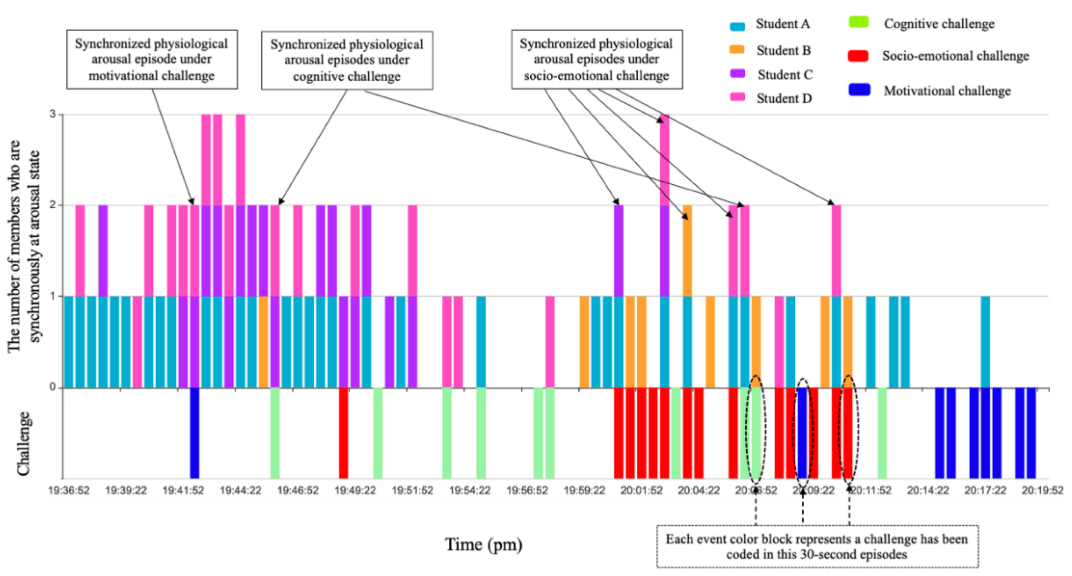


FIGURE 5
How group members' physiological arousal synchronized under different challenges (example of group 4. The episodes where two or more arousal color blocks and event color blocks are superimposed are synchronized physiological arousal episodes under challenges).

After determining the baseline HR, the next step was to find the individual arousal episodes during CA in which the HR was higher than the baseline. Since the study design included events that unfolded over time and there were specific time-locked events of interest, that is, coding challenges or social regulation in each 30-s episode (Mendes, 2009; Malmberg et al., 2019a), the platform also divided the individual HR data during CA into 30-s segments corresponding to the video. If the average HR within 30-s segments was higher than baseline, segments were identified as arousal episodes. Each participant's arousal episodes were thus obtained. After aligning all members' arousal episodes according to the timeline, the synchronized

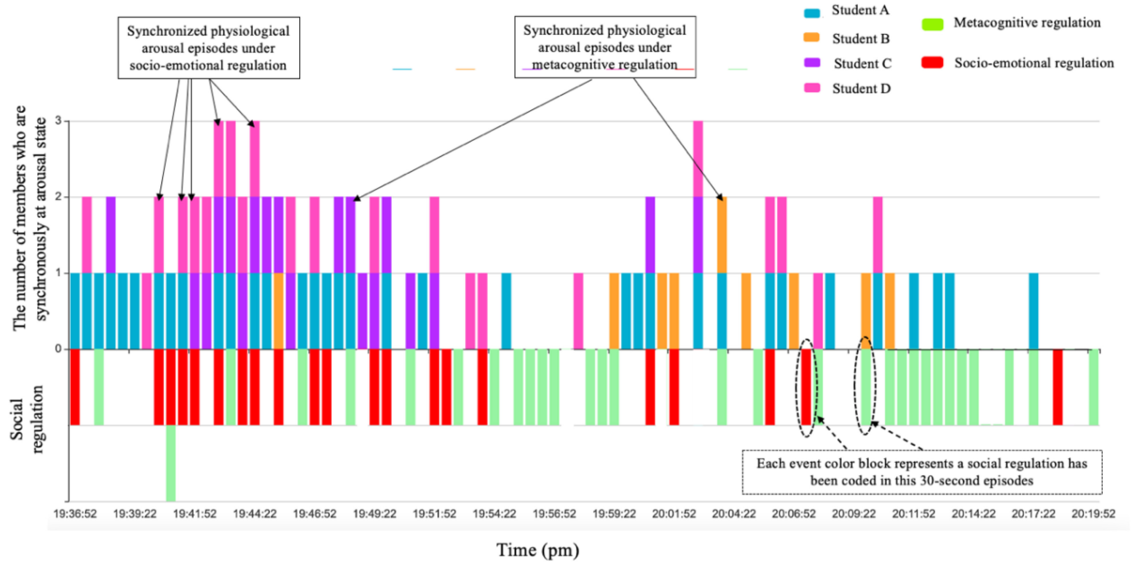


FIGURE 6
How group members' physiological arousal synchronized under different social regulations (example of group 4. The episodes where two or more arousal color blocks and event color blocks are superimposed are synchronized physiological arousal episodes under social regulation).

physiological arousal episodes of a group were acquired in which two or more members were synchronously in an arousal state (Figure 4).

3.4.2.2. Aligning it with coding challenges or social regulation

Since the segmented episodes of HR were consistent with the video, the last key step was to further align the group's arousal episodes with coding challenges or social regulation to reveal how group members' physiological arousal in HR synchronized under different conditions in terms of facing diverse challenges (cognitive, socio-emotional, motivational, and environmental challenges) (Figure 5) as well as engaging in different social regulation focuses (metacognitive vs. socio-emotional regulation) (Figure 6).

Following previous research, this study regarded episodes in which two or more members were in the arousal state as synchronized physiological arousal episodes (Palumbo et al., 2017; Malmberg et al., 2019a). Although the visualization in the platform could provide an intuitive representation of a group, the platform also outputted the result file to answer the research question. The resulting file included information about what challenge or social regulation was coded as well as the number of members who are synchronously in an arousal state in each 30-s episode during CA. To answer RQ1, the descriptive statistics of frequency and percentage of episodes with different arousal numbers (from none to four members) during the CA were presented. For RQ2 and RQ3, the frequency and percentage of episodes with different arousal numbers under different challenges and social regulations were first given. Then, the chi-square test was utilized to analyze the differences in the frequency of synchronized physiological arousal episodes when faced with different challenges and engaged in different social regulation focuses. Similarly, the RQ4 chi-square test was utilized to examine how does synchronize physiological arousal relates to the challenge being regulated. To ensure the accuracy of the results in RQ2–RQ4, episodes containing multiple challenges or social regulation codes were excluded.

TABLE 3 Frequency and percentage of episodes with different arousal numbers (from none to four members) during collaborative argumentation (CA) activity.

	None	1	2	3	4	Total episodes
Frequency of episodes	255	354	225	31	2	867
Percentage	29.41%	40.83%	25.95%	3.58%	0.23%	100.00%

4. Results

4.1. How often does group members' synchronized physiological arousal happen during collaborative argumentation?

Since the overall valid data were approximately 7.2 h, there were in total 867 30-s episodes calculated during CA of nine groups. Table 3 presents the frequency and percentage of episodes with different arousal numbers (from

TABLE 4 Frequency and percentage of challenges episodes with different arousal numbers (from none to four members).

	None		1		2		3		4		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
CC	15	31.3%	22	45.8%	10	20.8%	1	2.1%	0	0.0%	48	42.1%
SC	7	16.7%	16	38.1%	15	35.7%	4	9.5%	0	0.0%	42	36.8%
MC	12	54.5%	8	36.4%	2	9.1%	0	0.0%	0	0.0%	22	19.3%
EC	1	50.0%	1	50.0%	0	0.0%	0	0.0%	0	0.0%	2	1.8%
Total	35	30.7%	47	41.2%	27	23.7%	5	4.4%	0	0.0%	114	100.0%

TABLE 5 Differences in synchronized physiological arousal episodes frequency under different challenges.

	Non-synchronized physiological arousal episode		Synchronized physiological arousal episode		χ^2	*P*
	Frequency	%	Frequency	%		
CC	37	77.08%	11	22.92%	10.560	0.005
SC	23	54.76%	19	45.24%		
MC	20	90.91%	2	9.09%		

none to four members). It provides an overall picture of how group members' physiological arousal synchronized during collaborative argumentation. As shown in Table 3, the majority of episodes (609, 70.24%) were non-synchronized physiological arousal episodes (none or 1), and only one-third of episodes (258, 29.76%) were synchronized physiological arousal episodes (two or more). Episodes with four members who were all in arousal states were extremely rare (2, 0.23%).

4.2. What are the differences in synchronized physiological arousal episodes when groups face different types of challenges (e.g., cognitive, socio-emotional, and motivational challenges)?

Table 4 shows the frequency and percentage of challenge episodes with different arousal numbers. During collaborative argumentation, there were 114 challenge episodes, with the cognitive challenge (CC, 48, 42.1%), socio-emotional challenge (SC, 42, 36.8%), and motivational challenge (MC, 22, 19.3) occurring more frequently, while environmental challenges were rare (EC, 2, 1.8%). Therefore, environmental challenges were excluded in the following analysis.

The chi-square test further indicates that there was a significant difference in synchronized physiological arousal episodes frequency when faced with diverse challenges (χ^2 = 10.560, df = 2, p = 0.005). That is, when faced with socio-emotional challenges, the synchronized physiological arousal situations occurred the most (19, 45.24%), followed by cognitive challenges (11, 22.92%), and finally motivational challenges (2, 9.09%) (Table 5).

4.3. What are the differences in synchronized physiological arousal episodes when groups engage in different social regulation focuses (e.g., metacognitive vs. socio-emotional regulation)?

Table 6 shows the frequency and percentage of social regulation episodes with different arousal numbers. Overall, there were 444 social regulation episodes during CA, and the distribution of socio-emotional regulation (SR, 212, 47.7%) and metacognitive regulation (MR, 232, 52.3%) was relatively balanced.

The chi-square test further indicates that there was a significant difference in synchronized physiological arousal episodes frequency when engaged in different social regulation focuses (χ^2 = 7.250, df = 2, p = 0.007) as the synchronized physiological arousal happened more frequently when engaged in socio-emotional regulation (78, 36.79%) rather than in metacognitive regulation (58, 25%) (Table 7).

4.4. How does synchronized physiological arousal relate to the challenge being regulated?

Of 112 challenges, more than half challenges were regulated (63, 56.3%). The chi-square test further indicates that there was

TABLE 6 Frequency and percentage of social regulation episodes with different arousal numbers (from none to four members).

	None		1		2		3		4		Total	
	f	%	f	%	f	%	f	%	f	%	f	%
SR	57	26.9%	77	36.3%	68	32.1%	10	4.7%	0	0.0%	212	47.7%
MR	74	31.9%	100	43.1%	51	22.0%	7	3.0%	0	0.0%	232	52.3%
Total	131	29.5%	177	39.9%	119	26.8%	17	3.8%	0	0.0%	444	100.0%

TABLE 7 Differences in synchronized physiological arousal episodes frequency under different social regulations.

	Non-synchronized physiological arousal episode		Synchronized physiological arousal episode		χ^2	*P*
	Frequency	%	Frequency	%		
SR	134	63.21%	78	36.79%	7.250	0.007
MR	174	75.00%	58	25.00%		

a significant difference in synchronized physiological arousal episodes frequency between the challenge being regulated and not being regulated (χ^2 = 4.444, df = 2, p = 0.035). This is, in challenges being regulated episodes, more synchronized physiological arousal situations occurred (23, 35.48%) rather than in challenges not being regulated episodes (9, 18.37%) (**Table 8**).

5. Discussion

The current research aims to utilize HR to profile the synchronized physiological arousal and further explore how the synchronized physiological arousal is related to the social regulatory process in terms of two important perspectives identified in previous research, encountering challenges and social regulation focus. Toward that end, this study developed an arousal-regulation analysis platform, which could automatically detect synchronized physiological arousal episodes in HR and align them with coding challenges and social regulation based on the timeline. After applying it in a CA activity, several findings were obtained.

First, during the CA, only one-third of the episodes were synchronized physiological arousal episodes, and the situations where four members were all in arousal states were rare. Similar findings can be found in Pijeira-Díaz et al.'s (2019) study, in which the arousal was manifested by the EDA signal. They found that only in a small part of the time (≈5–40% of the lesson), the triad members were at the same arousal levels, and the time triad members were simultaneously in high arousal was rare. Similar findings not only verify that synchronized physiological arousal is hard to achieve in collaboration (Pijeira-Díaz et al., 2018) but also prove that the methods using HR to profile arousal in this study can also be an effective approach to detecting synchronized physiological arousal.

Second, the synchronized physiological arousal occurred differently under diverse challenges. Both cognitive and socio-emotional challenges were dominant challenges encountered by students, and when faced socio-emotional challenges, synchronized physiological arousal happened more often. Some recent studies highlighted the importance of socio-emotional challenges in collaboration for the reason that, unlike cognitive challenges, socio-emotional challenges could affect the group climate (Bakhtiar et al., 2018) and, if not appropriately regulated, they may lead to a negative climate, which is harmful and unworkable for accomplishing the shared goal (Isohätälä et al., 2018). This study further signifies the importance of socio-emotional challenges by providing more objective evidence from physiological measures as collaborating members expressed more synchronized physiological responses to socio-emotional challenges (Järvelä et al., 2019a). On the contrary, synchronized physiological arousal occurred the least under motivational challenges. This result could be explained by previous findings that the unequal participation of team members is often overlooked because students are more concerned with completing tasks rather than other members' engagement (Li et al., 2021). This finding extends previous research by further indicating that this disregard can also be reflected in students' physiological signals as members showed the least synchronized physiological response to motivational challenges.

Third, although students exhibited balanced socio-emotional and metacognitive regulation behavior, synchronized physiological arousal more frequently co-occurred with socio-emotional regulation. This finding contradicted Dindar et al.'s (2020) investigation, which found that the overall physiological synchrony index of a team was only related to students' self-report cognitive regulation, with no relationships found with emotional regulation. This inconsistency may be due to

TABLE 8 Differences in synchronized physiological arousal episodes frequency between the challenge being regulated and the challenge not being regulated.

	Non-synchronized physiological arousal episode		Synchronized physiological arousal episode		χ^2	*P*
	Frequency	%	Frequency	%		
Challenge being regulated	40	64.52%	23	35.48%	4.444	0.035
Challenge not being regulated	40	81.63%	9	18.37%		

the different methods of data collection and processing. As suggested by previous research, utilizing physiological data to provide temporal information about collaboration in this study allowed for a more fine-grained analysis of temporal dynamics and patterns of social regulatory processes, thereby revealing more nuanced discoveries (Järvelä et al., 2019a; Dindar et al., 2022). Furthermore, combining video and physiological measures may more objectively reflect learners' emotional and cognitive states in contrast to subjective self-reporting data (Roos et al., 2021). In addition, linking the findings of research questions 2 and 3, it can be unearthed that the shared physiological responses of collaborating members are more sensitive to the socio-emotional aspects of collaboration as the shared physiological arousal more frequently co-occurred with socio-emotional challenges and socio-emotional regulation. This result is also an improved answer to reflect the importance of socio-emotional aspects in collaboration. As revealed in previous investigations, students show an inability to regulate increasing tension (Sohr et al., 2018) and view the negative socio-emotional climate as a more difficult factor to control (Rogat and Linnenbrink-Garcia, 2011). The findings of this study not only reinforce the stance that relations matter in collaboration with physiological data (Järvelä and Rosé, 2022) but also suggests that shared physiological responses can be a signal of involvement in socio-emotional aspects of collaboration (Järvelä et al., 2019a).

The last significant finding is that the challenge being regulated episodes more frequently co-occurred with synchronized physiological arousal, which indicates a correlation between synchronized physiological responses with the challenges being regulated. Encountering challenges are a meaningful moment in collaborative learning because, if not appropriately regulated, the challenges could be detrimental to collaboration (Hadwin et al., 2018b). Although previous research has argued that the lack of regulation in responding to challenges is normally due to students' unawareness of challenges (Järvelä et al., 2019b), there is little evidence, possibly because data on this underlying process are hard to capture. The preliminary finding of the current study can be a piece of evidence by revealing that students showed less synchronized physiological responses to the challenges not being regulated.

6. Conclusion and pedagogical implications

The current study developed an automatic arousal-regulation analysis platform, which contributes to the existing methods of studying the social regulatory processes by successfully utilizing HR to profile synchronized physiological arousal. After applying it in a CA activity, some important characteristics were revealed, which could provide a better understanding and facilitate superior scaffolding of the social regulatory process. For example, only one-third of episodes were synchronized physiological arousal episodes, and they were more associated with the challenges being regulated. Therefore, scripts or awareness tools are needed to help students better identify challenges and become more aware of their occurrence to respond to and regulate them promptly. In addition, the least synchronized physiological arousal under the motivation challenge should attract educators' notice for the reason that this disregard for equal participation could lead to poor group performance in the long term (Li et al., 2021). Instructional intervention, such as directly emphasizing the value of equal participation, should be implemented to guide students to pay more attention to group members' engagement. Finally, synchronized physiological arousal is more sensitive to socio-emotional challenges as well as socio-emotional regulation. This finding not only reinforces the importance of socio-emotional aspects in collaboration but also calls for support to help students cope with socio-emotional challenges and facilitate socio-emotional regulation, thereby maintaining a favorable socio-emotional atmosphere.

7. Limitations and future work

The main limitation of this study is the sample size, uneven gender ratio, and the unitary learning context. As physiological measures can be affected by contextual changes (Järvelä et al., 2019a), the preliminary findings, as well as the function of the arousal-regulation analysis platform, need to be further validated in a larger, more gender-balanced sample and applied in more contexts. Furthermore, to answer the research question, the present study used 30 s as the analysis unit of HR to align with the videos and explore synchronized physiological arousal under different events. Future research can first identify the synchronized physiological arousal episodes without considering the video and then analyze the behavioral characteristics under these episodes in reverse. Finally, while arousal has been shown to be positively related to an individual learners' performance (Pijeira-Díaz et al., 2018), and there is also a consensus on the value of synchronized physiological arousal in collaboration (Dindar et al., 2022), empirical research on the relationship between synchronized physiological arousal and group performance is still scarce.

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 research involved human participants and followed the procedures of the Education Faculty's Ethics Committee. Informed consent was obtained for all participants.

Author contributions

XL led the research project, analyzed the data, and wrote the manuscript. WH conducted the data collection and analysis. YL supervised the research and revised the manuscript. ZM supported the instructional process. All authors read and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 62277006) and the International Joint Research Project of the Faculty of Education of Beijing Normal University (Grant No. ICER201903).

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

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Virginia: American Psychiatric Association.

Bakhtiar, A., Webster, E., and Hadwin, A. (2018). Regulation and socio-emotional interactions in a positive and a negative group climate. *Metacogn. Learn.* 13, 57–90. doi: 10.1007/s11409-017-9178-x

Critchley, H. D., Eccles, J., and Garfinkel, S. N. (2013). "Interaction between cognition, emotion, and the autonomic nervous system," in *Handbook of Clinical Neurology*, eds R. M. Buijs and D. F. Swaab (Amsterdam: Elsevier), 59–77. doi: 10.1016/B978-0-444-53491-0.00006-7

De Lecea, L., Carter, M. E., and Adamantidis, A. (2012). Shining light on wakefulness and arousal. *Biol. Psychiatry* 71, 1046–1052. doi: 10.1016/j.biopsych.2012.01.032

Dindar, M., Alikhani, I., Malmberg, J., Järvelä, S., and Seppänen, T. (2019). Examining shared monitoring in collaborative learning: a case of a recurrence quantification analysis approach. *Comput. Hum. Behav.* 100, 335–344. doi: 10.1016/j.chb.2019.03.004

Dindar, M., Järvelä, S., Nguyen, A., Haataja, E., and Çini, A. (2022). Detecting shared physiological arousal events in collaborative problem solving. *Contemp. Educ. Psychol.* 69, 1–13. doi: 10.1016/j.cedpsych.2022.102050

Dindar, M., Malmberg, J., Järvelä, S., Haataja, E., and Kirschner, A. P. (2020). Matching self-reports with electrodermal activity data: investigating temporal changes in self-regulated learning. *Educ. Inform. Technol.* 25, 1785–1802. doi: 10.1007/s10639-019-10059-5

Haataja, E., Malmberg, J., and Järvelä, S. (2018). Monitoring in collaborative learning: co-occurrence of observed behavior and physiological synchrony explored. *Comput. Hum. Behav.* 87, 337–347. doi: 10.1016/j.chb.2018.06.007

Hadwin, A. F., Järvelä, S., and Miller, M. (2018a). "Self-regulation, co-regulation, and shared regulation in collaborative learning environments: a social cognitive perspective," in *Handbook of Self-Regulation of Learning and Performance*, eds D. H. Schunk and J. A. Greene (London: Routledge), 83–106. doi: 10.4324/9781315697048-6

Hadwin, A. F., Bakhtiar, A., and Miller, M. (2018b). Challenges in online collaboration: effects of scripting shared task perceptions. *Int. J. Comput. Support. Collaborative Learn.* 13, 301–329. doi: 10.1007/s11412-018-9279-9

Isohätälä, J., Näykki, P., Järvelä, S., and Baker, J. M. (2018). Striking a balance: socio-emotional processes during argumentation in collaborative learning interaction. *Learn. Cult. Soc. Interact.* 16, 1–19. doi: 10.1016/j.lcsi.2017.09.003

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., and Kirschner, A. P. (2019a). What multimodal data can tell us about the students' regulation of their learning process? *Learn. Instr.* 72, 1–10. doi: 10.1016/j.learninstruc.2019.04.004

Järvelä, S., Järvenoja, H., and Malmberg, J. (2019b). Capturing the dynamic and cyclical nature of regulation: methodological progress in understanding socially shared regulation in learning. *Int. J. Comput. Support. Collaborative Learn.* 14, 425–441. doi: 10.1007/s11412-019-09313-2

Järvelä, S., and Rosé, P. C. (2022). Relations matter – CSCL research informing and developing CL competencies. *Int. J. Comput. Support. Collaborative Learn.* 17, 185–189. doi: 10.1007/s11412-022-09373-x

Järvenoja, H., Näykki, P., and Törmänen, T. (2019). Emotional regulation in collaborative learning: when do higher education students activate group level regulation in the face of challenges? *Stud. High. Educ.* 44, 1747–1757. doi: 10.1080/03075079.2019.1665318

Koivuniemi, M., Panadero, E., Malmberg, J., and Järvelä, S. (2017). Higher education students' learning challenges and regulatory skills in different learning situations. *Infancia Aprendizaje* 40, 19–55. doi: 10.1080/02103702.2016.1272874

Kreibig, D. S. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010

Kreibig, S. D., Gendolla, G. H. E., and Scherer, K. R. (2012). Goal relevance and goal conduciveness appraisals lead to differential autonomic reactivity in emotional responding to performance feedback. *Biol. Psychol.* 91, 365–375. doi: 10.1016/j.biopsycho.2012.08.007

Krummheuer, G. (1995). "The ethnology of argumentation," in *The Emergence of Mathematical Meaning: Interaction in Classroom Cultures*, eds P. Cobb and H. Bauersfeld (Mahwah, NJ: Erlbaum), 229–269.

Li, X. R., Li, Y. Y., Hu, W. Q., Li, K. R., and Gao, L. (2022). More socio-emotional regulation, more effective? Exploring social regulation of learning in collaborative argumentation among the high and low performing groups. *Metacogn. Learn.* doi: 10.1007/s11409-022-09329-4 [Epub ahead of print].

Li, Y. Y., Li, X. R., Zhang, Y., and Li, X. (2021). The effects of a group awareness tool on knowledge construction in computer- supported collaborative learning. *Br. J. Educ. Technol.* 52, 1178–1196. doi: 0.1111/bjet.13066

Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X. H., and Siipo, A. (2019a). Going beyond what is visible: what multichannel data can reveal about interaction in the context of collaborative learning? *Comput. Hum. Behav.* 96, 235–245. doi: 10.1016/j.chb.2018.06.030

Malmberg, J., Haataja, E., Seppänen, T., and Järvelä, S. (2019b). Are we together or not? The temporal interplay of monitoring, physiological arousal and physiological synchrony during a collaborative exam. *Int. J. Comput. Support. Collaborative Learn.* 14, 467–490. doi: 10.1007/s11412-019-09311-4

Malmberg, J., Järvelä, S., Järvenoja, H., and Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: progress of socially shared regulation among high-and low-performing groups. *Comput. Hum. Behav.* 52, 562–572. doi: 10.1016/j.chb.2015.03.082

Mason, L., Scrimin, S., Zaccoletti, S., Tornatora, C. M., and Goetz, T. (2018). Webpage reading: psychophysiological correlates of emotional arousal and regulation predict multiple-text comprehension. *Comput. Hum. Behav.* 87, 317–326. doi: 10.1016/j.chb.2018.05.020

Mendes, W. B. (2009). "Assessing autonomic nervous system activity," in *Methods in Social Neuroscience*, eds E. Harmon-Jones and J. S. Beer (New York, NY: Guilford Press), 118–147.

Näykki, P., Järvelä, S., Kirschner, A. P., and Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning—a process-oriented case study in a higher education context. *Int. J. Educ. Res.* 68, 1–14.

Ngajie, N. B., Li, Y., Tiruneh, T. D., and Cheng, M. (2020). Investigating the effects of a systematic and model-based design of computer-supported argument visualization on critical thinking. *Think. Skills Creat.* 38, 1–15.

O'Donnell, M. A., and Hmelo-Silver, E. C. (2013). "What is collaborative learning? An overview," in *The International Handbook of Collaborative Learning*, eds C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, and A. M. O'Donnell (London: Routledge), 15–28.

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., et al. (2017). Interpersonal autonomic physiology: a systematic review of the literature. *Pers. Soc. Psychol. Rev.* 21, 99–141.

Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., and Kirschner, A. P. (2019). Sympathetic arousal commonalities and arousal contagion during T collaborative learning: how attuned are triad members? *Comput. Hum. Behav.* 92, 188–197.

Pijeira-Díaz, H. J., Drachsler, H., Kirschner, P. A., and Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: how active are students? *J. Comput. Assist. Learn.* 34, 397–408.

Rogat, T. K., and Linnenbrink-Garcia, L. (2011). Socially shared regulation in collaborative groups: an analysis of the interplay between quality of social regulation and group processes. *Cogn. Instr.* 29, 375–415.

Roos, L. A., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., et al. (2021). Test anxiety and physiological arousal: a systematic review and meta-analysis. *Educ. Psychol. Rev.* 33, 579–618. doi: 10.1007/s10648-020-09543-z

Shalom, G. J., Israeli, H., Markovitzky, O., and Lipsitz, D. J. (2015). Social anxiety and physiological arousal during computer mediated vs. face to face communication. *Comput. Hum. Behav.* 44, 202–208.

Sobocinski, M., Järvelä, S., Malmberg, J., Dindar, M., Isosalo, A., and Noponen, K. (2020). How does monitoring set the stage for adaptive regulation or maladaptive behavior in collaborative learning? *Metacogn. Learn.* 15, 1–29. doi: 10.1007/s11409-020-09224-w

Sobocinski, M., Malmberg, J., and Järvelä, S. (2022). Exploring adaptation in socially-shared regulation of learning using video and heart rate data. *Technol. Knowl. Learn.* 27, 385–404. doi: 10.1007/s10758-021-09526-1

Sohr, R. E., Gupta, A., and Elby, A. (2018). Taking an escape hatch: managing tension in group discourse. *Sci. Educ.* 102, 883–916.

Su, Y., Li, Y. Y., Hu, H., and Rose, P. C. (2018). Exploring college English language learners' self and social regulation of learning during wiki-supported collaborative reading activities. *Int. J. Comput. Support. Collaborative Learn.* 13, 35–60.

Toulmin, S. (1958). *The Uses of Argument*. Cambridge, MA: Cambridge University Press.

Ucan, S., and Webb, M. (2015). Social regulation of learning during collaborative inquiry learning in science: how does it emerge and what are its functions? *Int. J. Sci. Educ.* 37, 2503–2532.

Wecker, C., and Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Comput. Educ.* 75, 218–228.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 18 January 2023
DOI 10.3389/fpsyg.2022.1093475

Check for updates

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University,
China

REVIEWED BY
Hafizhah Zulkifli,
National University of Malaysia, Malaysia
Muhammad Younas,
Soochow University,
China
Yuhua Wang,
Manchester Metropolitan University,
United Kingdom

*CORRESPONDENCE
Wenchao Xue
✉ wincherx@126.com

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

RECEIVED 09 November 2022
ACCEPTED 28 December 2022
PUBLISHED 18 January 2023

CITATION
Zhao X and Xue W (2023) From online to offline education in the post-pandemic era: Challenges encountered by international students at British universities. *Front. Psychol.* 13:1093475. doi: 10.3389/fpsyg.2022.1093475

From online to offline education in the post-pandemic era: Challenges encountered by international students at British universities

Xin Zhao and Wenchao Xue*

Information School, University of Sheffield, Sheffield, United Kingdom

Background: After 2 years of anti-pandemic struggles, universities in the United Kingdom have started to witness a reverse transition, a shift from online to offline education. This includes encouraging students to begin face-to-face programmes and allowing flexibility for remote learners, but later requiring all students to return to campus by a certain date.

Objectives: This paper aims to explore the challenges and impacts brought about by this new transition and provide recommendations for universities to enhance student experience for future adversity.

Method: This qualitative study conducted semi-structured interviews with 24 international students from a British university to explore their experiences during the transition. The results were analysed using thematic analysis.

Results: Our data revealed both internal and external challenges to students during the online-to-offline shift, which lead to a general resistance to said shift. Specifically, policy challenges (e.g., policy conflicts) imposed the most significant impacts on international students, resulting in psychological anxiety, financial losses, and negative learning experiences. The reduction of digital tools and learning materials during the shift also presented challenges to students who developed a reliance on digital resources while learning remotely. Other challenges have also been identified, including academic barriers and social engagement issues.

Conclusion: By highlighting these challenges, this paper has practical implications for university policy decisions and provides recommendations for supporting students' transition back to traditional offline learning.

KEYWORDS

online-offline transition, international student experience, post-pandemic, student transition, learning experience

1. Introduction

Worldwide educational institutions have been affected by the outbreak of the COVID-19 pandemic. According to a report from the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2021), over 220 million students were impacted by this unprecedented incident. The impact included social-distancing measures, lecture cancellations and teaching facility closures (Marinoni et al., 2020). Consequently, campus-based education activities were moved to online or hybrid environments, creating various challenges for students (Lemay et al., 2021). After over 2 years of anti-pandemic struggles, many countries have implemented less severe public health and social measures (WHO, 2022). Since the summer of 2021, the UK government announced its anti-COVID strategy, 'Living with COVID-19' (Cabinet Office, 2022). Following this national policy, the Department of Education of the United Kingdom started to encourage higher education institutions to return to delivering face-to-face teaching (Department of Education, 2022). As a result, most universities in the United Kingdom switched to face-to-face teaching, making students readjust to a reverse transition from online to offline learning, causing new challenges for students, especially those with international backgrounds. However, most research on student transitions related to the pandemic and post-pandemic education has focused on challenges during online transitions (Saikat et al., 2021; Gupta et al., 2022; Szopiński and Bachnik, 2022). A limited amount of research has been conducted on the reverse transition, in which students switched from online to offline learning.

This research aims to examine the experiences of Chinese international students during the post-pandemic transition from online to offline learning and asks the following questions:

- What challenges and constraints do Chinese international students face during the online-offline transition?
- What suggestions and improvements can be made to enhance the student experience for future adversity?

The project used a research university (S University) in the United Kingdom as an example before exploring the impact of the online-to-offline transition on a group of Chinese international students ($n = 24$). During the first half of the academic term, S University adopted a hybrid teaching approach, encouraging students to return to campus for face-to-face teaching while providing online alternatives for those temporarily based outside of the United Kingdom. Once students arrived on campus, they were required to attend face-to-face classes. During the second half of the academic term, S University cancelled all online teaching and completely shifted their courses offline. Students who could not reach the campus on time were required to take a Leave of Absence (LOA) and postpone their studies for a year. Depending on when these students arrive on campus, they may face different challenges during this reverse transition. Accordingly, this research defines the following student groups:

- Campus Starter (CS): Students who arrived at the university before the term started and attended the face-to-face courses (as encouraged).
- Remote Starter (RS): Students who arrived at the university after a period of distance learning and attended face-to-face courses before the start of the second half of the academic term.
- Leave of Absence Students (LOA): Students who completed the first half of the academic term remotely, but could not arrive at the university in time for the second half of the academic year for in-person teaching. They were placed on 'leave of absence' and had to postpone their studies for a year.

2. Literature review

2.1. Issues associated with online teaching during the pandemic

As a preventive measure to curb the spread of the coronavirus, strict social restrictions have been adopted by many countries, such as keeping in place social distancing, remote work, and regional lockdowns (Kaur et al., 2021). As a result, educational institutions around the world have been heavily impacted and forced to adopt a distance-learning approach (Gillis and Krull, 2020; Lederman, 2020; Zhao et al., 2020). However, research suggests that there are many concerns associated with the sudden shift to online teaching, such as technical issues, student disengagement, and time zone differences, which may cause social and technical challenges for both teachers and students (Fatoni et al., 2020; Lemay et al., 2021; Naddeo et al., 2021). For example, Daniel (2020) argues that instructors were unprepared for online teaching, due to a lack of training and support from universities. Naddeo et al. (2021) found that students' negativity is related to Internet connectivity and the performance of learning management systems (LMS). Similarly, Song et al. (2004) found that, in addition to connectivity issues, the students' experience can be negatively influenced by their unfamiliarity with the learning system. Research also reports a lack of peer interaction and student-teacher interaction online, which are both crucial for learning (Kamble et al., 2021). Furthermore, distance learning can also be a challenge for students who study academic subjects that require lab work (Dhawan, 2020; Radha et al., 2020). These issues may influence students' academic outcomes and overall experiences (Omar et al., 2021).

2.2. Opportunities for online teaching

Despite the challenges, the remote-learning environment has also brought about new educational opportunities (Hoss et al., 2022; Li et al., 2022). For example, Azorín (2020) argues that the crisis provides scarce opportunities to trial, improve, and rethink the role, content, and innovative delivery methods of education.

Meanwhile, Aguilera-Hermida (2020) believes that the remote-learning experience also promoted the development of digital literacy among students. Daniel (2020) argues that asynchronous learning, particularly suitable in digital formats, provides both teachers and students with flexibility and enables them to achieve a better balance between work and study. Moreover, online teaching also promotes flexibility while reducing expenses (Xhaferi and Xhaferi, 2020; Weldon et al., 2021). In addition to flexibility, researchers argue that online platforms and technology applications provide students with additional resources (Pokhrel and Chhetri, 2021; Sbaffi and Zhao, 2022) and enhance student academic achievements (Younas et al., 2022). Similarly, Fatoni et al. (2020) found that students can benefit from asynchronous materials (e.g., lecture recordings) to enhance their knowledge. There has also been an increase in the use of social media for sharing learning resources with students (Nasution et al., 2022). As Huang et al. (2021) point out, COVID-19 has created opportunities for universities and brings out innovative pedagogy and digital resources. However, it is uncertain how sustainable these benefits could be when universities start to transition back to 'ordinariness' and offer only face-to-face courses to students (Daniel, 2020).

2.3. Online to offline transition

Although COVID-19 is still spreading, coexistence with the virus has become an option for many countries, including the United Kingdom (Cabinet Office, 2022). Scholars, such as Hargreaves (2020), believe that temporary online education will move back towards the traditional campus-based model or a hybrid method. Meanwhile, Rashid and Yadav (2020) suggest that online teaching should remain an essential part of future teaching even when universities return to offline teaching.

In 2022, we have witnessed a shift to offline learning in the United Kingdom, as predicted by scholars (Daniel, 2020). As this is a new trend, there has been little research about the issues and challenges faced by students during this reverse transition. Among all university students, international learners suffer more challenges caused by external and transnational obstacles, such as border control, travel restrictions, flight cancellations, extra financial difficulties, and career issues (Hari et al., 2023). Therefore, this project will focus on the experience of a group of Chinese international students to explore their experiences during this transition.

3. Materials and methods

This research follows an interpretive paradigm and adopts qualitative research with an exploratory nature (Stebbins, 2001; Rudestam and Newton, 2014). It focuses on the experiences of Chinese international students during the online-offline transition to explore their perceptions and attitudes as well as the challenges they faced during this transition. Using a snowball sampling strategy, 24 semi-structured interviews were conducted with various cohorts of students, Campus Starters ($n = 8$), Remote Starters ($n = 8$), and Leave of Absence (LOA) students ($n = 8$). CS students began their on-campus life at the beginning of the first semester, which means they participated in more face-to-face courses than the other participant groups. While RS students travelled to the United Kingdom successively between Sep 2021 and Feb 2022, they engaged with both online and offline learning. Meanwhile, LOA students only experienced online learning and had to postpone their studies for a year. See Table 1 for detailed participant information.

The interviews were conducted online with the help of meeting tools (Google Meet and Tencent Meeting) that enabled Chinese international students outside of the United Kingdom to participate in the interview. Due to the nature of this study, the research was conducted in two languages: English and Chinese (Mandarin). A pilot study was conducted prior to data collection to refine and revise interview questions and enhance the reliability of the research. The interview questions were divided into two sections. As part of the first section, we asked questions about student programmes, study durations, and previous transitional experiences. Depending on students' learning modes (e.g., CS, RS or LOA), the second section contains different questions. Common questions focused on students' decision-making when choosing different learning modes, their overall feeling towards the university's online to offline transition, barriers they faced, specific events or activities that impacted their learning experiences, and their recommendations to the university. A design of the interview questions was submitted as Supplementary material.

The data from the semi-structured interview was then transcribed, translated, and analysed using a thematic analysis approach, which included becoming familiar with the data, generating initial codes, searching for themes, reviewing themes, defining themes, and producing a write-up (Braun and Clarke, 2006). This method is particularly suitable when analysing human experiences and perceptions, which also fits well with the objectives of this article. Specifically, informed by Kiger and Varpio (2020), our data analysis process involved the following steps: (1) familiarise all the interview transcripts and generate a general understanding of the dataset, (2) analyse each transcript in detail and note the initial codes according to the research questions, (3) examine all the initial codes and search for core themes, (4) reviewing the themes, interpreting their interrelationships, and answering the research questions, (5) defining and naming all the themes, identifying the issues, and making suggestions and (6) producing the report based on the themes.

The study received ethics approval from the university of [anonymised]. Informed consent was collected prior to the interviews. In addition, participants were informed of their right to withdraw from the study. The names of all participants have been anonymised. An ethics approval letter can be found in the Supplementary materials.

TABLE 1 Participant information.

Participant ID	Gender	Degree	Duration of online learning	Type of study
1	Female	Master	1 week	Remote start
2	Female	Master	6 months	Remote start
3	Male	Master	N/A	Campus start
4	Male	Master	N/A	Campus start
5	Female	Bachelor	N/A	Campus start
6	Female	Master	6 months	Remote start
7	Female	Master	6 months	LOA
8	Male	Master	6 months	LOA
9	Female	Pre-Master	2 months	Remote start
10	Female	Bachelor	4 months	LOA
11	Female	Bachelor	6 months	LOA
12	Female	Master	N/A	Campus start
13	Female	Bachelor	6 months	LOA
14	Female	Master	6 months	LOA
15	Male	Master	6 months	LOA
16	Male	PhD	N/A	Campus start
17	Female	Master	N/A	Campus start
18	Male	Master	6 months	Remote start
19	Female	Master	6 months	Remote start
20	Female	Master	2 months	Remote start
21	Female	Master	6 months	LOA
22	Female	Master	8 months	Remote start
23	Male	Master	N/A	Campus start
24	Male	Master	N/A	Campus start

4. Results

Our interview data revealed a number of challenges that students faced during the online to offline transition at the university, namely policy, infrastructure, academic and financial challenges.

4.1. Policy-related challenges

Policy issues create enormous difficulties for all types of students (CS, RS and LOA). The identified policy challenges consist of two sub-themes: policy constraints and policy conflicts.

4.1.1. Policy constraints

In the context of the COVID pandemic, policy constraints come from national-level travel restrictions and institutional-level teaching guidelines. Travel restrictions were enormous challenges for international students. During the transition, the RS and LOA students suffered from local lockdowns and strict flight controls, which forced them to be separated from the offline-learning environment. These issues were often unpredictable and left students with limited time to respond. One RS student said that:

> After my flight was suspended, I couldn't come in time and had to buy a new ticket. This was a bit troublesome. The travel procedures are quite complicated. If I didn't do it in advance, I couldn't arrive at the expected time. (Participant 6, RS, PGT)

The travel restrictions imposed a significant influence on students' choices of teaching approaches. Students were unable to shorten their physical distance from campus. Therefore, students had to choose from the available options. In other words, policy constraints deprived them of the opportunity for campus learning. One LOA student complained that:

> I was faced with various quarantine policies, [...] according to the domestic policy, there was no way for me to take public transportation, let alone apply for a visa to study abroad. (Participant 7, LOA, PGT)

Similarly, when asked about the reason for making the online choice, an RS student answered by saying that:

> My flight was suspended at that time. But actually, I wanted to come over. (Participant 6, RS, PGT)

In addition, some students expressed concerns about the challenges they may face after their studies in the United Kingdom were completed. To avoid quarantine upon their return to China, students were hesitant or even resisted the offline shift. One bachelor's student who was unwilling to transition said that:

> [...] if I go to the UK, I need to be prepared to not go home for the next two or three years, because China's strict policy makes it very hard to return home, which I couldn't accept at that time. (Participant 11, LOA, UG)

A master's student who postponed the journey to the United Kingdom shared a similar concern:

> Although the pandemic-management policy in Britain allows for total freedom in 2022, in my region, pandemic control is particularly strict. [...] If I got the coronavirus, it would be very difficult for me to return home, where I would face a long quarantine. So, I didn't want to come to the UK at that time. (Participant 6, RS, PGT)

Apart from their choice of learning methods and travel concerns, the travel restrictions imposed psychological and financial pressure on students, discouraging them from moving to offline learning. For example, one LOA student expressed that:

> [...] There were moments that I wished to go to the UK for on-campus learning. However, the fight tickets were very expensive, which led to a lot of pressure. [...] My flight could easily have been suspended and triggered the circuit cancel policy. [...] It ultimately made me anxious about the delay in my studies. [...] Everyone was extremely anxious. (Participant 15, LOA, PGT)

The institutional-level teaching guideline is another major challenge. The guideline refers to the mandatory offline-learning policy, which the university created without providing prior notice to students. This led to considerable impacts for RS and LOA students. This sudden policy shift also made it imperative for students to prepare for the transition in a short period of time. Data shows that the sudden change demotivated many students from transiting to offline learning. One LOA student said:

> I thought that they should have included online classes as an option in the second semester since the school had made such an inquiry. [...] So, I didn't have, or consider to have my visa, accommodations, and so on prepared. (Participant 11, LOA, UG)

Another identified impact of the shift is the adjustment. This is primarily related to RS students. Universities in the United Kingdom have had resources in place for years to support students' orientation at the beginning of the term. Therefore, CS students received sufficient support when familiarising themselves with offline teaching and the learning environments (e.g., campus facilities). Arriving at the campus halfway through the term meant that RS students needed to spend time and effort adapting to the new environment with less university support compared to CS students. As RS students had to make last-minute arrangements, travel expenses and overseas living costs were sometimes exceptionally high.

> Online classes are [...] very convenient, which saves both time and money. [...] When I first came here, I had to adapt to the life here [...] which might take me a week to adapt, [...] The transition period of that week will affect my courses, which is quite challenging. (Participant 6, RS, PGT)

4.1.2. Policy conflicts

Conflicts exist between travel restrictions and teaching guidelines. Specifically, the university policy requires face-to-face teaching, but the travel policy imposes many restrictions both in China and in the United Kingdom. This mismatch created a practical challenge for international students, especially those in the LOA and RS groups. Some LOA students were under a local lockdown and were forced to take a leave of absence. One LOA student said:

> I was forced to make that choice (leave of absence), [...] Many remote students were forced to go there and study on campus because they didn't want to take a break from college. We were forced to do that. (Participant 11, LOA, UG)

The conflicts between teaching guidelines and quarantine policy in the United Kingdom at the time were also challenging for certain CS students. The quarantine policy blocked face-to-face registration and sessions despite the arrival of CS students in the United Kingdom. The university also prevented them from attending online sessions as they were registered as CS students. One CS student said:

> When I first came to the UK, I was required to stay in quarantine for 14 days. However, I couldn't attend any classes in the Blackboard system until I completed registration offline to receive my student card. [...] This resulted in me missing some of my classes. (Participant 3, CS, PGT)

Conflicts also exist in policies between the university and academic departments. For example, the university provided remote-teaching alternatives during the first semester, but some schools or modules only offered offline exams, resulting in students taking a leave of absence. One LOA student claimed that:

> It was quite sudden that we were told to attend offline exams while students of other majors could take their exams online.

> We thought it was unfair. I even thought the school was way out of line. [...] I got 0 on that module and need to retest. [...] I could have received a good result after studying for so long. [...] I have to go over there next September and start my sophomore year all over again, which is a huge waste of time. (Participant 10, LOA, UG)

In addition to study interruptions and exam delays, students also reported that the sudden transition negatively impacted their career planning and psychology. The interruption of studies has necessitated postponing the original study work plan and introduced uncertainty. The fear of further uncertainty and potentially needing to take a gap year has created additional anxiety for LOA students. The following arguments support these identified impacts:

> My plan was disrupted! My plans for my studies and future career were completely messed up. And there was also a big impact on my life. I was at home by myself during the lockdown. Plus, the leave of absence stuff, I was depressed and not in a good mood [...] I couldn't help feeling anxious at night. (Participant 14, LOA, PGT)
>
> Due to a year of delay, I have concerns in terms of time. Also, I have age anxiety. (Participant 13, LOA, UG)

4.2. Infrastructure challenges

4.2.1. Technology constraint

Technical issues mainly occurred during the first half of the academic term. The university incorporated live-streaming technologies, which allow for face-to-face teaching to be broadcasted to remote students. All the international students experience some online components in the autumn semester, such as lecture recordings and asynchronous materials. Therefore, this paper infers that technical constraints are common challenges for all student groups. Concretely, the technical challenges manifest in unstable networks, unclear audio, and blurry pictures in the live class. These factors mainly affect the synchronisation and efficiency of the live-streaming sessions. When describing technical issues, students said:

> The problem was the network latency. [...] Occasionally, it was very laggy. (Participant 1, RS, PGT)
>
> Sometimes, there is noise in the livestream classes and recordings. This is inevitably with lots of students in a classroom. (Participant 4, CS, PGT)
>
> I need to use a VPN to attend classes, which can be troublesome and lags sporadically. (Participant 10, LOA, UT-Year2)

Students considered these technical issues to be temporary. Therefore, their overall learning outcomes and experiences were not hugely impacted. When asked about this, students reported that:

> It wasn't too bad. [...] I am satisfied with my studies during those six months. (Participant 8, LOA, PGT)
>
> I was quite content with my learning results. (Participant 15, LOA, PGT)

4.2.2. Campus facility constraint

The challenges caused by the inconvenience of campus facilities mainly affect RS students who newly completed the transition. After experiencing online learning, RS students needed to readjust to offline classrooms. Compared to CS students, the campus is new to them. Furthermore, there was no orientation week to support students who were unfamiliar with the campus. This unfamiliarity made them feel that the campus facilities were not easy to access. For example, one student explained:

> I couldn't find the classroom on my first day on campus. [...] The classroom was so far away from my apartment and located on the second floor of a basement. Twisting and winding, it was like a maze. [...] I looked for it for a long time. (Participant 1, RS, PGT)

The facility constraint impacts learning convenience. Preference for convenience makes students reluctant to accept new changes. Therefore, the physical inconvenience of offline facilities and a lack of university support regarding induction made them more unwilling to give up flexibility and easy access to teaching. This is one of the reasons for the resistance to the transition. One student claimed:

> Before the transition, I could sit at home and have classes without running to the campus. [...] The campus of my university is quite big, so it was easy to get lost. (Participant 5, CS, UG)

4.3. Academic challenges

4.3.1. Reduction of digital support

Online teaching facilitates a wealth of digital learning tools and can promote students' digital literacy skills. Apart from some efficient and innovative platforms, including learning management systems (e.g., Blackboard), Cloud-based collaboration tools (e.g., Google), and social media (e.g., YouTube), digital tools, such as recording software, auto-captions, and real-time translation, also played critical roles in improving academic results. Students were immersed in the digital environment for the first half of the academic term and were dependent on the digital tools that supported them in online learning. When they fully transitioned to offline learning, many digital tools became unavailable, which created a massive challenge for students. Almost all RS and CS participants stressed the importance of some digital tools for their learning. Among these tools, the course recording and language assistance software were most valuable to students. Course videos allow students to revisit lectures, while translation tools help them overcome language barriers, contributing to enhanced learning outcomes. Furthermore, international students rely heavily on captions and auto-translation to overcome language barriers and achieve academic progress. According to multiple students:

> I could turn to the recordings at any time, according to my learning progress, if there was anything I didn't figure out or I had missed in offline classes. [...] In the first semester [...] I relied heavily on the recordings and slides of the classes. (Participant 12, CS, PGT)
>
> I could turn to the subtitles when I couldn't understand the lecture. [...] I was able to understand more content in class compared with the second semester. (Participant 2, RS, PGT)
>
> The learning experience is better with online lectures as there are auto-captions that we can use. (Participant 18, RS, PGT)
>
> The chatbot feature allows the tutor to answer all questions one by one. This can benefit all students. Offline lectures can become crowded when students are asking questions. Tutors may not be able to address all of them. (Participant 24, CS, PGT)

Students' reliance on digital tools is directly linked to their learning experiences and outcomes. This reliance can also be arguably linked to students' resistance to offline learning. Face-to-face learning becomes more challenging if students experience language barriers. As one student argued:

> I became reliant on the subtitles when listening to online classes. I would habitually look at the subtitles and not think about what the teacher or my classmates were saying. And, when I switched to on-campus learning, I would find it hard to adapt in terms of listening and other aspects. (Participant 1, RS, PGT)

Students value the digital skills they have developed through online education. Our data suggests that online learning initiates innovative approaches and facilitates the utilisation of technology, such as digital assistance tools, online cooperation software, social media platforms, etc. These innovative and beneficial approaches also enhance the digital skills of students, which is indispensable in the Information Age. Some participants believe that the enhancements brought about by online learning contribute to their future development:

> The informative approach will become a major trend in the future. For instance, working from home or something like this. [...] it would be of great help if you have experience in online classes and are very skilled and seasoned in online information processing, operation, and learning. Because I worked as an intern during the LOA period and there were often online meetings or communication. [...] My previous online experience was valuable during my internship. (Participant 11, LOA, UG).

However, the transition to offline learning has significantly decreased the use of digital tools and materials for learning and teaching. Hence, reliance on digital support also drove some LOA students away from offline learning. One LOA student conveyed the following view:

> I can refer to the recordings, [...] It helps with the absorption and consolidation of knowledge. [...] Is it only in offline classes that you can learn well? Not necessarily. (Participant 14, LOA, PGT)

This challenge also affects the frequency and timing of students' utilisation of IT tools and platforms. For example, with the total shift back to offline learning, the recording and captioning tools are no longer available. Students' time spent on digital resources, such as E-Library and Blackboard, was also reduced.

> I spent more time on Blackboard and reading course materials. [...] I'm more frequently using online tools. [...] Then, during the second semester, [...] I didn't log on to Blackboard very often. (Participant 2, RS, PGT)

4.3.2. Reduced flexibility

Flexibility is a unique feature of online learning. This term describes flexibility in time and location, which could contribute to crisis resilience. Therefore, many students, after experiencing online learning, were reluctant to give up flexibility in teaching. Specifically, the recorded and asynchronous materials allowed students to learn more flexibly. Furthermore, these factors ensure that teaching is unaffected by external contingencies, such as the strike movement that occurred several times during the academic year.

> Last semester, [...] I could just turn on my computer and attend my classes, [...] when I came here, I had to walk a long distance to take the courses, which I found quite tiresome and time-consuming. (Participant 2, RS, PGT)
>
> I wasted my time running to the classroom, which ended up being empty (due to staff strikes), [...] The same thing could happen online, but it probably would not affect me as much. (Participant 3, CS, PGT)

Another reason is related to the previously identified factor: digital support. Lecture contents are easier to understand for international students with the help of digital tools. In addition, after transitioning to offline learning, students need to spend more time and effort on self-study to make up for the information missed during face-to-face lectures due to language barriers. One student points out:

> Because there is no recording, [...] I need to listen attentively, [...] I need to spend more time and energy, such as looking at the courseware again, or reading a lot of relevant materials to make up for some knowledge I missed in class. (Participant 5, CS, UG)

4.4. Financial challenges

The transition from online to offline also increased the total financial cost of education. This issue posed different levels of challenges depending on the cohort of students. For RS and LOA students, the increased cost is significant, especially in the context of COVID. The sudden transition means that students must spend a lot of money on living costs, accommodations, international flights, and COVID-related expenses. One LOA felt particularly stressed and complained by saying:

> A round trip ticket is very expensive, nearly £10,000. It is not a small amount of money for my family. [...] It is hard for me to make a casual choice. (Participant 15, LOA, PGT)
>
> Due to the lockdown, there was a delay in the postal service. As a result, I didn't get my visa and passport returned to me on time, and missed my flight as a result. There was no refund. I had to buy a new flight ticket. (Participant 22, RS, PGT)

These additional significant financial costs led to psychological stress for students and may have led to their decision to postpone their studies. Completing a degree at the minimum possible cost is very attractive to many students. Therefore, the cost challenge is a fundamental factor that drives their reluctance to transition.

> For international students, it requires a huge amount of money to live and study in another city. For those who have one year course, the online option can actually help them save a lot of money for their families. (Participant 11, LOA, UG)

The impact on CS students is less severe. The fact that they were already involved in offline classes means that they had considered the issue of education expenditure and made a choice with sufficient time to plan the trip. The additional high cost, due to COVID-19, is within their expectations.

5. Discussion

This research developed a nuanced understanding of the challenges that influence Chinese international students during the online-to-offline transition. Several challenges these students faced during the transition were identified, including policy issues, infrastructure constraints, academic barriers, and financial burdens. Among all the identified challenges, policy challenges posed the most serious issues for international students, and the impacts were far more severe than those brought about by other challenges. Issues such as academic, social and infrastructure only impact the transition's quality. In contrast, policy issues, especially policy conflicts, may determine the success or failure of some students' transitions. The emergence of LOA students represents the failure of their campus-oriented shift. The findings evidenced that LOA students were the most vulnerable among all students, as this group faced interruptions in their studies, delays in career planning, and enormous psychological pressure. This accords the finding of Sahu (2020), who suggests that policies, such as travel restrictions bring a series of secondary challenges, including delays in examinations, monetary problems, admission issues, study interruption, psychological health, and career plans.

All participants displayed resistance to a complete transition from online to offline learning. However, the resistance was more evident among the RS and LOA groups. This could be because these groups of students were unwilling to accept the challenges brought upon them by a sudden shift to offline learning. Another noticeable and common reason for resistance is the reliance on digital tools. Online learning tools are most effective in assisting with language barriers. With the help of recording, captioning, and live translation tools, participants reported significantly improved learning outcomes in lectures and eventually became reliant on them (Fatoni et al., 2020; Pokhrel and Chhetri, 2021). This is consistent with the finding of Weldon et al. (2021), who argue that lecture recordings significantly benefit students during the revision and assessment stages. Moreover, many efficient and innovative technologies were introduced during the online period, such as screensharing during group discussions or using learning materials from social media. Students also benefit from more frequent use of digital tools, as their digital skills are enhanced. This finding accords with the research conducted by Aguilera-Hermida (2020), who argues that the online-oriented transition promotes the utilisation of technology and digital literacy. Similarly, Noor et al. (2022) also highlights the benefits of information and computer technology (ICT) skills development for students as a result of a rapid growth of technology in education during the pandemic.

Based on the findings, this research has the following recommendations. This research suggests that a complete and sudden switch to offline teaching is not a perfect solution. Advance notice should be given to students at the offer stage rather than forcing a sudden change upon students midway through their studies. Lei and So (2021) argue that students are sensitive to the sudden switch between different learning modalities. In addition to considering national policies, universities in the United Kingdom should also take into account international policies that may affect international students' travels and studies. In addition, universities should implement a consistent policy across all academic departments to minimise the impact on students who study for joint degrees or attend modules offered by more than one department. In addition, the value of digital tools and resources should be acknowledged and sustained to enhance the student learning experience, academic outcomes, and the development of digital literacy. Our finding supports the view of previous research, which highlights the value of a blended teaching approach and sustained digital resources for teaching in a post-pandemic age (Hargreaves, 2020; Rashid and Yadav, 2020). Finally, universities should consider enhancing their resources to support the transition of those students returning from a leave of absence. Students from this group are often the most vulnerable and underrepresented. In the current system, students are most likely to receive support during the orientation or induction week. Dedicated support staff or resources could be made available for students who miss induction weeks, such as late arrivals and students who return from a leave of absence.

6. Conclusion

This research aims to study the challenges encountered by international students during the reverse transition from online to offline teaching in the post-pandemic era. The findings identified challenges faced by Chinese international students during this transition, namely, policy issues, infrastructure constraints, academic barriers, and financial burdens. First, the policy issues are most challenging for students resulting in a broad range of negative impacts. Our data suggests that LOA students are the most vulnerable to study interruptions, career plan disruptions, and psychological pressure. More support should be provided to this cohort when they resume their studies. Second, a consistent teaching method across the academic year is most valued by students. Almost all participants displayed negativity and resistance towards a sudden move to offline learning. Sudden changes in teaching can cause considerable disruptions to students. Thus, advanced notice should be provided to students if such sudden changes are unavoidable. Third, the complete reversion to the traditional face-to-face model is not a perfect option. Online modality has unique and irreplaceable benefits, promoting innovation and enhancing students' digital skills. Universities should retain the digital resources and pedagogy developed during the pandemic and incorporate them into offline teaching.

This research also presents several limitations. Due to the research scope, the subject is limited to Chinese international students. Particular challenges brought about by the policies in China are not representative of the entire international student community. Future studies may select a broader range of subjects. Further studies could also look at the readjustment issues faced by LOA students who have been found to be most vulnerable in this research. In addition, the perception of the participants may change over time. A longitudinal study is needed to determine the long-term impact of the transition on 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 University of Sheffield Information School 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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1093475/full#supplementary-material

References

Aguilera-Hermida, A. P. (2020). College students' use and acceptance of emergency online learning due to COVID-19. *Int. J. Educ. Res. Open* 1. doi: 10.1016/j.ijedro.2020.100011

Azorín, C. (2020). Beyond COVID-19 supernova. Is another education coming? *J. Prof. Capital Commun.* 5, 381–390. doi: 10.1108/JPCC-05-2020-0019

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Cabinet Office (2022). COVID-19 response-living with COVID-19. HM Government. Available at: https://www.gov.uk/government/publications/covid-19-response-living-with-covid-19/covid-19-response-living-with-covid-19 (Accessed August 25, 2022).

Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects* 49, 91–96. doi: 10.1007/s11125-020-09464-3

Department of Education (2022). Higher education COVID-19 operational guidance. Available at: https://www.gov.uk/government/publications/higher-education-reopening-buildings-and-campuses#full-publication-update-history (Accessed August 25, 2022).

Dhawan, S. (2020). Online learning: a panacea in the time of COVID-19 crisis. *J. Educ. Technol. Syst.* 49, 5–22. doi: 10.1177/0047239520934018

Fatoni, A. N., Nurkhayati, E., Nurdiawati, E., Fidziah Pamungkas, G., Adha, S., Irawan Purwanto, A., et al. (2020). University students online learning system during Covid-19 pandemic: advantages, constraints and solutions. *Syst. Rev. Pharm.* 11, 570–576.

Gillis, A., and Krull, L. M. (2020). COVID-19 remote learning transition in spring 2020: class structures, student perceptions, and inequality in college courses. *Teach. Sociol.* 48, 283–299. doi: 10.1177/0092055X20954263

Gupta, R., Aggarwal, A., Sable, D., Chahar, P., Sharma, A., Kumari, A., et al. (2022). Covid-19 pandemic and online education: impact on students, parents and teachers. *J. Hum. Behav. Soc. Environ.* 32, 426–449. doi: 10.1080/10911359.2021.1909518

Hargreaves, A. (2020). What's next for schools after coronavirus? Here are 5 big issues and opportunities. Available at: https://theconversation.com/whats-next-for-schools-after-coronavirus-here-are-5-big-issues-and-opportunities-135004 (Accessed August 25, 2022).

Hari, A., Nardon, L., and Zhang, H. (2023). A transnational lens into international student experiences of the COVID-19 pandemic. *Global Netw.* 23, 14–30. doi: 10.1111/glob.12332

Hoss, T., Ancina, A., and Kaspar, K. (2022). German university students' perspective on remote learning during the COVID-19 pandemic: a quantitative survey study with implications for future educational interventions. *Front. Psychol.* 13:734160. doi: 10.3389/fpsyg.2022.734160

Huang, R., Tlili, A., Wang, H., Shi, Y., Bonk, C. J., Yang, J., et al. (2021). Emergence of the online-merge-offline (OMO) learning wave in the post-COVID-19 era: a pilot study. *Sustainability* 13:3512. doi: 10.3390/su13063512

Kamble, A., Gauba, R., Desai, S., and Golhar, D. (2021). Learners' perception of the transition to instructor-led online learning environments: facilitators and barriers during the COVID-19 pandemic. *Int. Rev. Res. Open Dist. Learn.* 22, 199–215. doi: 10.19173/irrodl.v22i1.4971

Kaur, S., Bherwani, H., Gulia, S., Vijay, R., and Kumar, R. (2021). Understanding COVID-19 transmission, health impacts and mitigation: timely social distancing is the key. *Environ. Dev. Sustain.* 23, 6681–6697. doi: 10.1007/s10668-020-00884-x

Kiger, M. E., and Varpio, L. (2020). Thematic analysis of qualitative data: AMEE guide no. 131. *Med. Teach.* 42, 846–854. doi: 10.1080/0142159X.2020.1755030

Lederman, D. (2020). How teaching changed in the (forced) shift to remote learning. Inside Higher Ed. Available at: https://www.insidehighered.com/digital-learning/article/2020/04/22/how-professors-changed-their-teaching-springs-shift-remote (Accessed August 25, 2022).

Lei, S. I., and So, A. S. I. (2021). Online teaching and learning experiences during the COVID-19 pandemic–a comparison of teacher and student perceptions. *J. Hosp. Tour. Educ.* 33, 148–162. doi: 10.1080/10963758.2021.1907196

Lemay, D. J., Bazelais, P., and Doleck, T. (2021). Transition to online learning during the COVID-19 pandemic. *Comput. Hum. Behav. Rep.* 4:100130. doi: 10.1016/j.chbr.2021.100130

Li, J., Huang, X., Lei, X., Wen, J., and Lu, M. (2022). ICT literacy, resilience and online learning self-efficacy between Chinese rural and urban primary school students. *Front. Psychol.* 13:1051803. doi: 10.3389/fpsyg.2022.1051803

Marinoni, G., Van't Land, H., and Jensen, T. (2020). The Impact of COVID-19 on Higher Education Around the World IAU Global Survey Report. Available at: https://www.uniss.it/sites/default/files/news/iau_covid19_and_he_survey_report_final_may_2020.pdf (Accessed August 25, 2022).

Naddeo, A., Califano, R., and Fiorillo, I. (2021). Identifying factors that influenced wellbeing and learning effectiveness during the sudden transition into eLearning due to the COVID-19 lockdown. *Work* 68, 45–67. doi: 10.3233/WOR-203358

Nasution, A. K. P., Batubara, M. H., and Munandar, I.Nurmalina (2022). A systematic review of the benefits and challenges of online learning during the Covid-19 pandemic. *Jurnal Mantik* 6, 1534–1541. doi: 10.35335/mantik.v6i2.2365

Noor, U., Younas, M., Saleh Aldayel, H., Menhas, R., and Qingyu, X. (2022). Learning behavior, digital platforms for learning and its impact on university student's motivations and knowledge development. *Front. Psychol.* 13:933974. doi: 10.3389/fpsyg.2022.933974

Omar, H. A., Ali, E. M., and Belbase, S. (2021). Graduate students' experience and academic achievements with online learning during covid-19 pandemic. *Sustainability* 13:13055. doi: 10.3390/su132313055

Pokhrel, S., and Chhetri, R. (2021). A literature review on impact of COVID-19 pandemic on teaching and learning. *High. Educ. Futur.* 8, 133–141. doi: 10.1177/2347631120983481

Radha, R., Mahalakshmi, K., Kumar, V. S., and Saravanakumar, A. (2020). E-learning during lockdown of Covid-19 pandemic: a global perspective. *Int. J. Control Autom.* 13, 1088–1099.

Rashid, S., and Yadav, S. S. (2020). Impact of Covid-19 pandemic on higher education and research. *Indian J. Hum. Dev.* 14, 340–343. doi: 10.1177/0973703020946700

Rudestam, K. E., and Newton, R. R. (2014). *Surviving Your Dissertation: A Comprehensive Guide to Content and Process. 4th Edn.* London: SAGE Publications, Inc.

Sahu, P. (2020). Closure of universities due to coronavirus disease 2019 (COVID-19): Impact on Education and mental health of students and academic staff. *Cureus* 12:e7541. doi: 10.7759/cureus.7541

Saikat, S., Dhillon, J. S., Wan Ahmad, W. F., and Jamaluddin, R. A. D. (2021). A systematic review of the benefits and challenges of mobile learning during the COVID-19 pandemic. *Educ. Sci.* 11:459. doi: 10.3390/educsci11090459

Sbaffi, L., and Zhao, X. (2022). Evaluating a pedagogical approach to promoting academic integrity in higher education: an online induction program. *Front. Psychol.* 13:1009305. doi: 10.3389/fpsyg.2022.1009305

Song, L., Singleton, E. S., Hill, J. R., and Koh, M. H. (2004). Improving online learning: student perceptions of useful and challenging characteristics. *Internet High. Educ.* 7, 59–70. doi: 10.1016/j.iheduc.2003.11.003

Stebbins, R. A. (2001). *Exploratory Research in the Social Sciences. Vol. 48.* London: Sage.

Szopiński, T., and Bachnik, K. (2022). Student evaluation of online learning during the COVID-19 pandemic. *Technol. Forecast. Soc. Chang.* 174:121203. doi: 10.1016/j.techfore.2021.121203

The United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2021). New UNESCO global survey reveals impact of COVID-19 on higher education. The United Nations Educational, Scientific and Cultural Organization. Available at: https://www.unesco.org/en/articles/new-unesco-global-survey-reveals-impact-covid-19-higher-education (Accessed August 25, 2022).

Weldon, A., Ma, W. W. K., Ho, I. M. K., and Li, E. (2021). Online learning during a global pandemic: perceived benefits and issues in higher education. *Knowl. Manag. E-Learn.* 13, 161–181. doi: 10.34105/j.kmel.2021.13.009

World Health Organization [WHO] (2022). Public health and social measures (PHSM) index. Available at: https://covid19.who.int/measures

Xhaferi, B., and Xhaferi, G. (2020). Online learning benefits and challenges during the COVID 19 – pandemic-Students' perspective from SEEU. *SEEU Rev.* 15, 86–103. doi: 10.2478/seeur-2020-0006

Younas, M., Noor, U., Zhou, X., Menhas, R., and Qingyu, X. (2022). COVID-19, students satisfaction about e-learning and academic achievement: mediating analysis of online influencing factors. *Front. Psychol.* 13:948061. doi: 10.3389/fpsyg.2022.948061

Zhao, X., Kung, M., and Cai, L. (2020). "How instructors initially viewed teaching online in higher education in the UK during the COVID-19 pandemic," in *Learning Technologies and Systems.* (Cham: Springer), 277–286.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 February 2023
DOI 10.3389/fpsyg.2023.1080294

OPEN ACCESS

EDITED BY
Harrison Hao Yang,
State University of New York at Oswego,
United States

REVIEWED BY
Juilong Hung,
Boise State University,
United States
Lam-for Kwok,
HKCT Institute of Higher Education,
Hong Kong SAR, China

*CORRESPONDENCE
Di Sun
✉ sundi@dlut.edu.cn
Heng Luo
✉ luoheng@mail.ccnu.edu.cn

†These authors have contributed equally to this work and share first authorship

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

RECEIVED 26 October 2022
ACCEPTED 18 January 2023
PUBLISHED 06 February 2023

CITATION
Wang X, Sun D, Cheng G and Luo H (2023) Key factors predicting problem-based learning in online environments: Evidence from multimodal learning analytics.
Front. Psychol. 14:1080294.
doi: 10.3389/fpsyg.2023.1080294

Key factors predicting problem-based learning in online environments: Evidence from multimodal learning analytics

Xiang Wang[1†], Di Sun[2*†], Gang Cheng[3,4†] and Heng Luo[5*]

[1]Faculty of Education, Beijing Normal University, Beijing, China, [2]Faculty of Humanities and Social Sciences, Dalian University of Technology, Dalian, Liaoning, China, [3]Department of Information Technology, The Open University of China, Beijing, China, [4]Engineering Research Center of Integration and Application of Digital Learning Technology, Ministry of Education, Beijing, China, [5]School of Educational Information Technology, Central China Normal University, Wuhan, Hubei, China

Problem-based learning (PBL) has been used in different domains, and there is overwhelming evidence of its value. As an emerging field with excellent prospects, learning analytics (LA)—especially multimodal learning analytics (MMLA)—has increasingly attracted the attention of researchers in PBL. However, current research on the integration of LA with PBL has not related LA results with specific PBL steps or paid enough attention to the interaction in peer learning, especially for text data generated from peer interaction. This study employed MMLA based on machine learning (ML) to quantify the process engagement of peer learning, identify log behaviors, self-regulation, and other factors, and then predict online PBL performance. Participants were 104 fourth-year students in an online course on social work and problem-solving. The MMLA model contained multimodal data from online discussions, log files, reports, and questionnaires. ML classification models were built to classify text data in online discussions. The results showed that self-regulation, messages post, message words, and peer learning engagement in representation, solution, and evaluation were predictive of online PBL performance. Hierarchical linear regression analyses indicated stronger predictive validity of the process indicators on online PBL performance than other indicators. This study addressed the scarcity of students' process data and the inefficiency of analyzing text data, as well as providing information on targeted learning strategies to scaffold students in online PBL.

KEYWORDS

problem-based learning, peer engagement, learning process, multimodal learning analytics, online learning

Introduction

Problem-based learning (PBL) is a pedagogical philosophy covering a multitude of practices and has been employed in different institutes and diverse domains in the last 50 years (Kilinska and Ryberg, 2019). PBL aims to educate students through the process of solving problems (Neville, 2009). In PBL, students are empowered with full autonomy to interact with others and use their skills and knowledge to develop a viable solution (Savery, 2006).

Given the popularity of the Internet, PBL has increasingly been carried out in online environments or blended settings. Technology-enhanced settings enable students to use various tools to perform tasks and solve problems, which leads to the generation of large amounts of data in the learning process (Unal, 2019). These data are very valuable for investigating in-depth PBL information, however, the data themselves cannot show anything without effective analysis. An emerging field addressing this challenge is learning analytics (LA), which has the capability to auto-analyze large amounts of data and presents the analysis directly to related stakeholders (Pan et al., 2020). Researchers believe that this can, in turn, empower educators to be more aware of students' progress, assess their contributions based on evidence-based criteria, and identify patterns of low engagement and students at risk of failure (Foster and Siddle, 2020).

However, current applications of LA in PBL have not related LA results with specific PBL steps: Problem-solving performance and awareness were usually predicated by the overall LA without differentiation of the theoretical stages of PBL (Chen et al., 2019; Ludwig and Rausch, 2022). Additionally, text data generated from peer interaction are not mined effectively by LA methods. Recent reviews indicate that text mining and discourse analysis have not bee widely implemented and researched for educational purposes, compared to other analytic methods (Khalil and Ebner, 2016; Nkhoma et al., 2020). To identify students' learning progress, problem-solving performance, and need for assistance, as well as to provide fair assessment and proper scaffolding, this study employed multimodal learning analytics (MMLA) based on machine learning (ML) to quantify the process engagement of peer learning based on text data, identify log behaviors, self-regulation, and other factors to predict online PBL performance.

Literature review

Problem-based learning

PBL is a student-centered pedagogy triggered by an ill-structured problem-solving scenario (Permatasari et al., 2019), in which students are enabled to participate in "learning by doing" actively and to develop transversal and lifelong learning skills (Sohmen, 2020). PBL originates from constructivist conceptions, which view learning as the active construction of knowledge that occurs through social interaction and dialogue among learners (Saqr and Alamro, 2019). The principal idea behind PBL is that the subject matter content and skills to be learned are organized around shared problems (Saqr et al., 2020). In PBL, students need to articulate the problem and then search, evaluate, construct, and share information which is then applied to a problem-solving situation in the real world (Neville, 2009). PBL thus helps students to improve their critical thinking, problem-solving ability, cognitive skills, and overall performance (Joshi et al., 2020). It is also an effective way to cultivate students to achieve 21st-century skills, such as being able to communicate and collaborate to solve complex problems, innovate in response to new demands and changing circumstances, and use technology to build new knowledge (Binkley et al., 2012).

Researchers summarize four key elements of PBL: the design of ill-structured learning problems, the role of the instructor as a facilitator, students' self-regulation in the learning process, and peer learning to interact with others (Savery, 2006; Kilinska and Ryberg, 2019). The notion of ill-structured problems as the driving force for learning is a very central aspect of PBL (Kilinska and Ryberg, 2019). As opposed to presenting direct facts and conventional concepts in traditional instruction, complicated real-world problems are used in PBL to improve and promote student learning (Joshi et al., 2020). The role of the instructor during the process becomes that of facilitator to assist students to solve the problem (Horak and Galluzzo, 2017). Thus, students have to take the responsibility to be self-directed and self-regulated in their learning, which requires them to purposefully regulate their own cognitive, motivational, and emotional behavior, as well as that of others for optimal learning (Zimmerman, 2011). Peer learning and interaction are particularly meaningful in PBL. By working together in small groups, students are expected to actively communicate, share their expertise and previous knowledge, make joint decisions, and negotiate responsibilities, as well as to evaluate and modify the strategies of learning and group work through interactive dialogue (Hennessy and Murphy, 1999; Saqr et al., 2020).

PBL has been applied in multiple domains, and different models have been proposed across the world, such as the Alborg model with eight steps in Project Management; the Maastricht model with seven steps in Science, Healthcare, and Business; the Manchester model with eight steps in Medicine and Engineering; and the Samford model with seven steps in Business, Education, and Pharmacy, among others (Zotou et al., 2020). The details of the steps in these models are not the main part of the discussion here, but, generally, various models of PBL feature peer learning and four key steps for solving ill-structured problems: problem representation, solution development, making justifications, and monitoring and evaluating (Xun and Land, 2004).

With the rapid development of information technology and digital devices, online learning has become an acceptable educational format throughout the world. As a typical pedagogy, PBL has increasingly been carried out in online environments. Online discussion is a very important way to support PBL, especially peer learning, in online settings (Saqr and Alamro, 2019); it requires students to engage in active discussions in two types of dialogical spaces: content and relational spaces. The goal in the content space is to acquire a deeper understanding of the knowledge and skills in the domain by collecting information, discussing concepts, and proposing solutions to the problem; the relational space deals with interpersonal relationships and interactions among collaborators (Slof et al., 2010; Saqr and Alamro, 2019). There is overwhelming evidence of the value of PBL; however, offering students an ill-structured problem does not directly translate to effective interactions and high performance. PBL requires scaffolding by instructors, coordination of peer learning, and active engagement of students in a stimulating environment, which calls for a mechanism to monitor the efficiency of engagement and design a data-driven intervention that supports effective PBL (Pan et al., 2020). LA is an interesting emerging field that could address these challenges.

Learning analytics

LA is concerned with the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs (Siemens, 2013). The popularity of online learning, the challenge of extracting value from educational big data, and the demand to improve performance are the three driving forces in the emergence of LA (Ferguson, 2012). The data used in LA are mainly gathered through monitoring students' activity in online learning platforms (e.g., access to resources, logins, textual input) and materials from other various tools, techniques, or environments (e.g., forums, blogs, interactive whiteboards, social sites, and libraries) (Kilinska and Ryberg, 2019). LA methods are based on educational data mining, which includes relationship mining, prediction, modeling the user's knowledge domain, personalization and adaption, and structure discovery and analysis, as well as traditional evaluation and monitoring (Siemens, 2013). The LA domain can thus accumulate as much data as possible and enable stakeholders to understand the learning process, identify students' knowledge and skills, detect students' weaknesses and misconceptions, evaluate the assessment's efficiency, and ultimately improve learning (Zotou et al., 2020).

However, most LA research focused more on clickstream data than other types of data to measure learning and instruction, with few considering psychological characteristics and the text content generated by students, thereby making LA research seem like observational reports without enough learning and instructional guidance for practices (Tsai et al., 2020). Recently, MMLA has emerged as an exciting field within the LA domain, which focuses more on the diversity of data and methods in the research process than LA before (Di Mitri et al., 2018; Emerson et al., 2020).

MMLA builds upon multimodal human interaction, educational data mining, learning sciences, and many other fields to capture the complexity of learning through data-intensive approaches (Spikol et al., 2018). First, MMLA captures multimodal data from bodily movements, face tracking, affective sensors, hardware and software log files, and user and research-generated data (e.g., discourse data). Further, it focuses on developing a better understanding of the complexity of learning through advances in high-frequency multimodal data capture, signal processing, ML techniques, and statistical methods (Ochoa and Worsley, 2016). We, therefore, believe that MMLA may offer an opportunity to capture different insights about learning in PBL to provide effective support to help students achieve good performance. In this study, since LA were collected from a semester-long course, it was infeasible to incorporate certain multimodal data (e.g., emotion, motion, and biostats) generated from sensor-based technologies, despite their proven value as MMLA (Sharma et al., 2019). Consequently, the MMLA used in the current study mainly comprised three types: psychological characteristics, activity/engagement data, and student-generated text content.

Research on integrating LA into PBL

Although many researchers have proved and introduced LA or MMLA into PBL area, the relevant research is still limited (Zotou et al., 2020). So far, two kinds of attempts have been made in the pursuit of integrating LA with PBL. One is concerned with LA as a tool in learning platforms to collect learning and instructional information, generate relevant statistical data, and provide insights into the exchange of information for stakeholders within the learning platforms. For example, Hogaboam et al. (2016) investigated the use of LA tools to support instructors in facilitating an online PBL workshop for medical students (Hogaboam et al., 2016). The researchers collected multimodal data from videos, discussions, and whiteboards through PBL, and then built an LA dashboard to visualize student performance with a scrollable news feed, a graph of the discussion, and a word cloud. Similarly, the study by Spikol et al. (2018) also focused on applying LA to multimodal data deriving from diverse sensors (computer vision, user-generated content, and data from the learning objects) during PBL and presented an LA dashboard to visualize the results and help educators determine whether groups are performing well (Spikol et al., 2018). Triantafyllou et al. (2018) developed a platform in Moodle by employing LA to monitor students' learning pathways during PBL group work (Triantafyllou et al., 2018). This provided a communication and information channel between project supervisors and students, as well as between students belonging to the same group.

The other attempt has been to use LA as a research method to investigate academic topics in PBL, where the main attention is on the prediction of student performance. Tempelaar et al. (2014) proposed a dispositional LA infrastructure that combines learning dispositions data with student engagement/activity data from the learning management system, as well as data extracted from computer-assisted formative assessments (Tempelaar et al., 2014). The results showed that computer-assisted formative assessments were the best predictor of academic performance, while basic data from the learning management system did not substantially predict learning. Saqr et al. (2020) focused on interactivity relationships in online PBL, and employed social network analysis (SNA) to investigate which factors can improve the monitoring, facilitation, and prediction of student performance (Saqr et al., 2020). They found that SNA analysis can enable the prediction of performance in groups and support students with limited participation and interactions.

However, these studies, whether regarding LA as a tool or a method, did not relate the LA results and findings with specific PBL steps or pay sufficient attention to the interaction in peer learning. More particularly, they have not targeted the analysis of text data generated from peer interaction. Although several studies have used SNA to investigate the interactivity relationship in PBL and shown that SNA can help to map the patterns of interactions and quantify the structural properties of learning groups (Dado and Bodemer, 2017; Saqr and Alamro, 2019), all of these efforts are limited in the relational space and only leverage the content space of interaction to a limited degree.

The lack of in-depth analysis of interaction content imposes many challenges in detecting the state of students' knowledge, skills, and affection, and providing them with timely and proper facilitation. Students may suffer from poor cognition, lack of skills, low motivation, or self-suspicion. If these challenges are not effectively addressed, students may be disengaged from learning, inactive in collaboration, have low performance, and even withdraw from learning. However, with the current instrumentations, much time and effort are required for human beings to code text from interactions, interviews, or surveys, which means such methods cannot deliver automated analysis of and effective insights about the learning process to educators (Saqr and Alamro, 2019). Fortunately, after decades of development, the use of text analysis—or natural language processing (NLP)—has been attempted in education, and researchers have increasingly targeted text data. Text can easily be gathered from face-to-face and online activities, which constitutes one of the most promising modalities for MMLA and will likely accelerate discourse-based research, as well as opening up new possibilities for large-scale analysis of open-ended text corpora in education (Blikstein and Worsley, 2016).

Research methodology

Research questions

To address the gap in PBL research and take advantage of multimodal data, especially text data, this study proposes NLP-and ML-enhanced strategies for PBL in online settings. Based on the key elements of PBL (especially peer learning), the goal of this exploratory MMLA research is to better process, interpret, and present various student-generated data to support the online PBL process. Specifically, this study seeks to try effective ways to present content-and process-based indicators to facilitate PBL activities in online settings.

Demographic variables such as gender and prior knowledge were also included in our analysis as potential influencing factors, and we aimed to control for their cofounding impact on students' PBL performance. In particular, the following research questions guided our investigation:

RQ1. How do gender differences and prior knowledge influence students' PBL performance in online settings?

RQ2. How does students' self-regulation influence their PBL performance in online settings, holding gender differences and prior knowledge constant?

RQ3. How does students' peer learning engagement (i.e., log behaviors, and process engagement in problem-solving steps) predict their PBL performance in online settings, controlling for gender differences, prior knowledge, and self-regulation?

Participants and pedagogical design

The participants of this study were 104 fourth-year students in an online course on social work and problem-solving in 2021 at a Chinese university. This course was conducted on Moodle and aimed to introduce social work theories and train students' problem-solving skills based on three social work cases with ill-structured problems. PBL of each case in this course was designed based on the four-step problem-solving model by Xun and Land (2004), which includes: (a) problem representation, (b) generating and selecting solutions, (c) making justifications, and (d) monitoring and evaluating goals and solutions (Xun and Land, 2004). The question prompts of each step were listed in the online discussion module of Moodle to facilitate students' PBL process. The primary goal was to assist students in going through the four steps of the problem-solving process based on the repeated practice of the three cases with ill-structured problems. Finally, we expected them to effectively understand the relevant social work theories and confidently resolve real social work problems without the instructor's facilitation at the end of the semester. The participants aged from 22 to 24, with a mean age of 23.04.

The main part of the 16-week course process was peer learning based on online discussion. Participants were randomly assigned into 13 groups consisting of eight members. In the online discussion, based on each case, the instructor provided students with specific-domain question prompts for the four steps. Based on these prompts, students could initiate their group learning, organize their thoughts, share their ideas, and communicate with peers. Further, they could improve self-regulation to construct meaningful social work plans based on personal knowledge and problem-solving skills.

The research procedure and instruments of the present study were reviewed and approved by the Institutional Review Board of Central China Normal University (CCNU-IRB-201909021, approved on 2019/09/16). Students were made aware that their participation in the research study was completely voluntary and they had the right to withdraw from the study at any time without penalty. All their personal identifiable information would remain confidential and would not appear in any publications or presentations. The participants provided their written informed consent to participate in this study.

Research design and data

This study employed the MMLA method to investigate students' learning process in online PBL. The whole MMLA deployment process includes data collection and preprocessing, linguistic feature extraction, ML classification model construction, model performance evaluation, and building the hierarchical linear regression models. The details are shown in Figure 1 and described in the section on Data Analysis.

There are two main parts: the first is "ML Classification Model on Peer Interaction," which focuses on online discussion to identify peer learning engagement in each step through the problem-solving process. The discussion messages posted by students in the 2020 online PBL course were coded by human coders (the course instructor and a teaching assistant) and were labeled with their relevance with the course topic. We used human coding results to train and evaluate ML algorithms to automatically label the messages posted in the 2021 online PBL course. The second is "Hierarchical Linear Regression on online PBL Performance," which involves building hierarchical linear regression models to identify key factors predicting students' PBL

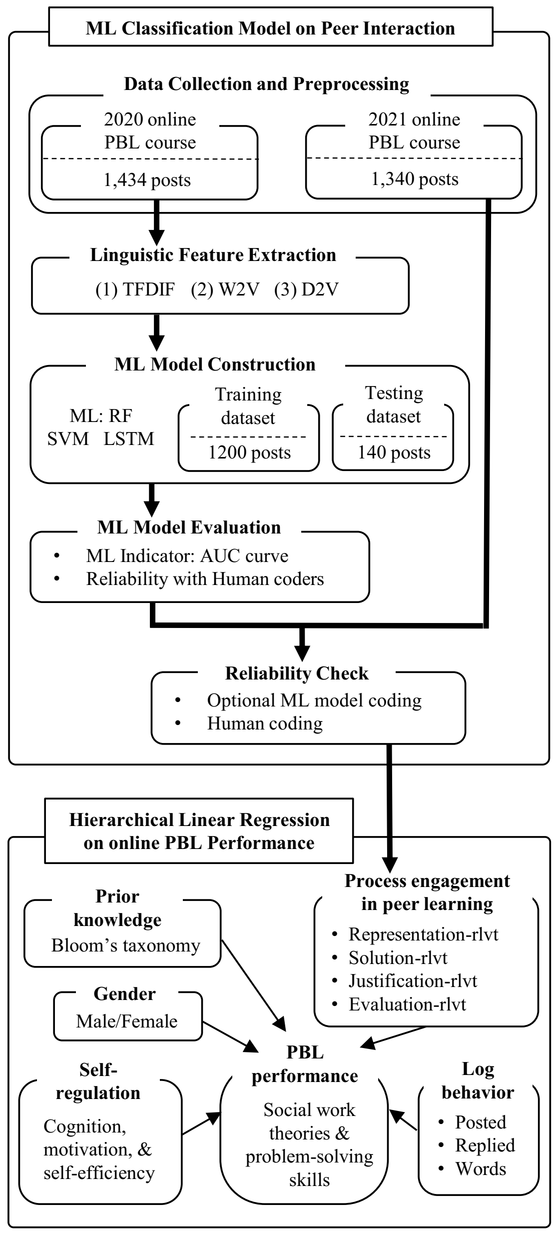


FIGURE 1
Research process.

performance in online learning and evaluating their importance by measuring their predicting capacity.

As mentioned above, the key principle affecting PBL is peer learning. Student discourse in online discussions forms the key data to identify the peer learning engagement of each step through PBL. A total of 1,434 online discussion posts from the 2020 semester and 1,340 messages from the 2021 semester were collected. The messages from 2020 were used to train and test the ML classification models based on NLP and ML algorithms. The optional model then categorized the 1,340 posts from 2021 into course-relevant and course-irrelevant data; this division formed the data used in the final hierarchical linear regression models on PBL performance. Moreover, students' log behaviors in Moodle, including the number of messages a student posted, and replied to, and the total words of messages were also collected as another kind of data indicating peer learning.

We also collected students' gender and self-regulation information through questionnaires, as well as prior knowledge from students' reports. The final PBL performance in this study was a combination of social work theories and problem-solving skills. Age is a demographic factor always related to learning performance; however, in this study, except for 7 students 24 years old and 3 students 22 years old, 94 of the 104 students were 23 years old, which means over 90% of students were of the same age. Therefore, we did not include age as a key influential factor in this study.

Instrument

Prior knowledge

At the beginning of the course, without any instruction on social work knowledge and problem-solving skills, students were required to write a design report to try to resolve the problems in the first case based on their prior knowledge. Then, two experts graded students" reports from 1 to 6 based on Bloom's taxonomy of cognitive levels, which includes recognition, understanding, analysis, application, synthesis, and evaluation. With acceptable inter-rater reliability (Spearman's Rho>0.7), the average ratings of the two experts were used as the final scores of students' prior knowledge.

Self-regulation

Like prior knowledge, student's self-regulation tendency and skills are considered as important personal trait that influences online learning performance and has a reciprocal relationship with online learning engagement, motivation, and interaction (Cho and Kim, 2013; Zheng et al., 2018; Miao and Ma, 2022). We developed a questionnaire to measure self-regulation that includes three constructs: cognition, motivation, and self-efficiency; Cronbach's Alpha is 0.856. Cognition consists of eight subscales adopted from the Community of Inquiry (CoI) framework (Arbaufvgh et al., 2008). Motivation is adapted from the scale developed by Lin et al. (2020). Self-efficiency included seven items adapted from the instrument developed by Artino and McCoach (2008).

Peer learning engagement

As mentioned above, we collected two kinds of data relevant to peer learning engagement. One was the process engagement of four problem-solving steps based on the classified discourse data through the optional ML classification model. The other was the log behaviors including the number of messages each student posted, replied to, and the total words of his/her messages.

PBL performance

In this study, PBL performance was a combination of social work theories and problem-solving skills. The full score of students' PBL performance was 100 points in two parts: multiple-choice questions on social work theories (20 points), and a final report on resolving an ill-structured social work problem by individual students themselves (80 points). The multiple-choice questions were adapted from the test banks of the National Graduate Entrance Exam, which aimed to examine students' recall and comprehension of social work knowledge taught in the course. The final report assignment was designed and graded by the course instructor, which aimed to test students' ability to apply the learned knowledge into solving authentic social work problems. The quality of final report was evaluated by four criteria: completeness of analysis, diversity of perspectives, justification of the final solution, and overall logic of reasoning.

Data analysis and results

Figure 1 illustrates the data analysis based on the two main parts of this study, with details described below.

Discourse data collection and preprocessing

To train the ML classification models, we collected 1,434 posts in online discussion data from students enrolled in this course in the 2020 academic year. These discourse data were dichotomously coded 1 as course-relevant, or 0 as course-irrelevant by the course instructor and teaching assistant. The Kappa value was near 0.93, which means a high consistency between coders. Among the coding results, 75% of the data were course-relevant.

Linguistic feature extraction

Three linguistic feature extraction (LFE) methods were employed in this study, including term frequency and inverse documentation frequency (TFIDF), Word2Vec (W2V), and Doc2Wec (D2V). TFIDF identifies the contribution of a word by calculating the frequency at which the word appears in the text and the whole corpus (Guo and Tao, 2016). W2V transforms each word into a multidimensional vector space; the distance between the vectors represents the similarity between the words (Lilleberg et al., 2015). D2V is an extension of W2V, extending the learning of the embeddings from words to word sequences, and agnostic to the granularity of the word sequence—it can be a word n-gram, sentence, paragraph, or document (Le and Mikolov, 2014).

ML classification model construction

Three ML algorithms were employed: random forest (RF), support vector machine (SVM), and long short-term memory (LSTM). The 1,434 online discussion posts from 2020 were randomly split into two subsets: 1,300 posts for training and 134 posts for testing. Based on the three LFE methods and three ML algorithms, nine different ML classification models were constructed.

Model performance evaluation

Generally, several common indicators are used to evaluate ML model performance, including accuracy, precision, recall, specificity, F1, and AUC (Powers, 2020). AUC represents the probability that the ML model can correctly classify randomly chosen course-relevant posts. AUC was adopted in this study because it can provide impartial evaluation even when the classification of the data is imbalanced. Based on the human coding results, 75% of messages were course-relevant and 25% were course-irrelevant. Therefore, AUC was the indicator chosen for ML model performance in this study. We evaluated the performance of the nine ML classification models through AUC with a 10-fold cross-validation method to prevent overfitting. The AUC values are listed in Table 1, and a larger AUC means the model is more powerful in classification.

We also employed the Kappa value to identify reliability and compare the consistency between the coding results from the ML classification models and those made by human coders. The details are also shown in Table 1; a higher Kappa indicates better consistency between the model results and human coding. To balance the efficacy of the ML classification model and consistency, we made a trade-off between AUC and Kappa to choose D2V_SVM as the final ML classification model to analyze the 1,340 online discussion posts from the 2021 academic year.

TABLE 1 ML model performance evaluation.

ML algorithm	Linguistic feature extraction	AUC	Kappa
RF	TFDIF	0.867	0.801
	W2v	0.861	0.789
	D2V	0.863	0.791
SVM	TFDIF	0.846	0.769
	W2v	0.876	0.810
	D2V	**0.880**	**0.816**
LSTM	TFDIF	0.852	0.734
	W2v	0.903	0.749
	D2V	0.891	0.760

Bold values indicate the optimal ML classification model based on a balanced performance indicated by AUC and Kappa values.

TABLE 2 Descriptive statistics of variables (*N*=104).

Variable	Mean	*SD*	Min	Max
PBL performance	75.82	11.409	42.41	95.19
Gender	0.37	0.485	0	1
Prior knowledge	3.406	0.466	3.0	5.0
Self-Regulation	3.754	0.457	2.78	5.00
Posted	9.30	4.703	1	25
Replied	4.28	1.717	3	9
Words	6,025	4,921	515	30,691
Representation-rlvt	3.62	0.981	2	6
Solution-rlvt	3.04	0.862	1	5
Justification-rlvt	2.65	0.926	1	5
Evaluation-rlvt	1.59	0.76	1	4

Reliability checking of data from 2021

The D2V_SVM model categorized the 1,340 posts from 2021 into course-relevant and course-irrelevant. We also randomly chose 100 posts to be coded by human coders. The Kappa value between the ML model coding and the human coding was 0.82, which indicates acceptable consistency. To gain an in-depth investigation of peer learning engagement in each problem-solving step, we related each courser-relevant message to a certain problem-solving step based on its log record. There were a total of 1,132 relevant messages distributed among the four PBL steps: Representation-relevant (Representation-rlvt, $n=376$), Solution-relevant (Solution-rlvt, $n=316$), Justification-relevant (Justification-rlvt, $n=275$), and Evaluation-relevant (Evaluation-rlvt, $n=165$). A larger number of relevant messages generated by students indicates greater relevance of peer interaction and discourse to the PBL task, and thus shows a higher level of peer learning engagement during the online PBL course.

Hierarchical linear regression analyses on problem-solving performance

Based on the instructional settings and the preparation of the ML classification model, we finally selected the possible factors that predict students' problem-solving performance in online PBL, including gender, prior knowledge, self-regulation, three log behavior variables, and peer learning engagement during the four steps of PBL as measured by the relevance of messages posted in each step. The descriptive statistics are shown in Table 2.

Five hierarchical linear regression models were built to investigate the unique and combined influences of these variables on online PBL performance (See Table 3). The scores for the variance inflation factors (VIFs) in each regression model were lower than 2.5, which indicates there were no collinearity issues among the independent variables in this study. The results of Model 1 showed that individual differences in gender and prior knowledge accounted for 0.057 of variance in online PBL performance, and prior knowledge was the only significant predictor ($\beta=0.241$, $p=0.016$). In Model 2, after adding the self-regulation factor, 20% more variance in online PBL performance was explained ($\Delta R^2=0.204$, $p<0.001$), holding the Model 1 variables constant. This result suggested that students' self-regulation played a significant role in online PBL ($\beta=0.452$, $p<0.001$). The R^2 and adjusted R^2 of Model 2 were 0.261 and 0.239.

In Model 3A and Model 3B, we examined the respective influences of peer learning of three log behaviors and peer learning engagement in four problem-solving steps beyond Model 2. We observed that adding three log behaviors in Model 3A accounted for an additional 0.291 of variance explained in online PBL performance ($\Delta R^2=0.291$, $p<0.001$), holding the Model 2 variables constant. The result demonstrated that students who frequently posted messages had better academic performance ($\beta=0.202$ $p=0.021$), as well as students with more message words ($\beta=0.446$, $p<0.001$). The R^2 and adjusted R^2 of Model 3A were 0.552 and 0.524. Adding peer learning engagement in four problem-solving steps to Model 3B accounted for an additional 0.323 of variance

 10.3389/fpsyg.2023.1080294

TABLE 3 Hierarchical linear regression analysis to predict PBL performance.

	Model 1					Model 2				
	B	*SE*	*β*	*p*	*VIF*	*B*	*SE*	*β*	*p*	*VIF*
Intercept	55.622	8.380		0		13.088	11.044		0.239	
Gender	0.331	2.303	0.014	0.886	1.016	−0.142	2.052	−0.006	0.945	1.018
Prior knowledge	5.895	2.395	0.241	0.016	1.016	5.997	2.131	0.245	0.006	1.016
Self-regulation	–	–	–	–	–	11.285	2.161	0.452	0.000	1.002
R^2 (Adjusted R^2)	0.057 (0.039)					0.261 (0.239)				
ΔR^2	0.057					0.204				
F	5.044*					11.650**				
	Model 3A					Model 3B				
	B	*SE*	*β*	*p*	*VIF*	*B*	*SE*	*β*	*p*	*VIF*
Intercept	29.368	9.230		0.002		9.895	8.820		0.265	
Gender	0.663	1.631	0.028	0.685	1.030	−0.733	1.593	−0.031	0.646	1.046
Prior knowledge	2.263	1.805	0.093	0.213	1.167	2.014	1.728	0.082	0.247	1.139
Self-regulation	7.371	1.808	0.295	0	1.122	8.028	1.739	0.321	0	1.106
Posted	0.490	0.209	0.202	0.021	1.597	–	–	–	–	–
Replied	0.010	0.458	0.002	0.982	1.018	–	–	–	–	–
Words	0.001	0.000	0.446	0	1.680	–	–	–	–	–
Representation-rlvt	–	–	–	–	–	3.321	0.973	0.286	0.001	1.598
Solution-rlvt	–	–	–	–	–	3.918	1.159	0.296	0.001	1.753
Justification-rlvt	–	–	–	–	–	0.283	1.056	0.023	0.79	1.677
Evaluation-rlvt	–	–	–	–	–	2.841	1.199	0.189	0.02	1.455
R^2 (Adjusted R^2)	0.552 (0.524)					0.584 (0.553)				
ΔR^2	0.291					0.323				
F	19.722**					19.043**				

	Model 4				
	B	*SE*	*β*	*p*	*VIF*
Intercept	17.848	7.969		0.028	
Gender	0.158	1.352	0.007	0.907	1.06
Prior knowledge	1.252	1.531	0.051	0.416	1.256
Self-regulation	5.820	1.517	0.233	0.000	1.182
Posted	0.673	0.186	0.277	0.000	1.88
Replied	−0.011	0.396	−0.002	0.978	1.139
Words	0.000	0.000	0.185	0.030	2.245
Representation-rlvt	2.225	0.868	0.191	0.012	1.791
Solution-rlvt	2.962	0.995	0.224	0.004	1.816
Justification-rlvt	0.635	0.908	0.051	0.487	1.744
Evaluation-rlvt	2.683	1.089	0.179	0.016	1.688
R^2 (Adjusted R^2)	0.713 (0.682)				
ΔR^2	0.161(M4 vs. M3A)/0.130(M4 vs. M3B)				
F	22.899**				

*indicates $p < 0.05$, ** indicates $p < 0.001$.

explained in online PBL performance beyond Model 2 ($\Delta R^2 = 0.323$, $p < 0.001$). This result indicated that students who engaged more in the problem representation step were more likely to have better online PBL performance ($\beta = 0.286$, $p = 0.001$), as did students who engaged more in the problem solution step (0.296 0.001) and the problem evaluation step ($\beta = 0.189$, $p = 0.02$). The R^2 and adjusted R^2 of Model 3B were 0.584 and 0.553, respectively.

Model 4 was the integrated model, investigating the combined influences of all of the study variables on online PBL performance. The R^2 and adjusted R^2 of Model 4 were 0.713 and 0.682, respectively. The results showed that students' prior knowledge no longer significantly predicted performance in online PBL ($\beta = 0.051$, $p = 0.416$). Students' self-regulation was still significant in predicting performance in online PBL ($\beta = 0.233$, $p < 0.001$). However, peer learning engagement (log behaviors and process engagement) is highlighted as the main explanation of online PBL performance. Model 4 explained significantly more of the variance than Model 3A ($\Delta R^2 = 0.161$, $p < 0.001$) and Model 3B ($\Delta R^2 = 0.13$, $p < 0.001$). However, the model comparison results also suggested that process engagement in peer learning (classified by ML models as course-relevant to each problem-solving step) had stronger predictive validity than peer learning of log behavior in explaining online PBL performance, holding all other variables constant.

Discussion

Influence of gender and prior knowledge on online PBL performance

To answer the first research question, Model 1 indicated that students' prior knowledge was a significant indicator of final PBL performance in online settings. Previous studies have illustrated that prior knowledge could reduce cognitive load and lead to better learning engagement, so it is one of the powerful influential factors in determining final learning achievement (Dong et al., 2020; van Riesen et al., 2022). The result of the current study is in line with previous findings and especially highlights that prior knowledge is influential on students' performance in online PBL settings. Focusing on this study, among the cognitive levels in Bloom's taxonomy, most students' prior knowledge was at level 3, analysis. This means that students were short of enough knowledge and skills in application, synthesis, and evaluation to resolve ill-structured social work problems. However, such higher-level cognition abilities are essential in PBL. Therefore, the course instructor designed specific question prompts to scaffold students in the problem-solving process. We would like to insist that understanding the state of students' prior knowledge is very important in PBL to provide students with appropriate and adequate facilitation.

Influence of student self-regulation on online PBL performance

For the second research question, the result of the current study was consistent with general experience and our expectations: in Model 2, students' self-regulation was validated as a positive and powerful predictor of online PBL performance. Due to the COVID-19 pandemic, online learning has become inescapable support to traditional face-to-face instruction, which is more popular in a variety of educational institutes. At present, online learning is regarded as a routine to effectively carry out learning and instruction in case of unexpected events. Self-regulation or self-direction is always a research focus in the area of online learning. Because of the separation of time and space in online learning, improving and sustaining students' self-regulation ability is a complex topic that has attracted the attention of many researchers. As a particular format of online learning, online PBL is triggered by ill-structured problems and requires students to fix the problem with higher-order thinking and skills; this means online PBL sets higher requirements for students' self-regulation in the problem-solving process. Therefore, helping students to regulate their attention on cognition, motivation, and affection through the problem-solving process can enhance their learning performance (Zimmerman, 2011). In this study, the instructor provided four steps to clarify the process of solving ill-structured problems. The study findings suggested that such process-oriented strategies scaffolded students to regulate their learning, communicate with their peers, and direct their attention to solving problems.

Influence of peer learning engagement on online PBL performance

To answer the third research question, Models 3A, 3B, and 4 were developed. In Model 3A, after controlling for student gender, prior knowledge, and self-regulation, three log behaviors of peer learning showed a significant impact on problem-solving performance in online PBL. Messages posted and the total words of messages were, in particular, significant influential factors on problem-solving performance. The number of messages a student replied to was not a significant predictor of problem-solving performance. After checking the relevant data, we found that students posted almost twice as many messages as they replied to. We speculated that students were more used to sharing their thoughts with their peers than answering others' questions. This should remind instructors to pay attention to encouraging students to help their peers in group learning and improve relevant abilities of their own.

Similarly, in Model 3B, after controlling for the same variables as in Model 3A, peer learning engagement in four steps of problem-solving significantly predicted online PBL performance. The results were in line with the literature, which has found that those students who actively communicated with peers, shared their knowledge, discussed problems, negotiated solutions, and evaluated strategies would achieve better performance (Hennessy and Murphy, 1999). However, we found that in the four steps of the problem-solving process, only engagement in Representation, Solution, and Evaluation were significant predictors of problem-solving performance. Justification did not exert a powerful influence. We would like to clarify that these results do not mean that Justification is not an important factor for problem-solving performance. On the contrary, it might imply that students need to improve their knowledge and skills to justify their solution plans for ill-structured problems.

In Model 4, the integrated Model, we found that students' prior knowledge was not a significant predictor of PBL performance; students' self-regulation, log behaviors of posted and messages words, and process engagement (Representation, Solution, and Evaluation) were still the significant predictors, especially the powerful factors of process engagement. This indicated that, with adequate coding consistency with human coders, employing the ML classification model to identify the dynamic engagement of the PBL process can mine in-depth information throughout the entire semester than using psychometric measurements or human coding. In sum, the pedagogical integration of four-step online PBL with MMLA models can provide a comprehensive understanding of online PBL experience regarding gender, prior knowledge, self-regulation, and peer learning. It also supported the notion that ML-based LA methods can be used as an alert or diagnostic module to provide in-depth information and facilitate students in moving through different online PBL phases.

Conclusion and limitations

This study constructed multimodal learning analytical models to investigate how peer learning engagement, especially dynamic process engagement in Representation, Solution, Justification, and Evaluation, along with self-regulation affects problem-solving performance under online PBL settings, holding gender and prior knowledge constant. The results revealed that self-regulation, messages posted, message words, process engagement in Representation, and Evaluation were predictive of online PBL performance. ML classification models were also built and evaluated to classify discourse data in online discussions. The optional model effectively and objectively identified peer learning engagement in different problem-solving steps in online PBL contexts and indicated stronger predictive validity of process indicators on online PBL performance than other indicators.

The study findings may contribute both theoretical and practical improvements to online PBL. Concerning theory building, this study adds to the literature that the process engagement of peer learning in online problem-solving can be incorporated into an integrated MMLA model to predict online PBL performance, which is a response to the scarcity of studies tracing students' progress in PBL (Zotou et al., 2020). Regarding practical implications, this study illustrates that ML classification models based on NLP can be trained effectively to identify process engagement from discourse data instead of conducting intensive human coding; the MMLA model can thus be applied to monitor and detect students with poor peer learning engagement through the PBL process. The MMLA model's predictions can help stakeholders to design strategies, make decisions, and conduct evaluations to foster students' problem-solving abilities in online settings.

This study shows that gender does not have an impact social work problem-solving performance in online learning, but prior knowledge is the basis and start for further learning. We, therefore, suggest that before or at the beginning of the online PBL, students should have a pre-test to illustrate the state of their knowledge and skills relevant to the learning topic, which can help the instructors to set up effective scaffolds. Throughout the process of online PBL, instructors can take advantage of ML, MMLA, or other technology to monitor and evaluate students' self-regulation and peer learning engagement, and then provide personalized facilitation to improve student performance.

There are, however, some limitations to this study. First, the participants in this study were drawn from only one of the social work courses at the undergraduate level, which might not represent students in other disciplines. Second, we only trained and tested limited NLP methods and ML classification algorithms, and the sample size was not very large. In the future, we would like to develop more ML classification models as well as further improve reliability. Third, although the process engagement based on discourse data was coded into course-relevant or course-irrelevant, and linked to the four problem-solving steps, the discourse may be coded into multiple classes in the future to investigate more in-depth information from the valuable data. Finally, as PBL and online learning both rely on self-regulation—especially intrinsic motivation in learning—future research could include more multimodal data from various learning settings, such as social media, intelligent agents, online platforms, onsite classrooms, or face-to-face environments, to integrate self-regulation and peer learning to foster a comprehensive MMLA model of online PBL.

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 research procedure and instruments of the present study were reviewed and approved by the Institutional Review Board of Central China Normal University (CCNU-IRB-201909021, approved on 2019/09/16). The participants provided their written informed consent to participate in this study.

Author contributions

DS and HL: conceptualization and writing–review and editing. XW and DS: methodology, investigation, and writing–original draft preparation. DS and GC: formal analysis. HL and GC: resources. DS and GC: project administration. DS: funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This work was funded by Research on Intelligent Big Data and Community Peacebuilding, Grant of "Social Science Foundation of Liaoning Province in China" (L21BSH002).

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

References

Arbaufvgh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., et al. (2008). Developing a community of inquiry instrument: testing a measure of the Community of Inquiry framework using a multi-institutional sample. *Internet High. Educ.* 11, 133–136. doi: 10.1016/j.iheduc.2008.06.003

Artino, A. R., and Mccoach, D. B. (2008). Development and initial validation of the online learning value and self-efficacy scale. *J. Educ. Comput. Res.* 38, 279–303. doi: 10.2190/EC.38.3.c

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., et al. (2012). *Defining Twenty-first Century Skills. Assessment and Teaching of 21st Century Skills.* Dordrecht: Springer.

Blikstein, P., and Worsley, M. (2016). Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. *J. Learn. Anal.* 3, 220–238. doi: 10.18608/jla.2016.32.11

Chen, L., Yoshimatsu, N., Goda, Y., Okubo, F., Taniguchi, Y., Oi, M., et al. (2019). Direction of collaborative problem solving-based STEM learning by learning analytics approach. *Res. Pract. Technol. Enhanc. Learn.* 14, 14–24. doi: 10.1186/s41039-019-0119-y

Cho, M.-H., and Kim, B. J. (2013). Students' self-regulation for interaction with others in online learning environments. *Internet High. Educ.* 17, 69–75. doi: 10.1016/j.iheduc.2012.11.001

Dado, M., and Bodemer, D. (2017). A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educ. Res. Rev.* 22, 159–180. doi: 10.1016/j.edurev.2017.08.005

Di Mitri, D., Schneider, J., Specht, M., and Drachsler, H. (2018). From signals to knowledge: a conceptual model for multimodal learning analytics. *J. Comput. Assist. Learn.* 34, 338–349. doi: 10.1111/jcal.12288

Dong, A., Jong, M. S.-Y., and King, R. B. (2020). How does prior knowledge influence learning engagement? The mediating roles of cognitive load and help-seeking. *Front. Psychol.* 11, 1–10. doi: 10.3389/fpsyg.2020.591203

Emerson, A., Cloude, E. B., Azevedo, R., and Lester, J. (2020). Multimodal learning analytics for game-based learning. *Br. J. Educ. Technol.* 51, 1505–1526. doi: 10.1111/bjet.12992

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *Int. J. Technol. Enhanc. Learn.* 4, 304–317. doi: 10.1504/IJTEL.2012.051816

Foster, E., and Siddle, R. (2020). The effectiveness of learning analytics for identifying at-risk students in higher education. *Assess. Eval. High. Educ.* 45, 842–854. doi: 10.1080/02602938.2019.1682118

Guo, A., and Tao, Y. (2016). Research improvement of feature words weight based on TFIDF algorithm. IEEE Information Technology, Networking, Electronic and Automation Control Conference, 2016 Chongqing, China. 415–419.

Hennessy, S., and Murphy, P. (1999). The potential for collaborative problem solving in design and technology. *Int. J. Technol. Des. Educ.* 9, 1–36. doi: 10.1023/A:1008855526312

Hogaboam, P., Chen, Y., Hmelo-Silver, C. E., Lajoie, S. P., Bodnar, S., Kazemitabar, M., et al. (2016). Data dashboards to support facilitating online problem-based learning. *Q. Rev. Dist. Learn.* 17, 95–97.

Horak, A. K., and Galluzzo, G. R. (2017). Gifted middle school students' achievement and perceptions of science classroom quality during problem-based learning. *J. Adv. Acad.* 28, 28–50. doi: 10.1177/1932202X16683424

Joshi, A., Desai, P., and Tewari, P. (2020). Learning analytics framework for measuring students' performance and teachers' involvement through problem based learning in engineering education. *Proc. Comp. Sci.* 172, 954–959. doi: 10.1016/j.procs.2020.05.138

Khalil, M., and Ebner, M. (2016). What is learning analytics about? A survey of different methods used in 2013-2015. ArXiv, abs/1606.02878.

Kilinska, D., and Ryberg, T. (2019). Connecting learning analytics and problem-based learning–potentials and challenges. *J. Prob. Based Learn. High. Educ.* 7, 1–24. doi: 10.5278/ojs.jpblhe.v7i1.2545

Le, Q., and Mikolov, T. Distributed representations of sentences and documents. The 31st international conference on machine learning (ICML 2014), (2014). Beijing, China. 1188–1196.

Lilleberg, J., Yun, Z., and Zhang, Y.. Support vector machines and Word2vec for text classification with semantic features. IEEE international conference on Cognitive Informatics and Cognitive Computing. (2015). Beijing, China.

Lin, L., Ginns, P., Wang, T., and Zhang, P. (2020). Using a pedagogical agent to deliver conversational style instruction: what benefits can you obtain? *Comput. Educ.* 143:103658. doi: 10.1016/j.compedu.2019.103658

Ludwig, S., and Rausch, A. (2022). The relationship between problem-solving behaviour and performance–Analysing tool use and information retrieval in a computer-based office simulation. *J. Comput. Assist. Learn.*, 1–27. doi: 10.1111/jcal.12770

Miao, J., and Ma, L. (2022). Students' online interaction, self-regulation, and learning engagement in higher education: the importance of social presence to online learning. *Front. Psychol.* 13:815220. doi: 10.3389/fpsyg.2022.815220

Neville, A. J. (2009). Problem-based learning and medical education forty years on. *Med. Princ. Pract.* 18, 1–9. doi: 10.1159/000163038

Nkhoma, C., Dang-Pham, D., Hoang, A.-P., Nkhoma, M., Le-Hoai, T., and Thomas, S. (2020). Learning analytics techniques and visualisation with textual data for determining causes of academic failure. *Behav. Inform. Technol.* 39, 808–823. doi: 10.1080/0144929X.2019.1617349

Ochoa, X., and Worsley, M. (2016). Augmenting learning analytics with multimodal sensory data. *J. Learn. Anal.* 3, 213–219. doi: 10.18608/jla.2016.32.10

Pan, Z., Li, C., and Liu, M. (2020). Learning Analytics Dashboard for Problem-based Learning. *The Seventh ACM Conference on Learning @ Scale (L@S '20)*. New York, NY, USA.

Permatasari, B., Gunarhadi, G., and Riyadi, R. (2019). The influence of problem based learning towards social science learning outcomes viewed from learning interest. *Int. J. Eval. Res. Educ.* 8, 39–46. doi: 10.11591/ijere.v8i1.15594

Powers, D. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv* 1504:854. doi: 10.48550/arXiv.2010.16061

Saqr, M., and Alamro, A. (2019). The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Med. Educ.* 19:160. doi: 10.1186/s12909-019-1599-6

Saqr, M., Nouri, J., Vartiainen, H., and Malmberg, J. (2020). What makes an online problem-based group successful? A learning analytics study using social network analysis. *BMC Med. Educ.* 20:80. doi: 10.1186/s12909-020-01997-7

Savery, J. (2006). Overview of problem-based learning: definitions and distinctions. *Interdisc. J. Problem Learn.* 1, 9–20. doi: 10.7771/1541-5015.1002

Sharma, K., and Papamitsiou, Z. And Giannakos, M. (2019). Building pipelines for educational data using AI and multimodal analytics: a "grey-box" approach. *Br. J. Educ. Technol.*, 50, 3004–3031, doi: 10.1111/bjet.12854

Siemens, G. (2013). Learning analytics: the emergence of a discipline. *Am. Behav. Sci.* 57, 1380–1400. doi: 10.1177/0002764213498851

Slof, B., Erkens, G., Kirschner, P., Jaspers, J., and Janssen, J. (2010). Guiding students' online complex learning-task behavior through representational scripting. *Comput. Hum. Behav.* 26, 927–939. doi: 10.1016/j.chb.2010.02.007

Sohmen, V. (2020). "Project-based learning (PBL) in a higher education project introduction of an accelerated PBL (A-PBL) model" in *Handbook of Research on Adult Learning in Higher Education*. eds. M. C. P. O. Okojie and T. C. Boulder (Pennsylvania: IGI global).

Spikol, D., Ruffaldi, E., Dabisias, G., and Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *J. Comput. Assist. Learn.* 34, 366–377. doi: 10.1111/jcal.12263

Tempelaar, D. T., Rienties, B., and Giesbers, B. (2014). "Computer assisted, formative assessment and dispositional learning analytics in learning mathematics and statistics" in *Caa 2014: Communications in Computer and Information Science*. eds. M. Kalz and E. Ras (Zeist, The Netherlands: Springer, Cham), 67–78.

Triantafyllou, E., Xylakis, E., Nilsson, N. C., and Timcenko, O. Employing learning analytics for monitoring student learning pathways during problem-based learning group-based group work: a novel approach. The 7th international research symposium on PBL: Innovation, PBL and competences in engineering education, (2018). Beijing, China. 542–551.

Tsai, Y.-S., Rates, D., Moreno-Marcos, P. M., Muñoz-Merino, P. J., Jivet, I., Scheffel, M., et al. (2020). Learning analytics in European higher education—trends and barriers. *Comput. Educ.* 155:103933. doi: 10.1016/j.compedu.2020.103933

Unal, E. (2019). Web 2.0 technologies supporting problem based learning: a systematic literature review. *J. Prob. Based Learn. High. Educ.* 7, 25–50. doi: 10.5278/ojs.jpblhe.v7i1.2845

Van Riesen, S. A. N., Gijlers, H., Anjewierden, A. A., and De Jong, T. (2022). The influence of prior knowledge on the effectiveness of guided experiment design. *Interact. Learn. Environ.* 30, 17–33. doi: 10.1080/10494820.2019.1631193

Xun, G., and Land, S. M. (2004). A conceptual framework for scaffolding III-structured problem-solving processes using question prompts and peer interactions. *Educ. Technol. Res. Dev.* 52, 5–22. doi: 10.1007/BF02504836

Zheng, C., Liang, J. C., Li, M., and Tsai, C. C. (2018). The relationship between English language learners' motivation and online self-regulation: a structural equation modelling approach. *System* 76, 144–157. doi: 10.1016/j.system.2018.05.003

Zimmerman, B. (2011). "Motivational sources and outcomes of self-regulated learning and performance" in *Handbook of Self-regulation of Learning and Performance*. eds. B. J. Zimmerman and D. H. Schunk (Abingdon-on-Thames: Taylor & Francis).

Zotou, M., Tambouris, E., and Tarabanis, K. (2020). Data-driven problem based learning: enhancing problem based learning with learning analytics. *Educ. Technol. Res. Dev.* 68, 3393–3424. doi: 10.1007/s11423-020-09828-8

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 21 February 2023
DOI 10.3389/fpsyg.2023.1118174

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University, China

REVIEWED BY
Tingting Yang,
University of Georgia, United States
Gege Li,
Hokkaido University of Education, Japan

*CORRESPONDENCE
Geping Liu
✉ liugp2@swu.edu.cn

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

RECEIVED 07 December 2022
ACCEPTED 30 January 2023
PUBLISHED 21 February 2023

CITATION
Han J, Liu G and Zheng Q (2023) Prior knowledge as a moderator between signaling and learning performance in immersive virtual reality laboratories.
Front. Psychol. 14:1118174.
doi: 10.3389/fpsyg.2023.1118174

Prior knowledge as a moderator between signaling and learning performance in immersive virtual reality laboratories

Jining Han[1], Geping Liu[1]* and Qiyu Zheng[2]

[1]Faculty of Education, Southwest University, Chongqing, China, [2]Chaoyang School, Chongqing, China

The purpose of this study is to investigate the effects of signaling and prior knowledge on the cognitive loads, motivations, and learning of college students in an immersive virtual reality (IVR) environment. This study applied a 2 (signaling vs. no signaling) by 2 (high vs. low prior knowledge levels) between-subjects factorial design. The results revealed that signaling directed the attention of students with low prior knowledge levels, effectively helped them select relevant information and reduced their cognitive loads, whereas signaling had no significant effect on the cognitive loads, intrinsic motivation, and learning performance of learners with high levels of prior knowledge. These results suggest that IVR environments for students with low prior knowledge levels should reduce cognitive load and improve learning, and signals in the form of text annotations and color changes are recommended for additional support. Students with high prior knowledge levels do not require additional signals to support learning; therefore, the IVR environment needs to be designed in such a way as to be tailored to the individual differences of students.

KEYWORDS

immersive virtual reality, signaling, prior knowledge, cognitive load, learning performance

1. Introduction

According to the European Commission's Education and Training 2020 Strategy and the report on the agenda for the modernization of higher education, one of the main challenges for higher education is to improve the quality of teaching and learning. Studies show that there is a high demand for laboratory teaching in higher education, and due to limitations on staff and equipment that were induced by the COVID-19 pandemic, students have only limited access to equipment and machines in the laboratory courses offered in higher education (Mishra et al., 2020). By using virtual reality (VR) in class, students are able to directly observe experimental phenomena and independently master the learning process in a VR environment, increasing the effectiveness and efficacy of acquiring knowledge and skills (Jensen and Konradsen, 2018; Radianti et al., 2020). VR provides contextual and diverse learning resources, but it also inevitably brings new technological problems, such as the unreasonable presentation of visual elements and maladaptation with multichannel perception, which causes learners to experience increased cognitive loads and experience a negative impact (Baceviciute et al., 2021; Parong and Mayer, 2021b); therefore, instructional aid is needed to help learners.

With the increasing fidelity and visual richness of the teaching and learning environment, especially in immersive virtual reality (IVR) environments, the visual load for learners increases significantly (e.g., Birbara and Pather, 2021; Parong and Mayer, 2021a). Many studies have found that signaling can improve learning and assist learners in their visual search for important information (e.g., Johnson et al., 2014; Arslan-Ari, 2018; Albus et al., 2021), thereby reducing learning time. The learner's prior knowledge is an important determinant of the learning effect (Johnson and Lawson, 1998). It was formerly widely assumed that learners constructed concepts based on their prior knowledge (Novak, 1990), but further research has revealed that prior knowledge affects not only learners' subsequent conceptual learning but also their perception and attention (Kim and Rehder, 2011; Buckingham et al., 2016). Therefore, changes in the way learners interpret visual representations also largely depend on their prior knowledge. However, only a few studies have investigated the effects of learners' different prior knowledge levels on the efficacy of signaling (e.g., Richter et al., 2016). At the same time, there is a lack of studies on how interactive IVR laboratories impact the cognitive processes of learners interacting with learning situations, which is an important perspective for research on VR technology in education.

2. Literature review

2.1. Virtual reality in multimedia learning

Mayer (2020) published the third edition of Multimedia Learning, and the principles of multimedia instructional design were reorganized into three sections. In the context of multimedia learning, the first section offers the principles for reducing extraneous processing, the second for managing essential processing, and the third for fostering generative processing (Mayer, 2020). With the rapid development of VR in education, many studies have discussed multimedia learning principles in VR environments and provided implications for designing VR for use in multimedia learning.

Makransky et al. (2019a) used the redundancy principle to design virtual learning resources in a study in which two groups of learners learned with either virtual reality head-mounted display (VR-HMD) or desktop presentation and received two learning simulations, one with on-screen text and one with on-screen text with narration. The study found that the redundancy principle had no effect on the learning process. In another study by Parong and Mayer (2018), an empirical study was conducted on the effectiveness of the segmenting principle. They compared the learning effect of a self-paced slideshow with that of a continuous VR animation. The learners who received the slideshow scored higher on the factual questions than the VR group learners but did not score as high on the conceptual questions. Baceviciute et al. (2021) explored the application of the modality principle in VR by comparing audio-visual and visual-only presentations, and the results showed an inverse modality effect, whereby learning under visual-only presentations was more effective than auditory and visual presentations. Petersen et al. (2020) investigated pretraining principles in VR. The experimental group received narrative pretraining followed by a VR exploration tour, while the control group went straight to VR learning. The pretraining group showed a significant increase in test scores for knowledge transfer. Makransky et al. (2019b) examined the embodiment principle and found that gender-specific design in VR learning environments can impact learning performance, retention, and transfer. Makransky et al. (2020) explored the formulating generative learning strategies in VR, and they found that compared to video, VR can better facilitate the transfer of procedural knowledge with generative learning strategies.

The transferability and effectiveness of multimedia design principles in IVR are visible in studies that show a relatively significant effect on learners in IVR learning environments, mainly reflected in cognitive processing and cognitive loads, but no clear conclusions can be drawn about the specific learning process at present (Han et al., 2021). More empirical findings are needed to further enrich the effectiveness of multimedia design principles on the design of IVR learning environments.

2.2. Signaling

Some studies have pointed out that adding visual cues to learning materials can reduce cognitive load and effectively improve learning effects (Mayer et al., 2003). The signaling principle, which promotes goal-oriented learning by highlighting the organizational structure, can be followed in the design and development of learning materials to enhance learning by highlighting important information (e.g., underlining, marking, etc.,) to attract learners' attention (de Koning et al., 2009).

The salient visual cues in the traditional multimedia learning environment are the signaling models, which can be divided into two categories: textual signaling and graphic signaling (Van Gog, 2014). The experimental results of the studies that have been conducted show that the two forms of signaling have different effects on learners. For example, a study showed that when color flashes were used as signals to draw learners' attention to various laboratory tools in the virtual environment, learners experienced a negative effect on their learning, and some color flashes had greater negative impacts than others (Jeung et al., 1997). Contrary to the findings of a previous study, another study found that the application of color coding in mind maps had a more positive effect on learning (Ferrara and Butcher, 2011). Hwang and Shin (2018) combined dynamic and static visualizations into one medium by adding transparent static images (graphical signals) to virtual animations. The results showed that combining dynamic and static visualization signals was beneficial in reducing the cognitive load of learners.

Due to the highly immersive feature of the IVR learning environment, the textual and graphic signals used are quite different from those designed and produced in traditional multimedia materials (Albus et al., 2021), resulting in a large difference in the users' experience, and it is not yet possible to conclude with certainty whether the effect of textual or graphic signals in IVR environments is consistent with the findings of previous studies. Therefore, further research on the application of signaling in IVR experiments is needed to provide more diverse perspectives on the design of IVR environments to enhance the practicability and effectiveness of IVR technology.

2.3. Prior knowledge impacts the effect of signaling

Learners use their prior knowledge to select relevant information from visual cues and to find, retrieve and add information from their prior knowledge and finally construct mental models (Braune and Foshay, 1983). Because the signaling effect and difference in the level of prior knowledge are very significant and there is an interaction effect (Johnson et al., 2014; Arslan-Ari, 2018), it is necessary to

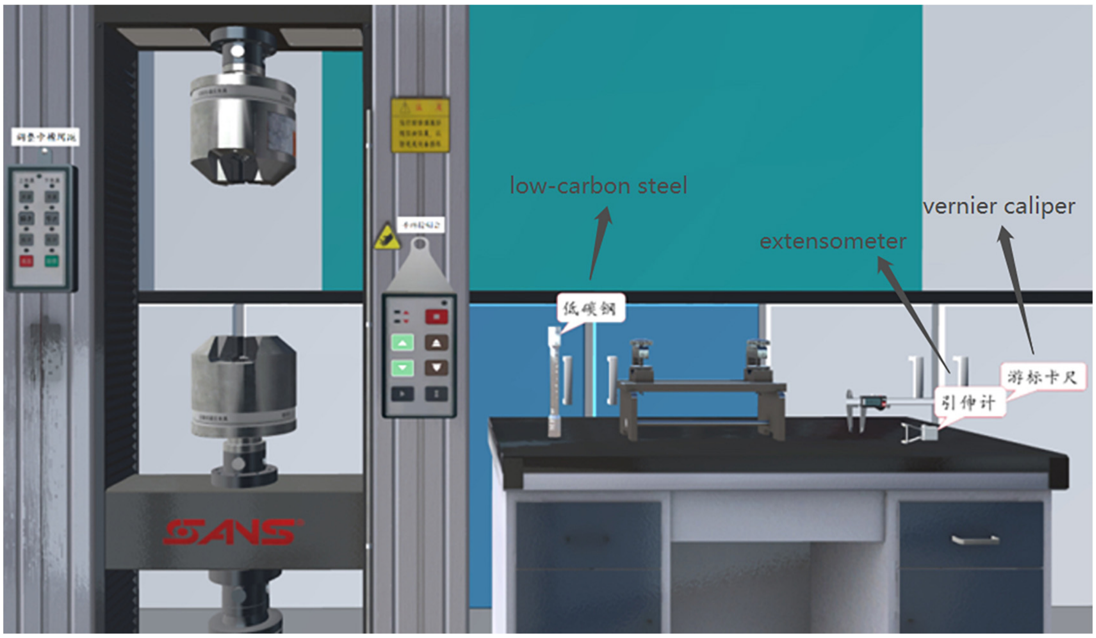


FIGURE 1
Text signals in the immersive virtual reality (IVR) experiment.

explore the synergetic effect of signaling and prior knowledge level in the IVR environment.

Cook (2006) investigated the instructional design of visual representations in the science classroom, conducting empirical research, and integrating theoretical concepts related to cognitive load. The study concluded that individual differences—especially learners' prior knowledge level—are a key factor in determining the impact of visual representations on learners' cognitive structures and learning processes because prior knowledge can determine the ease with which learners perceive and interpret visual representations in working memory. Johnson et al. (2013, 2014) found that learners with low prior knowledge levels performed better with signals than without them, but signals were not very helpful for learners with high prior knowledge. They found that there was a significant interaction effect when the signals were presented together with an Animation Pedagogical Agent (APA) in the learning material; however, the arrow signals had no significant effect on learners when the learning material did not include an APA. Khacharem (2017) reported that red circle signals in static images of football matches were effective in directing the attention of novice learners; however, this strategy was ineffective for learners with relevant sports experience.

Arslan-Ari (2018) investigated the effects of signaling and prior knowledge levels on student learning from animations with narration, and the results showed that prior knowledge and signaling had significant interaction effects on the learning effect. The effect of providing signaling in instructional animations varies with learners' prior knowledge levels. Specifically, when visual signals were provided, learners with low prior knowledge benefited more, while visual signals did not have an effect on learners with high prior knowledge. Vogt et al. (2020) investigated the effects of signals and graphic signals on learning performance and their intrinsic interaction mechanisms. The findings suggest that prior knowledge is an important moderating factor and that although learners with a low prior knowledge level can benefit from both types of signals, the gains in the learning effect they obtain are not significant, whereas learners with a moderate prior knowledge level benefited from the combined effect of both types of signals. For those with a higher prior knowledge level, signals can instead reduce their learning effect. Therefore, to deeply understand the role of the signaling principle in supporting or hindering different learners, it is necessary to conduct a more refined analysis of signal reception and cognitive construction during the learning process and to explore how learners obtain positive effects and cognitive feedback.

Although most studies have shown the moderating effect of prior knowledge level in multimedia learning, existing studies have not been concerned with the interaction between prior knowledge level and signaling design in IVR. The effects of visual selection, organization, and integration processes on the learning process in IVR environments also vary with individual learners; therefore, this study explores the interaction effect of signaling and prior knowledge level on learning effect by measuring learners' cognitive load, intrinsic motivation, and learning performance in IVR. This study contributes to the empirical research on signaling principles in IVR laboratories in theory and practice and provides valuable information for instructional designers and practitioners.

2.4. Purpose of the study

The presentation of learning materials in highly IVR environments requires explicit signal selection and organization to assist learners in information reception and selection, including from text and images, in different sensory modes in an IVR environment. For learners with different prior levels of ability, signaling may have different advantages. This study provides further insight into the effect of multimedia design principles in IVR environments and delves into the effects of signaling and prior knowledge levels on learning effects in IVR experimental conditions.

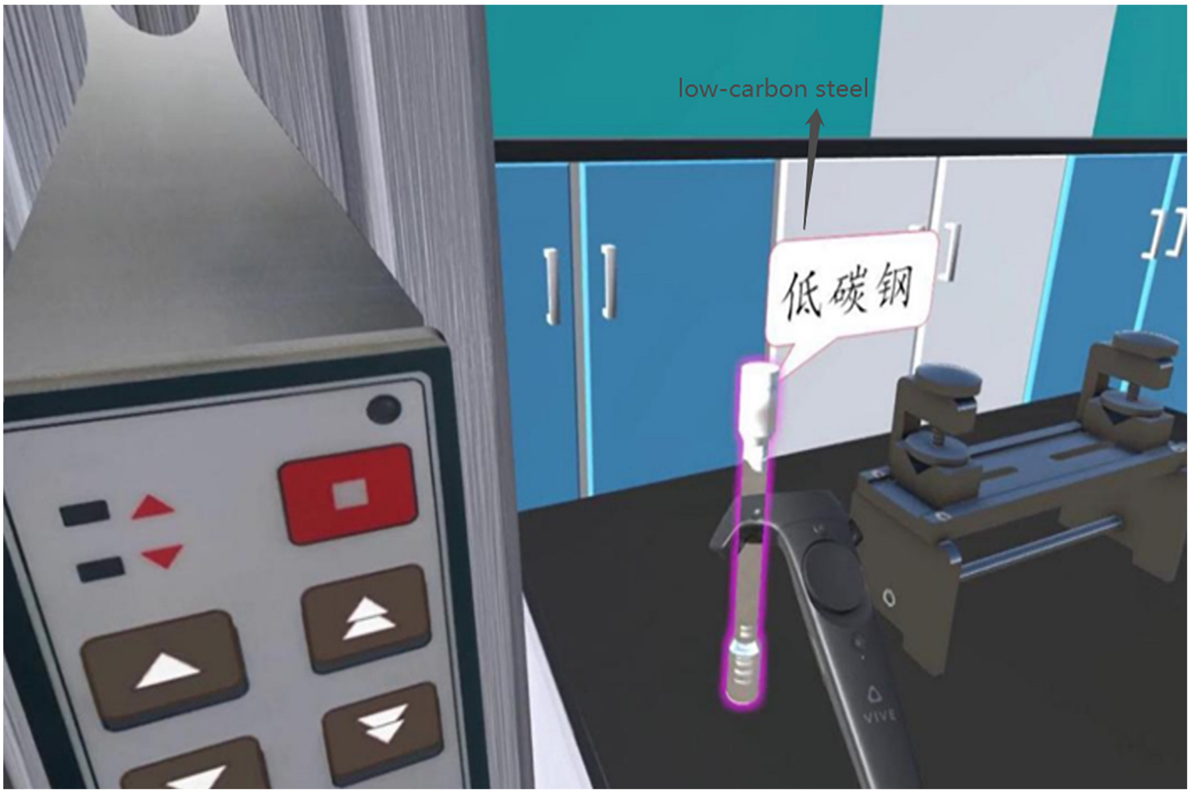


FIGURE 2
Highlight signals for the immersive virtual reality (IVR) experiment.

(1) What are the effects of signaling and prior knowledge on the cognitive load of college students studying in an IVR environment?

(2) What are the effects of signaling and prior knowledge on the motivations of college students studying in an IVR environment?

(3) What are the effects of signaling and prior knowledge on the learning performance of college students studying in an IVR environment?

3. Materials and methods

3.1. Participants

A 2 × 2 factorial ANOVA design was used in this study, with the independent variables being signaling and prior knowledge, resulting in four experimental conditions: no signaling/low prior knowledge level, no signaling/high prior knowledge level, signaling/low prior knowledge level, and signaling/high prior knowledge level. The participants were recruited *via* the social media postings of a university. Forty volunteer undergraduate students (21 female and 19 male) from various majors at a large public university in the southwestern part of China participated in the study. All the participants completed and submitted informed consent forms. The age range of the participants was 21–30 years old (M = 23.975, SD = 1.60). The study used the HTC Vive Pro including a VR headset, base stations, controllers, mounting kit, etc. Therefore, participants were required to physically come to the laboratory to participate in the study. Due to the impact of the epidemic, 40 participants were recruited. Han et al. (2021) conducted a systematic review for IVR studies from 2004 to 2021, and the study reported that the sample sizes ranged from 8 to 174, with a median participant number of 58. Although the sample size is small, the research team added qualitative data, multiple data sources could help in answering the research questions (Twining et al., 2017).

3.2. Instruction materials

In this study, the laboratory course "Low-carbon Steel Tensile Experiment" of the architectural engineering technology major was used as the learning task in the IVR experiment. Unity 3D software was used in this study to build the IVR laboratory and conduct interactive design. The VR environment was composed of multiple simulation models. The VR laboratory environment in this study was simulated as an indoor environment. To simulate a sense of the environment, the lighting components of the software were used to simulate the lighting effect of a real environment. The computer used a camera to render and display the environment. This study simulated the IVR laboratory environment of architectural engineering technology through scene elements, which included game objects, directional light, and a unity camera. The design principles were as follows: (1) the simulated IVR environment should improve learners' sense of immersion and presence in operation; (2) based on human-computer interactions, the operation steps in the IVR experiment should be consistent with those of the actual experiment; and (3) the human-computer interface should be both simple and practical to ensure smooth human-computer interactions.

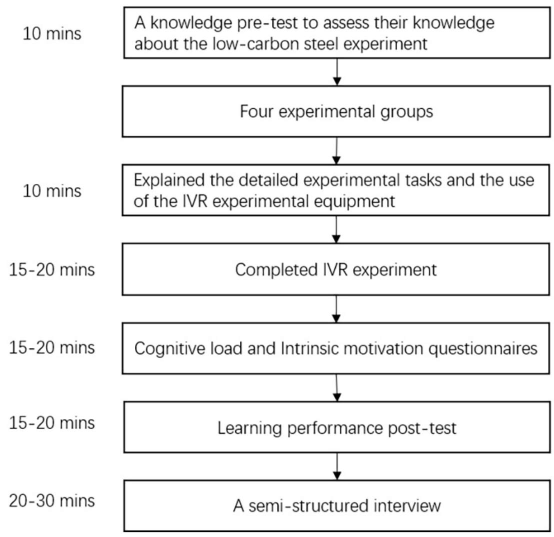


FIGURE 3
Experimental procedure.

The main contents included observing four stages of the failure process of low-carbon steel, the shape of the truncated surface and the stress-displacement curve and understanding the relationship between the force and displacement to which the low-carbon steel is subjected during the tensile strength test. The experiment took 15–20 min to complete, and the five key steps were as follows: (1) correctly install the low-carbon steel on the electronic universal testing machine; (2) install the extensometer; (3) open the operating software, set parameters and fill in the preparatory information into the experimental report; (4) observe the low-carbon steel tensile test; and (5) fill in the results of the experimental report. The IVR experiment with signals is the same in learning content as that without signals, but the IVR experiment with signals has added text annotations (see **Figure 1**) and highlights the selected object (see **Figure 2**).

3.3. Instruments

3.3.1. Prior knowledge test

The prior knowledge multiple-choice test included five questions related to the experimental contents of this study. The questions were used to test the participants' basic knowledge of architectural mechanics related to the experiment, but no questions were posed that were similar to the experimental tasks to avoid influencing the subsequent experiments and tests. The test was positively scored, with higher scores indicating a higher level of prior knowledge of the learners. The questions on the prior knowledge test were compiled by the members of the research team based on relevant knowledge and were cross-reviewed by experts in architectural engineering technology for final confirmation.

3.3.2. Cognitive load questionnaire

The NASA Task Load Index (NASA-TLX) scale was used in this study to measure the cognitive load of learners in completing the IVR experiment. The six dimensions measured by the scale were mental demand, physical demand, temporal demand, performance, effort, and frustration level. The Cronbach's alpha of the scale was 0.85, indicating that the scale is a reliable assessment for measuring learners' subjective mental workload.

3.3.3. Intrinsic motivation scale

This intrinsic motivation scale, adapted from Ryan's (1982) questionnaire and used in Lin and Li's (2018) study, consisted of four subdimensions, including interest, competence, value, and pressure. The learners' intrinsic motivation scale used a 7-point Likert scale, ranging from 1 ("Completely False") to 7 ("Completely True"). Negative items were reverse scored, so higher scores reflect more positive motivation. The Cronbach's alphas for interest, competence, value, and pressure were 0.90, 0.74, 0.78, and 0.76, respectively.

3.3.4. Learning performance

After the four experimental groups finished the IVR experiment, they completed the learning performance post-test, which was set up as a question-type knowledge retention test. The six multiple-choice questions in this test were written by experts in architectural engineering technology and mainly included questions testing observation (e.g., "How many stages is the stress−strain diagram of low-carbon steel divided into during the tensile test?," "What is the shape of the fracture of low-carbon steel?"), memory retention questions (e.g., "Please choose the correct sequence of low-carbon steel tensile experiments"), knowledge transfer questions (e.g., "What is the deformation of low-carbon steel like during its failure process before it is subjected to the maximum tensile stress? What does this deformation say about its tensile capacity?"), which were used to check whether the learners had mastered the knowledge related to the low-carbon steel tensile experiment.

3.4. Procedure

One week prior to the experiment, the participants received a knowledge pretest to assess their knowledge about the low-carbon steel experiment, and they were subsequently and accordingly grouped into high or low prior knowledge learners. A median score was used in this study to differentiate high from low levels of prior knowledge. Specifically, participants with test scores below or equal to the median score of the prior knowledge test (Mdn = 25) were grouped as having low prior knowledge (n = 20), and participants with test scores above the median score were grouped as having high prior knowledge (n = 20). The independent sample t-test revealed a significant difference in the mean scores of the pretest between the low prior knowledge group (M = 9.25, SD = 7.65) and the high prior knowledge group (M = 41.5, SD = 7.79), $t = -13.20$, $p < 0.01$. One week after taking the prior knowledge test, participants participated in the IVR experiment task to complete the learning and assessment parts of the study.

The participants were semi-randomly assigned to no signaling (n = 20) and signaling (n = 20) conditions based on the results of the prior knowledge level pretest. The four experimental groups were as follows: no signaling/low prior knowledge level (n = 10), no signaling/high prior knowledge level (n = 10), signaling/low prior knowledge level (n = 10), and signaling/high prior knowledge level (n = 10). To carry out the experiment successfully, the participants first needed to be familiar with the knowledge related to the contents of the experiment. Reading materials for the low-carbon steel tensile

 10.3389/fpsyg.2023.1118174

TABLE 1 Analysis of variance (ANOVA) results.

	Dependent variable	Type III sum of squares	df	Mean square	*F*	Sig	η^2
Signaling	Cognitive load	0.251	1	0.251	0.632	0.432	0.017
	Intrinsic motivation	0.437	1	0.437	2.351	0.134	0.061
	Post-test	140.625	1	140.625	1.246	0.272	0.033
Prior knowledge	Cognitive load	0.017	1	0.017	0.044	0.836	0.001
	Intrinsic motivation	0.003	1	0.003	0.018	0.895	0.000
	Post-test	1,050.625	1	1,050.625	9.310	0.004	0.205
Signaling* prior knowledge	Cognitive load	46.584	1	46.584	117.388	0.000	0.765
	Intrinsic motivation	0.331	1	0.331	1.777	0.191	0.047
	Post-test	180.625	1	180.625	1.601	0.214	0.043

a. R^2 = 0.766 (Adjusted R^2 = 0.747). b. R^2 = 0.103 (Adjusted R^2 = 0.029). c. R^2 = 0.252 (Adjusted R^2 = 0.190).
*Interaction effect between signaling and learners' prior knowledge level.

TABLE 2 Descriptive statistical analysis results.

	VR version	Prior knowledge	Mean	Standard deviation	*N*
Intrinsic motivation	Signal	Low	4.500	0.423	10
		High	4.300	0.479	10
		Total	4.400	0.452	20
	No signal	Low	4.109	0.475	10
		High	4.273	0.332	10
		Total	4.191	0.408	20
	Total	Low	4.305	0.481	20
		High	4.286	0.401	20
		Total	4.295	0.438	40
Cognitive load	Signal	Low	3.067	0.545	10
		High	5.1837	0.388	10
		Total	4.1257	1.180	20
	No signal	Low	5.383	0.4161	10
		High	3.183	0.983	10
		Total	4.283	1.347	20
	Total	Low	4.225	1.279	20
		High	4.183	1.258	20
		Total	4.204	1.252	40
Post-test	Signal	Low	37.00	6.749	10
		High	43.00	8.233	10
		Total	40.00	7.947	20
	No signal	Low	29.00	13.904	10
		High	43.50	12.030	10
		Total	36.25	14.679	20
	Total	Low	33.00	11.402	20
		High	43.25	10.036	20
		Total	38.13	11.804	40

experiment were provided by the researchers. Researchers explained the detailed experimental tasks and the use of the IVR experimental equipment so that the participants could be initially familiar with the requirements for operating virtual experimental environments and tasks. The IVR experiment was individually conducted for each participant. A research assistant helped participants to put on and adjust the HTC Vive Pro headset, and the participants began to learn and operate in an IVR laboratory environment. The learning content and tasks of the version with signals and those of the version without signals were the same. The version with signals included text annotations and object highlighting, which helped the learners select the experimental operation and access the key emphatic information of the selected object. Once preparations were complete, the experimental group and the control group wore HTC Vive Pro devices and started the IVR experiment. After the learners finished the low-carbon steel tensile strength experiment, they completed the learning performance test questionnaires, cognitive load scale, and intrinsic motivation scale and finally were led in a semi-structured interview. The quantitative data were analyzed by SPSS. The complete experiment time was approximately 100–120 min, and the experimental procedure is shown in **Figure 3**.

3.5. Interview

Three students were randomly selected from each of the four experimental groups, for a total of 12 students, and each student was interviewed for 20–30 min. Riches et al. (2019) proposed a theoretical framework with which to explore the factors affecting sense of presence in a VR environment, which included three dimensions: thoughts, emotions, and behaviors. Based on the theoretical framework, the interview questions comprised three categories. The first category involves the participants' thoughts on signaling, e.g., are signals helpful to you? The second concerns the participants' views on learning in an IVR environment, e.g., what parts of the IVR experiment scene appeal to you the most? The third is the comparison of the IVR laboratory with the traditional laboratory in terms of participants' behaviors in the IVR, e.g., comparing the differences and making suggestions for improvement. The research team transcribed the interviews into text, and the three researchers coded the interviews and subsequently summarized them into themes. The three researchers compared the codes and themes and negotiated by presenting evidence (in the event of disagreements) to finally determine the following themes: operational feelings, recognition, and inadequacy.

4. Results

4.1. Effects of prior knowledge and signaling on cognitive load

To test whether signaling makes a difference in the cognitive load of learners with different prior knowledge levels, an ANOVA was conducted. The results showed that the main effect of signaling was not significant, $F(3, 34) = 0.632$, $p = 0.432$, $\eta 2 = 0.017$, as shown in **Table 1**. The main effect of learners' prior knowledge level was not significant, $F(3, 34) = 0.044$, $p = 0.836$, $\eta 2 = 0.001$. The interaction effect between signaling and learners' prior knowledge level was significant, $F(3, 34) = 117.388$, $p = 0.000$, $\eta 2 = 0.765$, *p*-value less than 0.05. From the range of effect sizes, 0.14 or above indicated a large effect on the dependent variable. The effect size of learners' prior knowledge level was 0.765, which is much larger than 0.14, thus indicating that the interaction effect of signaling and learners' prior knowledge level in the IVR laboratory environment was the main factor causing significant changes in learners' cognitive load. Specifically, for learners with a low prior knowledge level, signaling allows them to focus more on meaningful visual elements and reduces cognitive load in the IVR laboratory. On the other hand, learners with a high prior knowledge level had a lower cognitive load in the environment without signaling. Descriptive statistics showed that learners in the experimental condition without signaling ($M = 4.283$, $SD = 1.347$) had a higher cognitive load than those with signaling ($M = 4.125$, $SD = 1.180$), but the difference was not significant, as shown in **Table 2**. In terms of generated cognitive load, learners with a high prior knowledge level ($M = 4.183$, $SD = 1.258$) generated a lower cognitive load in the IVR laboratory than learners with a low prior knowledge level ($M = 4.225$, $SD = 1.279$), but the difference was not significant.

4.2. Effects of prior knowledge and signaling on intrinsic motivation

To test the different influences of signaling and prior knowledge level on learners' intrinsic motivation in the IVR laboratory, an ANOVA was conducted with the intrinsic motivation of the subjects as the dependent variable. The results showed that the main effect of learners' prior knowledge level was not significant, $F(3, 34) = 0.018$, $p = 0.895$, $\eta 2 = 0.000$, *p*-value greater than 0.05. The main effect of signaling on intrinsic motivation was not significant, $F(3, 34) = 2.351$, $p = 0.134$, $\eta 2 = 0.061$, *p*-value greater than 0.05. The interaction effect between signaling and learners' prior knowledge level was not significant, $F(3, 34) = 1.777$, $p = 0.191$, $\eta 2 = 0.047$, *p*-value greater than 0.05. While descriptive statistical analysis revealed that learners with low prior knowledge levels were more motivated to learn ($M = 4.305$, $SD = 0.481$) than learners with high prior knowledge levels ($M = 4.286$, $SD = 0.401$), the difference was not significant. The results of the descriptive statistical analysis showed that while the IVR laboratory with signaling ($M = 4.400$, $SD = 0.452$) was more likely to stimulate learners' intrinsic motivation than that without signaling ($M = 4.191$, $SD = 0.408$), the difference was not significant. The results of the data analysis indicated that there was no significant difference between the conditions of signaling and prior knowledge level in the IVR laboratory in terms of triggering learners' intrinsic motivation, which remained at a high level.

4.3. Effects of prior knowledge and signaling on learning

To test whether signaling can affect learning performance and whether it differently affects learners who have different prior knowledge levels, an ANOVA was conducted with the participants' learning performance as the dependent variable. The results show that the main effect of learners' knowledge and experience level was significant, $F(3, 34) = 9.310$, $p = 0.004$, $\eta 2 = 0.205$, *p*-value less than 0.05. The presence of signaling was not significant, $F(3, 34) = 1.246$, $p = 0.272$, $\eta 2 = 0.033$, *p*-value greater than 0.05. The interaction effect of signaling and learners' prior knowledge level was not significant, $F(3, 34) = 1.601$, $p = 0.214$, $\eta 2 = 0.043$, *p*-value greater than 0.05. At the same time, based on effect size, learners' prior knowledge levels in the IVR laboratory were the main factor causing changes in learners' learning performance with a large effect size of 0.205, which was above 0.14. This result indicated that changes in learning performance were mainly influenced by learners' prior knowledge level, and there was no significant difference between the presence of signaling and the interaction effect of signaling and the learners' prior knowledge level on their learning performance. Based on the descriptive statistical analysis results, prior knowledge level had a more significant effect on the subjects' learning performance in the IVR laboratory, and learners with a high prior knowledge level (M = 43.25, SD = 10.036) obtained better learning performance compared to the post-test performance of learners with a low prior knowledge level (M = 33.00, SD = 11.402). In the IVR laboratory with signaling (M = 40.00, SD = 7.947) learners achieved higher learning performance than those in that with no signaling (M = 36.25, SD = 14.679), but the difference was not significant.

4.4. Qualitative results

This study adopted the directed qualitative analysis method, which is guided by a more structured process than in a conventional approach (Hickey and Kipping, 1996). Directed qualitative analysis is a more appropriate approach when there are already existing theories about the phenomenon under investigation (Hsieh and Shannon, 2005). Based on Riches et al. (2019)'s theoretical framework, the study reported three themes: (1) learners' operational feelings about the virtual experiment (thoughts), (2) learners' attitudes of the virtual experiment (emotions), and (3) based on learners' behaviors and experiences in VR, their opinions about the inadequacies of the virtual experiment.

Regarding learners' operational feelings about the virtual experiment, the students in the IVR laboratory environment with signals reported that the "object highlighting" was very helpful for experimentation and that there were some differences between experimentation in the virtual environment and the real scene, mainly in the way the objects were touched. In this study, the IVR controller was activated to pick up and move the object and select the button. Some students said, for example:

"The object highlighting is a quick way to determine whether the object has been touched."

They were able to operate more smoothly than learners in an IVR laboratory environment without signals, avoiding the impact of uncertainty on experimental procedures. Some learners also found the "text annotation" signal helpful, and this was generally true for learners with a low prior knowledge level. These less experienced learners thought the text annotations were intuitive and concise and found it easy to understand and proceed to the next step due to the clear cues throughout the experiment. However, a small number of students with a high prior knowledge level thought that the text annotations interfered with their own thinking and implementation of the experimental procedure, for example:

"*I was thinking about what to do next, but the text annotations interrupted my thinking.*"

Many students said, for example:

"*The content of the experiment is very interesting, and it is a novel experience to be able to conduct experiments virtually.*".

This is because most of the learners were performing an IVR experiment for the first time and had a great interest in IVR. Therefore, prior knowledge level and the VR version had little effect on the intrinsic motivation of learners, and learners were very willing to perform similar experimental operations in immersive virtual scenes. Six students with high prior knowledge levels reported that the design of the virtual laboratory scenes was very realistic, for example:

"*The electronic universal testing machine, Vernier callipers and other experimental instruments are vividly rendered, and the setting of the laboratory makes people feel the authenticity of the scene, with a strong sense of situational immersion.*"

Thus, the sense of immersion is an important factor in stimulating learners' interest and motivation. Object highlighting, text annotations and experimental content are the main factors that affect learners' implementation of the procedures in the IVR laboratory. These signals further affect the extraneous cognitive load and intrinsic motivation of learners during the experiments, while different signals in the experiments influence learners' cognitive processes and determine whether they are alert to the details of low-carbon steel stretching during the experiment (e.g., low-carbon steel failure process, fracture shape, etc.), thus impacting the post-test learning performance.

Regarding learners' attitudes of the virtual experiment, most of the students in the interviews affirmed the effectiveness of the immersive virtual environment designed in this study in the following ways: (1) They thought the experiment was interesting. Some students said, for example:

"*I like operating the experimental equipment in the virtual scene*" and "*It made me want to understand how this experiment works.*"

(2) It contributes to inquiry-based learning. For example, a student stated the following:

"*I can control my own learning pace by doing experiments in this kind of virtual scene, and I can repeat steps that I am not familiar with.*"

(3) It is helpful for observation and memory. For example, a student mentioned the following:

"*In the immersive virtual laboratory, I am very relaxed while learning the content and can observe the experimental phenomena at close range to form a strong memory of knowledge such as the fracture section shape of low-carbon steel when it was stretched.*"

(4) They are willing to learn the experiment again. Most students believed that doing the experiment was of certain value to themselves and expressed their willingness to complete the experiment in an IVR laboratory. According to the overall analysis, the IVR laboratory had a relatively positive impact on the learners, which in turn stimulated their motivation and interest in learning.

Regarding inadequacies of the virtual experiment, by sorting out the questions raised by the interviewees about the IVR experiment, the following two reflections on inadequacies are proposed: (1) In terms of VR technology, IVR technology can provide high-quality display resolution, a wider field of view and an experience that stimulates multiple sensory modalities, but at the same time, this experience may bring physiological discomfort to some participants, such as those with vertigo. In this experiment, the IVR controller was used to achieve movement in the virtual scene, and the sensitivity of the IVR controller led to a certain discomfort caused by the fast roaming speed in the scene, which had the potential to produce vertigo. Vertigo would affect the whole experience of the IVR laboratory and might increase the extraneous cognitive load of the learners. The HTC headset would cause a certain discomfort for some learners who wear glasses, and at the same time, its weight would make the students uncomfortable. (2) In this particular experiment, learners who were unfamiliar with the low-carbon steel stretching experiment itself (learners with a low prior knowledge level) were more focused on the operation of the VR itself in the virtual environment and might have overlooked some details of the experiment that should have been observed. Some learners reported that there was a gap between their spatial cognition ability in the virtual scene and their ability in real life, and they could not distinguish left from right, which affected their own operation and judgment.

5. Discussion

This study explores the effects of signaling and different prior knowledge levels (high vs. low) on learners' cognitive load, intrinsic motivation, and learning performance in the IVR laboratory. Several findings have been presented to support that the factor of signaling has a significant effect on reducing learners' cognitive load (e.g., Arslan-Ari, 2018; Albus et al., 2021). This study found that the interaction effect of signaling and prior knowledge level on learners' cognitive load in the IVR laboratory was significant and that an expertise/experience reversal effect on cognitive load emerged, i.e., signaling had a negative effect on the cognitive load of learners with high prior knowledge levels but had a positive effect on that of learners with low prior knowledge levels. The empirical findings of this study demonstrate that signaling in the IVR laboratory can effectively help learners with low prior knowledge operate experiments and observe experimental phenomena, providing greater support for reducing learners' cognitive load. The findings provide suggestions and guidance for the design of IVR experiments and their application in teaching practice.

5.1. Effect of signaling

This study analyses the effects of signaling on students in an IVR laboratory and finds no significant differences in cognitive load, intrinsic motivation, or learning performance between the two groups of participants with and without signaling. From the results of multivariate ANOVA, in the analysis of the effects of only signaling, learners' learning performance is higher in the VR version with signals (M = 40.00, SD = 7.947) than in the VR version without signals (M = 36.25, SD = 14.679), learners' intrinsic motivation is higher in the VR version with signals (M = 4.400, SD = 0.452) than without signals (M = 4.191, SD = 0.408), and learners' cognitive load is lower in the VR version with signals (M = 4.126, SD = 1.180) than without signals (M = 4.283, SD = 1.347), but there is no significant difference in the effect of signaling in the IVR laboratory on learning performance, intrinsic motivation and cognitive load.

This result is inconsistent with cognitive load theory and the cognitive–emotional theory of multimedia learning, both of which suggest that designing signals in multimedia learning materials that are highly relevant to the teaching content can help learners reduce visual retrieval and lower cognitive load (Fiorella and Mayer, 2014). However, there are still some researchers whose findings are consistent with the present study. Nelson et al. (2014) hypothesized that the use of visual signals could reduce the overall cognitive load of students when completing virtual world-based tasks, yet the findings showed only small significant differences in cognitive load between the groups. The reason given in their analysis was that the virtual scenes themselves were designed to contain fewer interactive objects and limited virtual space for exploration, which has been found to reduce learners' cognitive load. It has also been suggested that signaling only improves learning performance with complex learning materials (Albus et al., 2021). Similarly, Jeung et al. (1997) suggested that adding visual signals to learning materials that required high levels of visual searching facilitated learning, while adding visual signals to conditions with low levels of visual searching had little effect on learning. The IVR experiment scene designed in this study is as simple and direct as possible, with no redundant or irrelevant scene elements. Thus, the visual impact on learners does not reach the level of so-called "complexity" that would increase cognitive load. However, it is also clear that the design of IVR laboratories should not be overloaded with visual overlays, as this would result in a visual input burden and cognitive processing overload for learners. This suggests that the application of signaling needs to be designed in conjunction with the actual virtual scene to provide the most concise and effective guidance for learners.

In terms of the types and content of signaling, this study uses signals in the form of textual annotations and color cues, which allow learners to receive signals (Baceviciute et al., 2020) in a visually demanding virtual environment through short and concise signals. Knowing which information is relevant now and which needs to be looked at more closely can reduce unnecessary searches and draw attention to relevant aspects. However, as some of the text involves technical terms (e.g., manual control box, fixture, etc.), it can be difficult for learners with a low prior knowledge level to understand, and they need to invest mental effort in the operation of the experiment, thus potentially increasing the cognitive load and impacting learning effects. Furthermore, the findings of Richter et al. (2016) suggest that compared with multimedia integrated signals, learning materials with a more prominent visual appearance, such as color coding, may be more accessible to learners than discursive signals as well as descriptive references. In this study, signals are presented in the form of dialog boxes in the scene. The study could have better embedded the signals into the environment, and learners' sense of learning experience may have been affected (Albus et al., 2021). Thus, in signal design, the learning situation can be combined with the laboratory design, dynamic signals can be combined with static signals, different depths and elaborations of the integration of text and images can be made, and other multimedia design principles, such as the redundancy principle and the spatial proximity principle, can be combined to provide good signal effects for learners (Lai et al., 2019).

5.2. Effect of prior knowledge level

It is found that learners' prior knowledge level plays an important role in influencing learning performance, intrinsic motivation and cognitive load in IVR laboratories. From the results of multivariate ANOVA, in the analysis of only the effect of learners' prior knowledge level, the learning performance of learners with a high prior knowledge level (M = 43.25, SD = 10.036) is significantly higher than that of learners with a low prior knowledge level (M = 33.00, SD = 11.402), indicating that prior knowledge has a significant impact on learning performance in an IVR laboratory. Intrinsic motivation was slightly lower in learners with a high prior knowledge level (M = 4.286, SD = 0.401) than in learners with a low prior knowledge level (M = 4.305, SD = 0.481). Cognitive load was slightly lower in learners with a high prior knowledge level (M = 4.183, SD = 1.258) than in learners with a low prior knowledge level (M = 4.225, SD = 1.279), indicating that learners with high prior knowledge levels had higher learning performance and lower cognitive load overall when learning in the immersive virtual laboratory, which is a finding consistent with most studies (Kalyuga et al., 2004; Kriz and Hegarty, 2007). Another study showed that learners' attitude-learning-confidence is strongly related (Taçgın, 2020) and that learners with high prior knowledge levels have more positive attitudes and interest in learning in virtual learning environments. In summary, it is recommended that novice learners familiarize themselves with the IVR technology before using the system as a learning tool and improve their learning effects by increasing their confidence or attitude with repeated practice. These findings are almost unanimous in studies that address only the effect of prior knowledge level on learning effects, but with the addition of some intervening factors for moderation, there may be an interaction effect, and the results may vary.

5.3. Interaction effect of prior knowledge level and signaling

Signaling is able to help learners with low prior knowledge levels reduce their cognitive load generated by completing the immersive virtual experiment task, enhance their intrinsic motivation, and achieve better performance on the final test, while learners with high prior knowledge levels had better learning performance without these additional signals (Kalyuga, 2007; Richter et al., 2016). The results of the data analysis support the hypothesis that the effectiveness of signaling in the IVR laboratory differs for learners with different prior knowledge levels, indicating that the expertise reversal effect of signaling on learners is as valid in an IVR laboratory as in the traditional multimedia learning environment (Johnson et al., 2013, 2014; Khacharem, 2017; Arslan-Ari, 2018). Therefore, in conducting IVR laboratories, it is quite effective to direct the attention of learners with low prior knowledge levels to help them select relevant information. From this study, the design of synchronous textual signals on the microcomputer-controlled electronic universal testing machine used for completing experimental tasks can direct learners' attention to the aspects of the scene related to conducting the experiment. Thus, these synchronous signals reduce external load and help learners use their mental resources to build coherent mental representations of the target learning material and integrate these representations into long-term memory (Arslan-Ari, 2018). In addition to text signals, color changes are used as signals in this study, which play an important role in IVR experiments. Based on learner feedback, the largest difference between immersive and realistic laboratories is the selection of relevant objects. Most existing virtual reality technology uses the IVR controller for interaction, and learners have no real sense of touch when they interact with the corresponding scene elements (Streppel et al., 2018). There can be confusion about whether they have picked up the object, which affects the thinking and cognitive load and can also reduce the learning experience. Using color changes to signal that an object has been touched helps learners carry out the operation steps smoothly and coherently. Multimedia learning cognitive–affective theory, cognitive load theory and other theories on media learning consider the learner's prior knowledge base to be the most important moderator. Previous studies have proven that signal transmission can enhance students' knowledge retention ability (Boucheix et al., 2013) and knowledge transfer ability in problem-based learning (Ouwehand et al., 2014). Although the interaction effect of the two is not significant for learners' intrinsic motivation and learning performance, the results of the data analysis suggest that the joint effect of signaling and prior knowledge level on learners is an area worthy of more research and an existing research trend (Arslan-Ari, 2018). Learners with low prior knowledge levels are slightly more intrinsically motivated than learners with high prior knowledge levels in the condition with signaling, while in the condition without signaling, learners with high prior knowledge had higher intrinsic motivation.

The results of this study also demonstrate the need for personalized learning and the need for developers to adopt a different mindset in the design and development of immersive virtual experiments to cater to differences between learners with different prior knowledge levels and other characteristics (Kalyuga, 2007). Based on the above findings and the characteristics of immersive virtual laboratories, developers need to reintegrate the association between the prior knowledge level and signaling (Kalyuga, 2008), design signal types that are highly relevant to the experimental content and virtual scenes, tailor the available information to the prior knowledge levels of target students, and ensure that the new material is consistent with their original mental representations. VR designers could offer VR participants the option to toggle between signals and no-signals in the same way that a TV viewer can toggle between closed captioning and no closed captioning. If a student is getting distracted/annoyed by the dialog boxes in this low carbon steel tensile strength experiment, for example, they could just opt to disable the dialog boxes.

6. Conclusion and implications

The core purpose of this study is to explore the effect of signaling and prior knowledge on learning effects in an IVR laboratory. This study suggests that the design of IVR experiments should be combined with the characteristics of different learners. For learners with high prior knowledge levels, only the necessary signals should be designed and offered as resources, as additional signals will result in higher cognitive load with a negative effect on the learning experience and potentially the final learning performance. When learners' prior knowledge is low, appropriate signals can help to reduce cognitive load and improve learning performance. Future research may recruit more participants to obtain a better effect size. The study applied NASA-TLX to measure participants' cognitive load. Despite being widely used as a measure of cognitive load, the NASA-TLX has several limitations. More empirical research on measuring cognitive load is needed. Future empirical studies should also consider using multimodal channels to measure learners' cognitive load, such as EEG, ECG, EDA, DDT, and eye tracking, which could more accurately gather data and deeply explore the mechanisms related to cognitive load. The IVR experimental scene in this study is designed with only two types of signals, text annotation and color, and future research can further analyze the effects of signal type, complexity, level of detail, and information content on learners.

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 Dr. Shengquan Luo, Faculty of Education, Southwest University. The patients/participants provided their written informed consent to participate in this study. 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

JH: conceptualization, investigation, methodology, data collection, writing – review and editing, and validation. GL: conceptualization, writing – review and editing, and supervision. QZ: data collection and formal analysis. All authors have read and agreed to the published version of the manuscript.

Funding

This study was financially supported by the 2020 Chongqing Higher Education Teaching Reform Research Project "Research on the Deep Integration of Information Technology and University Teaching" under Grant "201004", "the Fundamental Research Funds for the Central Universities" under Grant "SWU1909764", the "Southwest University Education and Teaching Reform Research Project" under Grant "2020JY038", and the "Southwest University Continuing Education Research Project" under Grant "SWU2108028".

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

Albus, P., Vogt, A., and Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Comput. Educ.* 166:104154. doi: 10.1016/j.compedu.2021.104154

Arslan-Ari, I. (2018). Learning from instructional animations: How does prior knowledge mediate the effect of visual cues? *J. Comput. Assist. Learn.* 34, 140–149. doi: 10.1111/jcal.12222

Baceviciute, S., Mottelson, A., Terkildsen, T., and Makransky, G. (2020). "Investigating representation of text and audio in educational VR using learning outcomes and EEG," in *Proceedings of the 2020 CHI conference on human factors in computing systems*, (New York, NY: ACM). doi: 10.1145/3313831.3376872

Baceviciute, S., Terkildsen, T., and Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. *Comput. Educ.* 164:104122. doi: 10.1016/j.compedu.2020.104122

Birbara, N. S., and Pather, N. (2021). Instructional design of virtual learning resources for anatomy education. *Adv. Exp. Med. Biol.* 1317, 75–110. doi: 10.1007/978-3-030-61125-5_5

Boucheix, J.-M., Lowe, R. K., Putri, D. K., and Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learn. Instr.* 25, 71–84. doi: 10.1016/j.learninstruc.2012.11.005

Braune, R., and Foshay, W. R. (1983). Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. *Instruct. Sci.* 12, 121–145. doi: 10.1007/BF00122453

Buckingham, G., Michelakakis, E. E., and Rajendran, G. (2016). The influence of prior knowledge on perception and action: Relationships to autistic traits. *J. Autism Dev. Disord.* 46, 1716–1724. doi: 10.1007/s10803-016-2701-0

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Sci. Educ.* 90, 1073–1091. doi: 10.1002/sce.20164

de Koning, B., Tabbers, H., Rikers, R., and Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educ. Psychol. Rev.* 21, 113–140. doi: 10.1007/s10648-009-9098-7

Ferrara, L. A., and Butcher, K. R. (2011). Visualizing feedback: Using graphical cues to promote self-regulated learning. *Cogn. Sci.* Soc. 33, 1880–1885.

Fiorella, L., and Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemp. Educ. Psychol.* 39, 75–85. doi: 10.1016/j.cedpsych.2014.01.001

Han, J., Zheng, Q., and Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *J. Adv. Inf. Technol.* 12, 302–310. doi: 10.12720/jait.12.4.302-310

Hickey, G., and Kipping, C. (1996). Issues in research. A multi-stage approach to the coding of data from open-ended questions. *Nurse Res.* 4, 81–91.

Hsieh, H. F., and Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qual. Health Res.* 15, 1277–1288.

Hwang, Y., and Shin, D. (2018). Visual cues enhance user performance in virtual environments. *Soc. Behav. Pers.* 46, 11–24. doi: 10.2224/sbp.6500

Jensen, L., and Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Educ. Inf. Technol.* 23, 1515–1529. doi: 10.1007/s10639-017-9676-0

Jeung, H., Chandler, P., and Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educ. Psychol.* 17, 329–345. doi: 10.1080/0144341970170307

Johnson, A. M., Ozogul, G., and Reisslein, M. (2014). Supporting multimedia learning with visual signalling and animated pedagogical agent: Moderating effects of prior knowledge. *J. Comp. Assist. Learn.* 31, 97–115. doi: 10.1111/jcal.12078

Johnson, A., Ozogul, G., Moreno, R., and Reisslein, M. (2013). Pedagogical agent signaling of multiple visual engineering representations: The case of the young female agent. *J. Eng. Educ.* 102, 319–337. doi: 10.1002/jee.20009

Johnson, M. A., and Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *J. Res. Sci.Teach.* 35, 89–103. doi: 10.1002/(SICI)1098-2736(199801)35:1<89::AID-TEA6>3.0.CO;2-J

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educ. Psychol. Rev.* 19, 509–539. doi: 10.1007/s10648-007-9054-3

Kalyuga, S. (2008). When less is more in cognitive diagnosis: A rapid online method for diagnosing learner task-specific expertise. *J. Educ. Psychol.* 100, 603–612. doi: 10.1037/0022-0663.100.3.603

Kalyuga, S., Chandler, P., and Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Hum. Fact.* 46, 567–581. doi: 10.1518/hfes.46.3.567.50405

Khacharem, A. (2017). Top-down and bottom-up guidance in comprehension of schematic football diagrams. *J. Sports Sci.* 35, 1204–1210. doi: 10.1080/02640414.2016.1218034

Kim, S., and Rehder, B. (2011). How prior knowledge affects selective attention during category learning: An eyetracking study. *Memory Cogn.* 39, 649–665. doi: 10.3758/s13421-010-0050-3

Kriz, S., and Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *Int. J. Hum. Comput. Stud.* 65, 911–930. doi: 10.1016/j.ijhcs.2007.06.005

Lai, A.-F., Chen, C.-H., and Lee, G.-Y. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *Br. J. Educ. Technol.* 50, 232–247. doi: 10.1111/bjet.12716

Lin, L., and Li, M. (2018). Optimizing learning from animation: Examining the impact of biofeedback. *Learn. Instr.* 55, 32–40. doi: 10.1016/j.learninstruc.2018.02.005

Makransky, G., Andreasen, N. K., Baceviciute, S., and Mayer, R. (2020). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *J. Educ. Psychol.* 113, 719–735. doi: 10.1037/edu0000473

Makransky, G., Wismer, P., and Mayer, R. (2019b). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *J. Comput. Assist. Learn.* 35, 349–358. doi: 10.1111/jcal.12335

Makransky, G., Terkildsen, T. S., and Mayer, R. E. (2019a). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learn. Instruct.* 60, 225–236. doi: 10.1016/j.learninstruc.2017.12.007

Mayer, R. E. (2020). *Multimedia learning*, 3rd Edn. Cambridge: Cambridge University Press, doi: 10.1017/9781316941355

Mayer, R., Sobko, K., and Mautone, P. (2003). Social cues in multimedia learning: Role of speaker's voice. *J. Educ. Psychol.* 95, 419–425. doi: 10.1037/0022-0663.95.2.419

Mishra, L., Gupta, T., and Shree, A. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *Int. J. Educ. Res. Open* 1:100012. doi: 10.1016/j.ijedro.2020.100012

Nelson, B. C., Kim, Y., Foshee, C., and Slack, K. (2014). Visual signaling in virtual world-based assessments: The SAVE science project. *Inf. Sci.* 264, 32–40. doi: 10.1016/j.ins.2013.09.011

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *J. Res. Sci. Teach.* 27, 937–949. doi: 10.1002/tea.3660271003

Ouwehand, K., Gog, T., and Paas, F. (2014). Effects of gestures on older adults' learning from video-based models. *Appl. Cogn. Psychol.* 29, 115–128. doi: 10.1002/acp.3097

Parong, J., and Mayer, R. (2018). Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology* 110, 785–797. doi: 10.1037/edu0000241

Parong, J., and Mayer, R. E. (2021a). Learning about history in immersive virtual reality: Does immersion facilitate learning? *Educ. Technol. Res. Dev.* 69, 1433–1451. doi: 10.1007/s11423-021-09999-y

Parong, J., and Mayer, R. E. (2021b). Cognitive and affective processes for learning science in immersive virtual reality. *J. Comput. Assist. Learn.* 37, 226–241. doi: 10.1111/jcal.12482

Petersen, G. B., Klingenberg, S., Mayer, R. E., and Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *Br. J. Educ. Technol.* 51, 2098–2114. doi: 10.1111/bjet.12991

Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Comput. Educ.* 147:103778. doi: 10.1016/j.compedu.2019.103778

Riches, S., Elghany, S., Garety, P., Rus-Calafell, M., and Valmaggia, L. (2019). Factors affecting sense of presence in a virtual reality social environment: A qualitative study. *Cyberpsychol. Behav. Soc. Netwo.* 22, 288–292. doi: 10.1089/cyber.2018.0128

Richter, J., Scheiter, K., and Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educ. Res. Rev.* 17, 19–36. doi: 10.1016/j.edurev.2015.12.003

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *J. Pers. Soc. Psychol.* 43, 450–461. doi: 10.1037/0022-3514.43.3.450

Streppel, B., Pantförder, D., and Vogel-Heuser, B. (2018). "Interaction in virtual environments - How to control the environment by using VR-glasses in the most immersive way BT," in *Proceedings of the virtual, augmented and mixed reality: Interaction, navigation, visualization, embodiment, and simulation*, eds J. Y. C. Chen and G. Fragomeni (Las Vegas, NY: Springer International Publishing), 183–201.

Taçgın, Z. (2020). The perceived effectiveness regarding Immersive Virtual Reality learning environments changes by the prior knowledge of learners. *Educ. Inf. Technol.* 25, 2791–2809. doi: 10.1007/s10639-019-10088-0

Twining, P., Heller, R. S., Nussbaum, M., and Tsai, C. C. (2017). Some guidance on conducting and reporting qualitative studies. *Comput. Educ.* 106, A1–A9. doi: 10.1016/j.compedu.2016.12.002

Van Gog, T. (2014). "The signaling (or cueing) principle in multimedia learning," in *The Cambridge handbook of multimedia learning*, 2nd Edn, ed. R. E. Mayer (Cambridge: Cambridge University Press), 263–278. doi: 10.1017/CBO9781139547369.014

Vogt, A., Klepsch, M., Baetge, I., and Seufert, T. (2020). Learning from multiple representations: Prior knowledge moderates the beneficial effects of signals and abstract graphics. *Front. Psychol.* 11:601125. doi: 10.3389/fpsyg.2020.601125

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 15 March 2023
DOI 10.3389/fpsyg.2023.1139884

Check for updates

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University,
China

REVIEWED BY
Dengkang Chen,
Central China Normal University,
China
Nyet Moi Siew,
Universiti Malaysia Sabah,
Malaysia

*CORRESPONDENCE
Eva Schmidthaler
✉ eva.schmidthaler@jku.at

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

RECEIVED 07 January 2023
ACCEPTED 16 February 2023
PUBLISHED 15 March 2023

CITATION
Schmidthaler E, Schalk M, Schmollmüller M, Hinterplattner S, Hörmann C, Anđić B, Rottenhofer M, Lavicza Z and Sabitzer B (2023) The interdisciplinary implementation of poly-universe to promote computational thinking: Teaching examples from biological, physical, and digital education in Austrian secondary schools.
Front. Psychol. 14:1139884.
doi: 10.3389/fpsyg.2023.1139884

The interdisciplinary implementation of poly-universe to promote computational thinking: Teaching examples from biological, physical, and digital education in Austrian secondary schools

Eva Schmidthaler[1]*, Maritta Schalk[1], Mathias Schmollmüller[1], Sara Hinterplattner[2], Corinna Hörmann[1], Branko Anđić[1], Marina Rottenhofer[1], Zsolt Lavicza[1] and Barbara Sabitzer[1]

[1]Johannes Kepler University, School of Education, Linz, Austria, [2]Dynatrace, Linz, Austria

Today's teaching and didactical methods are progressively aiming to integrate digital technologies, computational thinking (CT), and basic computer science concepts into other subjects. An innovative and creative way of combining and integrating CT and teaching cross-curricular skills without digital devices is to include the game Poly-Universe (PolyUni). According to previous research, the game is expected to have a positive effect on visual perceptual progress, including isolation, and the development of shape-background skills. So far, however, comparatively few attempts have been made to explore the educational possibilities of PolyUni for different school levels and subjects, besides mathematics. Therefore, this article aims to close this gap by exploring how PolyUni can be used to promote CT in three subjects: physical education (PE), digital education (DGE), and biology (B). Furthermore, it evaluates whether the pre-defined learning objectives in those subjects have been achieved, and examines how PolyUni combines the requirements of the different curricula in Austrian secondary school, based on self-designed tasks. Additionally, further aspects of PolyUni such as engagement and collaboration are discussed. To explore the above-mentioned benefits, a mixed-methods study was implemented, whereas the workshops and accompanying teaching materials (e.g., worksheets) were developed based on the COOL Informatics concept. The participant observation method was employed for qualitative data collection, and a self-designed assessment grid as well as additional picture analysis were used for the quantitative data. PolyUni was introduced in three different workshops at Austrian secondary schools with 80 students observed and analyzed. Based on the present data, it can be assumed that PolyUni supports achieving the requirements of the different curricula and pre-defined teaching and learning objectives in a playful way. Furthermore, the game not only promotes CT in secondary school but also encourages enjoyment and collaboration between peers in biological, digital, and physical education lessons.

KEYWORDS

poly-universe, computational thinking, digital education, game-based learning, biology, physical education, stem

1. Introduction

Computational thinking (CT) is a term that has become more and more present in the last years and decades, especially since Jeanette Wings used the expression in 2006 (Selby and Woollard, 2013). However, when looking at the term, one does not find a clear unambiguous definition. Often CT is used as a catch-all term for problem-solving strategies (Selby and Woollard, 2013; Curzon and McOwan, 2018). Abstraction, decomposition, generalization, and evaluation around algorithms are techniques of these mentioned strategies. Selby and Woollard (2013) have shown in their paper how multi-layered the term is and that it can be viewed from four perspectives - thinking term, problem-solving term, computer science term, and imitation term (Selby and Woollard, 2013). Despite this multi-layered approach after a search for a definition, no clear explanation was found. In this context, the question arises whether the term will not change in the next few years anyway, since CT is also evolving. In this study, as according to Curzon and McOwan (2018), CT is considered a problem-solving technique, which includes the following important elements: decomposition, pattern recognition, generalization, abstraction, algorithms, and evaluation. Basically, it is about ideas of reasoning and choosing a good representation of data for the problems at hand (Curzon and McOwan, 2018). Furthermore, creativity also plays an important role in the scientific thought process and CT (Curzon and McOwan, 2018; Israel-Fishelson et al., 2021).

Regardless of an imprecise definition of the term, its importance in school is undisputed. CT has already found its way into higher education, but this is not quite the case in primary and secondary education (Settle et al., 2012). In Austria, digital literacy has been a statutory mandatory subject in secondary schools since the 2022/2023 school year, and CT is an important pillar of its curriculum (BMBWF, 2022c). Despite the further development and the integration of CT in various curricula, it is evident that creative and multifaceted problem-solving strategies are indispensable for daily life but also for individual school subjects (Barr and Stephenson, 2011; Selby and Woollard, 2013). Looking at STEAM (science, technology, engineering, arts, and mathematics, or applied mathematics) subjects in secondary school, "Biology" and "Physical Education" may not be perceived as having an obvious connection to CT. However, scientists are showing that the integrated CT concepts in a learning-by-modeling environment facilitate not only a deeper understanding of the subject-specific concepts and processes, especially biological processes (Naulakha and Bar-Joseph, 2011; Hutchins et al., 2020) and physical activity (Fritz, 2022), but may also support developing foundational computing skills and knowledge to support the future learning of advanced computer science (CS) concepts.

There are programs such as "Science through Sports" (Hammrich et al., 2021) but the scientific considerations for the implementation of CT in the sector of physical education are also increasing. This can be done by training analyses but also through problem-solving tasks. A study from Galoyan et al. (2022) combines athletic fields, such as track and field and basketball, with scientific topics, design thinking principles, and CT. In doing so, they also demonstrate many ways in which CT can be implemented through sports into everyday school life (Galoyan et al., 2022).

The same applies to the biological field: During the course "Introduction to Bioinformatic," biology students were taught concepts and skills related to bioinformatics and CT (such as problem-solving, abstract thinking, and pattern recognition). As a result of the lecture, participants were more engaged in the learning process, and they were able to understand new concepts (biological and CT) (Hong, 2009). As well, Goldberg et al. presented an interdisciplinary course in which CT and computer science concepts were integrated into students' lectures (e.g., math, biology, and art). To increase engagement and interest in CT concepts and the CS field later, it was tested in areas that might be beneficial and fit the requirements of the biology curriculum (e.g., algorithms for DNA-sequences, or data analysis in health education) (Goldberg et al., 2012). In a more recent study (2020), high school students showed science and computer-assisted learning gains after studying computer modeling within a science unit (Arastoopour Irgens et al., 2020).

1.1. The poly-universe game and application In education

"Poly-Universe" (PolyUni) originally was a geometric skill development game, designed by Saxon and Stettner (2019). The game PolyUni consists of 72 unique flat tiles in the form of triangles, squares, and (almost) circles (Figure 1). The novelty of PolyUni lies in the "scale-shifting" symmetry inherent to its geometric forms and color combination system (Saxon, 2018). During the Erasmus+ project called "Poly-Universe in School Education (PUSE)," further work was done on the game and its implementation was researched (Dardai, 2018). The game, originally designed for mathematics classes, is intended to promote geometric understanding and combinatorics, but can further be used in other contexts (e.g., entertainment and non-formal and formal education) (Saxon, 2018; Saxon and Stettner, 2019).

In addition to the analogue game, PolyUni was further developed so that it could also be used as a digital educational tool in primary and secondary education, aiming to enable the application (app) of this game in an online or hybrid learning environment. The online version of the game provides almost all the benefits that the game has in a physical environment, except for direct physical contact with the game sets. The online tasks of this game are enriched with animations based on the rules of the game. Participants can interact with the animations and at the end of the animation, check their knowledge of the game through several quiz questions. Poly-Universe in a digital environment could be also used in the GeoGebra platform (GeoGebra, 2022) for solving problems connected to the game. The methodology of using ICT (information and communications technology) tools during the teaching/learning procedure becomes more and more popular. With GeoGeogebra's assistance, it is possible to create interactive programs for the creation and further consideration (such as a 3D extension) of the spatial-geometric components that reflect Poly-Universe's characteristics (Saxon, 2018).

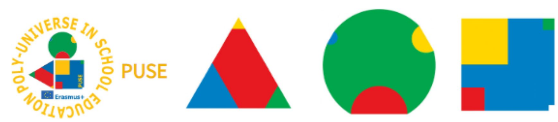


FIGURE 1
The Project's Logo (left), and the three Shapes of Poly-Universe Tiles (left) (Saxon, 2018).

There are no limits to using these individual game elements in the educational field: Besides the game's initial geometric-oriented tasks like "find the small triangles of the same size and connect them" or "find the small triangles of the same color and connect them" (Saxon, 2018), a variety of creative laying exercises are also practicable and are constantly being developed. During the Erasmus+ PUSE Project in 2008, some studies on the use and the effects of the game PolyUni in the classroom were included (Dardai, 2018). Hoffmann (2020) points out the potential of PolyUni for visualization and sees an opportunity to connect the subjects of "Geography," "Literature," "Sociology," "Political Science," and "Mathematics" with it. Within the PUSE Methodology and the PUSE study three teaching examples in the subject biology are presented: "substrate and enzyme connection," "examples of molecules model," and "characteristics of (flowering) plant families" could be created with the game elements by the student. The "PUSE Methodology" is a material collection of mathematical tasks for primary and secondary school students (suitable for six up to eighteen-year-olds), created within the PUSE project and developed by the project partners (Saxon and Stettner, 2019). The PUSE study was conducted in four different countries (Finland, Spain, Hungary, and Slovakia) to investigate the effects of the game including the following aspects: memory and mental rotation tests, visual perception, attention span, and attitude towards mathematical topics. A variety of methods, like cognitive tests, online questionnaires, and qualitative methods were used for the investigation. The most significant differences before and after testing were found in visual perception, including the isolation and development of shape background ability but also in the students' attitudes towards mathematics and related tasks. In both areas, the tests showed that the skills and opinions of the students improved significantly throughout PolyUni or changed positively because of the game (Dardai, 2018).

Moreover, besides the implementation in different school subjects, the game PolyUni can also be used within different age groups: from kindergarten (De Vasconcelos Martins et al., 2022) up to secondary school (Saxon, 2018; Saxon and Stettner, 2019). In a 2022 study, the Poly-Universe tiles were used even in Residential Care. The children were in kindergarten and primary school age (3 years and higher). The participants liked the game, because of the colorful elements, simplicity, and universality. The authors even claimed that PolyUni has the possibility to promote social and sensorimotor skills, spatial vision, and algorithmic thinking (De Vasconcelos Martins et al., 2022).

1.2. The potential of poly-universe to promote CT

To promote CT, various materials, such as educational games, can be used cross-curricular in different school subjects. Studies and reviews on game-based learning are sometimes contradictory because topic-related research is highly susceptible to a muddle of approaches and methodologies, as to whether games really contribute to learning success or whether they only promote enjoyment in the classroom (Vandercruysse et al., 2012; Hamari and Koivisto, 2015; Crocco et al., 2016; Hamari et al., 2016; Qian and Clark, 2016). A study conducted in 2020 looked at how to develop CT skills through game-based learning: It developed links between CT-based solutions and real-world problems using an adaptive learning game to promote CT skills and CT concepts (Hooshyar et al., 2021). Another option for promoting CT is through the analog educational game Poly-Universe.

In addition to the advantages from a subject-oriented point of view, this game also has an added value in CT. Due to the different variants of combining the parts, different strategies of problem-solving must be applied. "Which parts can I connect to another?," "If I need a square, what color and shape combinations must be included so that it fits into the figure?," "If I use this circle now, does that mean that I need a circle in the other color but with the same color as the small circle?" All these questions can happen when students try to create figures with PolyUni (Dardai, 2018). Experimenting with colors, shapes, and sizes in the form of the game offers many possibilities and opens space for creative work around CT and a wide variety of subjects and age groups (Dardai, 2018; Saxon, 2018; Saxon and Stettner, 2019).

Due to the game's uniqueness and possibility to promote CT als an innovative educational tool (Saxon, 2018; De Vasconcelos Martins et al., 2022), this article aims to show that teachers can utilize the game in other subjects, such as biology (B), digital education (DGE), and physical education (PE). Therefore, three workshop examples, and accompanying teaching and learning materials for B, PE, and DGE were developed by the authors or adapted from existing examples of the "PUSE Methodology. Poly-Universe in school education" by Saxon and Stettner (2019).

1.3. COOL informatics – Core concept for poly-universe workshop and material development

The teaching units with the game Poly-Universe were developed as three workshops (W1-3) on the basis of the "COOL Informatics" concept: discovery, individuality, cooperation, and activity. "COOL" as an abbreviation stands for different meanings: firstly for "COoperative Open Learning," an Austrian teaching model, and second, for "COmputer-supported Open Learning." Lastly, "COOL" stands for a popular sense of being interesting, motivating, fun, and effective (Sabitzer, 2013; Sabitzer and Stefan Pasterk, 2013).

The concept is a guide for teachers and provides suggestions for innovative, motivating, brain-friendly and supportive teaching that prepares for the demands of the 21st century (21st-century skills) (e.g., communication, collaboration, creativity, and critical thinking skills; Kennedy and Sundberg, 2020). It provides guidance, ideas, and examples for the preparation (planning and material development), the design of teaching units as well as the development and testing of special competencies. Previous research results on the basic concept of COOL Informatics (especially from computer science lessons) show that the consideration of these principles in the design of materials and teaching sequences can strengthen the motivation of learners, contribute to a better understanding of the content, and increase learning success. Figure 2 illustrates the COOL Informatics concept with its four main pillars. Each pillar includes the main teaching and learning methods as well as the neurodidactical foundation, such as pattern recognition, connecting new information to previous knowledge, joy, and constructivism (Piaget, 1971; Sabitzer, 2013; Sabitzer and Stefan Pasterk, 2013).

Within the workshops (W1-3), using step-by-step instructions (e.g., rules of the game), the students can utilize exploratory learning to solve a wide variety of solution-based tasks in their own learning rhythm (discovery). Furthermore, based on competence-based learning, and additional optional tasks with the game PolyUni, the individual needs, strength (e.g., endurance in PE), and interests of the students can be addressed (individuality). With the help of partner work as a teaching and learning methodology, and plenary discussions at the end of the cross-curricular workshops, cooperation, and collaboration within the class are strengthened, joy (together) is conveyed, and knowledge is deepened in the lesson (cooperation). The fourth pillar "activity" is the most important in the "COOL informatics" concept: within hands-on materials, and playing, tickling, and experimenting with the PolyUni tiles, the students are actively learning basic IT concepts and CT skills, by connecting new knowledge with previous one (Sabitzer, 2013; Sabitzer and Stefan Pasterk, 2013).

Moreover, after the task development, planning, and concepting the three workshops, all self-designed tasks, materials (e.g., task sheets, video), and the processes of the workshops were evaluated again or adapted by the research team if it was necessary. After the evaluation, all materials were made available as an open educational resource (OER) to (Austrian) teachers (JKU COOL Lab Materialbörse, 2022; Table 1).

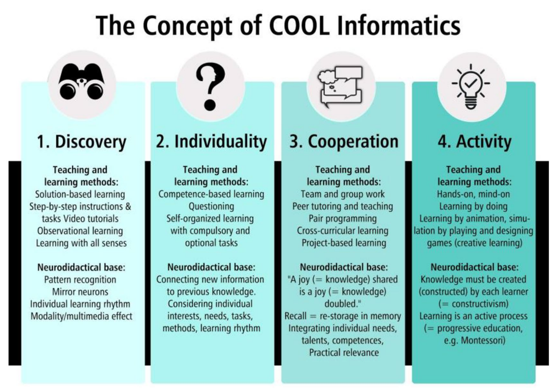


FIGURE 2
The four pillars of the core concept of COOL Informatics (Sabitzer, 2013).

2. Materials and methods

2.1. Research aim

The main goal of this research is to examine creative and innovative possibilities of using the Poly-Universe game in biological, digital, and physical education, to teach simple CS concepts (such as algorithms and coding) and CT skills (such as pattern recognition and abstraction) across disciplines.

For this purpose, three workshops, based on the COOL Informatics concept, were developed. In the first workshop, the Poly-Universe elements were used to map developmental steps of invertebrates in biology lessons (W1). In the second, mathematical tasks were combined with endurance exercises in physical education (W2). Finally, the PolyUni parts were used for dance programming in the third workshop (W3). Therefore, the following research questions are addressed in this article:

1. RQ: How can the Poly-Universe be utilized to promote CT and biological concepts in secondary school biology classes?
2. RQ: How can PolyUni tasks be used to combine physical education with CT in secondary school?
3. RQ: How can the game teach PolyUni CT concepts through dance programming in secondary school?

In addition to the main aim of this work, the following sub-goals were set: Firstly, to investigate how PolyUni can be used to achieve learning goals from the teaching curriculum based on self-designed

TABLE 1 Overview of all workshops, the CT concepts/skills conveyed in W1-3, and how they were guided by the COOL informatics concept (Sabitzer, 2013; Sabitzer and Stefan Pasterk, 2013).

Workshop	CT-Skills and concepts	Main activities of each phase and how they are guided by COOL informatics
Workshop 1 (W1) Poly-Universe in Biological Education	Abstraction, Generalization, Coding, Decomposition, Pattern Recognition, Algorithmic Thinking, Debugging, Experimenting, Problem-Solving	Explanation Phase: Step-by-Step instruction & tasks descriptions/explanations, observational learning (*discovery*) Developing Phase: learning with all senses, learning by doing, tinkering, experimenting with tiles, learning by playing, and designing creative learning, optional tasks, enjoyment, fun (*discovery, activity, cooperation, and individuality*) Debugging Phase: self-organized learning, group work, competence-based learning (*discovery, activity, individuality, and cooperation*) Evaluation and Feedback: team, group work (*cooperation*)
Workshop 2 (W2) Poly-Universe in Physical Education	Coding, Decomposition, Pattern Recognition, Algorithmic Thinking, Debugging, Experimenting, Problem-Solving	Warm-Up Phase: Step-by-Step instruction & tasks, observational learning, sports activity (*discovery and activity*) Developing Phase: learning with all senses, learning by doing, tinkering, experimenting, learning by playing and designing creative learning, optional tasks, competence-based learning, enjoyment (*discovery, activity, individuality, and cooperation*) Evaluation and Feedback team, group work (*cooperation*)
Workshop 3 (W3) Poly-Universe in Digital Education	Abstraction, Automation, Generalization, Coding, Decomposition, Pattern Recognition, Algorithmic Thinking, Debugging, Experimenting, Problem-Solving	Explanation Phase: Step-by-Step instruction & tasks, observational learning Practice Phase: dance activity, learning with all senses, learning by doing learning (*discovery, activity, individuality, and cooperation*) Developing Phase and Performances: tinkering, experimenting, learning by playing and designing creative learning, project-based learning, fun, enjoyment (*discovery, activity, individuality, and cooperation*) Evaluation and Feedback team, group work (*cooperation*)

tasks. Secondly, to evaluate their contribution to the achievement of student learning outcomes, and other possible advantages (e.g., engagement, enjoyment, collaboration).

2.2. Data collection and processing

To examine the above-mentioned questions, a mixed-method approach was embodied. The observation of the participants, based on the methodology of the participating observation (DeWalt et al., 1998), was used for qualitative data collection, and an assessment grid and additional picture analysis were used for the quantitative data. The collection process took place in three workshops between June and December 2022. The data was processed *via* an evaluation of a self-designed computational thinking curriculum assessment grid (CTC-AG).

In regards to the participating observation method (DeWalt et al., 1998), the participants' options regarding the entertainment factor of the game, the misunderstanding of the tasks, the collaboration with peers, and the problem-solving process, were observed and documented by the research team. The participant observation is a method, which has its origins in the field research of De Gérando (2016), was carried out at secondary schools, in a natural environment for the students. Participatory observation was chosen as the data collection method, in order to survey the participating students in natural way whilst they were using the game (in terms of fun, enjoyment of the tasks, behavior within class), and while solving the tasks (understanding the question and rules of the game, correctness of the tasks, errors, debugging). Therefore, this approach collected qualitative data that was very useful in allowing the research team to receive a clearer picture of how the students were using the tiles of the PolyUni game, and additionally, how the participants were interacting with each other. Further, the researchers were openly recognizable to all students and actively participated in the lesson for, e.g., taking pictures during the lectures, explaining the game PolyUni and tasks, and assisting throughout the exercises. The research team took notes during their observation process. In addition, data was documented *via* photographs throughout all workshops, taken by the students themselves in W1 and W3, and the research team in W1-3. At the end of W1, all photos taken during the lectures were sent to the workshop leader for evaluation and analysis.

The CTC-AG was answered manually by all members of the research team for each workshop, partly during and the rest after the lecture, using the notes from the observation (e.g., written recording, notes) and pictures. The questions regarding the game and task design, collaboration within the group, and enjoyment could be answered due to the observation and this documentation process during or shortly after the workshops. The rest of the questions, regarding the accuracy of performing CT tasks, biology practices, the dance, or sports exercises, and further the achievement of the pre-defined teaching and learning goals in all subjects, built on the Austrian secondary school curriculum (BMBWF, 2022a) which is based on bloom taxonomy (Bloom et al., 1956), were processed and checked *via* the CTC-AG after analyzing the pictures and the task sheets (only in W2).

The evaluation criteria of the CTC-AG are based on the Austrian school law (BMBWF, 2022a) and the AHS (German: "Allgemeinbildende höhere Schulen"; General secondary school) curriculum for secondary schools for the subjects of digital literacy, biology, and physical education by the ministry (BMBWF, 2022b). The CTC-AG contains on the one hand of general and demographic questions regarding the number of participants and groups, the location and date of the workshops, age, and gender of the students, and on the other hand of 21 questions that are summarized into three main topics: game and task Design ("Task"), the curriculum biology, physical education or digital Education ("Curriculum"), and "CT" (shown in the Appendix).

The questions can be answered on a scale from "I-V" Likert-5-Scale (Likert, 1932). The degree of correctness or fulfillment of a questionnaire item was rated according to the Austrian school grading system from "I" to "V" (1–5) by the scientific team. "I" stands for "the question has been answered far beyond the essentials," "II" for "beyond the essentials," "III" for "fully met in the essential areas," "IV" for "largely fulfilled in the essential areas" and "V" for "not fulfilled in the essential areas" (BMBWF, 2022a).

2.3. Sampling of the workshops (W1-3)

To analyze and evaluate the utilization of the game PolyUni in secondary school, three different PolyUni-workshops were held between June and December 2022 an Austrian secondary school (BRG Steyr/Upper Austria): "Poly-Universe in Biological Education" (W1), "Poly-Universe in Physical Education" (W2) and "Poly-Universe in Digital Education" (W3). All of the participants had no previous knowledge and did not know the game PolyUni until the day of the workshop.

In sum, 80 11–12 year–old Austrian students participated in this research, 69 identified themselves as female, 19 as male. Each of the three workshops lasted 100 min, and two researchers were always present per lecture. On 15th June, 2022, twenty-nine 11–12-year-old students (female = 24, male = 5) attended the first biology workshop (W1). On 30th June 2022, the second workshop for physical education took place, and twenty-seven students (female = 23, male = 4) participated (W2). Lastly, the third workshop took place on the 2nd December, 2022. In this lesson, twenty-four 11–12-year-old students (female = 14, male = 10) participated in the dance programming workshop for digital education (W3) (Figure 3).

2.4. Task and course design: W1 "poly-universe in biological education"

The first workshop "Poly-Universe in Biology Education" was intended for the subject Biology, but can also be taught interdisciplinary in the subject DGE. In W1, information about invertebrates was conveyed and discussed. Some of the biological topics were already known by the students (such as the morphology of insects) and some were not taught before (such as the morphology of arachnids and the life cycle of insects). The topics were given to the students as worksheets to supplement their textbooks. The different tasks and the worksheet were, based on the Austrian biology AHS curriculum, self-designed, and "Bio@school 2" (Weisl and Schermaier, 2021) was used as the textbook, because "Bio@school 2" is the biology textbook at this Austrian school and all students have this available. The students should use the information to solve the following three tasks, using the PolyUni game: Task one: What is the morphology of

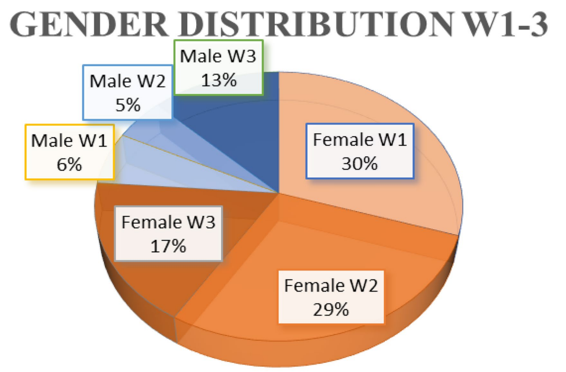


FIGURE 3
Gender Distribution of the participants in all held workshops.

FIGURE 4
Students solving biology tasks with PolyUni files in W1.

a spider compared to an insect? Task two: What is the life cycle of a cockchafer? Task three: What are the differences and similarities between bugs and beetles? The desired teaching and learning objectives for each subject, biology, and digital literacy were recorded based on the Austrian Curriculum (BMBWF, 2022b) in advance so that the teaching unit can then be evaluated using the CTC-AG.

The following learning and teaching goals have been defined: The students recognize, describe, and distinguish the structure and shape of insects and arachnids. They can recognize, name, and describe the differences and similarities between bugs and beetles, and the development cycle of a beetle. The students discover similarities, rules, and patterns in PolyUni. The students use and create codes with the game, to help them name, understand and describe biological processes from the everyday life of an insect and a spider. The students follow the game's clear instructions and carry them out. Furthermore, the students can recognize and correct mistakes in their and other codes themselves.

In addition to these goals, the course of the workshop was also determined in advance, and its duration (100 min) was set as follows: In the first phase (10-20 min), the Erasmus+ project PUSE, the worksheet, the game, and the tasks were presented in the Explanation Phase and the students were divided into five groups. The groups were separated by drawing lots of the different PolyUni game pieces. Each group received a set of the PolyUni tiles, wooden sticks, blank cards, pencils, and a worksheet with the same three tasks. To solve the tasks, the students had to use the game PolyUni, the worksheet, and the textbook. Pens, cards, and wooden sticks were optional. In the second Development Phase (50-60 min) the students solved the tasks with the game and took a photo for documentation afterwards. The first two tasks had to be completed by the students in 100 min, the third task was optional. After the completion of each task, the group members looked at the results of the other groups, to find any mistakes, and analyzed different approaches, color, shape, or size codes used. Thus, they could revise, change, or even contribute to their solutions to start their solving process. After this Debugging Phase (10 min) the students took a second photo. In the final Evaluation and Feedback Phase (10–20 min) the (biological) content was repeated and the students' results were discussed.

The workshop at the secondary school proceeded as planned: Explanation Phase (20 min), Development Phase (60 min), Debugging Phase (5 min), and Evaluation and Feedback Phase (15 min). The only deviation took place within the Debugging Phase, which had been integrated into the Developing Phase. Only one of the 5 groups took a photo after their first attempt to solve the task, then observed the results of the other groups, and afterwards discussed and modified their results, and documented their final approaches within a second picture. All remaining groups of students wandered through the room during the entire processing time of the Development Phase, observed the other solutions, and thus constantly changed their approaches without visually recording the individual steps, but only the final results (Figure 4).

2.5. Task and course design: W2 "poly-universe in physical education"

The second workshop "Poly-Universe in Physical Education" was developed for the subject PE and interdisciplinary for the DGE. The main focus of W2 was on the potential of physical activity to enhance students' cognitive functions by improving their CT skills. In addition, the students should be enabled to expand their motor skills in a variety of ways and to further develop their conditional skills. The lesson was also intended to improve the students' coordination abilities and to raise their awareness of their own movement behavior in terms of movement quality and movement economy. To connect all these predefined teaching goals, an alternating task model was used in which cognitive and physical tasks rotated with each other according to a predefined time window.

For the physical task area, an obstacle course (as shown in Figure 5) with four different movement challenges was chosen. Each challenge contained a physical component (basic movements such as jumping, running, and/or a combination) and a cognitive part (e.g., movement reaction to certain colors). In relation to the cognitive section of the lesson, the task "Tagram" from the script "PUSE Methodology Poly-Universe in school education" by Saxon and Stettner (2019) was used. "Tagram" is about reconstructing given shapes with connections of the same size and color by using the PolyUni triangle set.

The lesson had a total duration of 100 min and was divided into a Warm-up phase (20–25 min), a Development phase

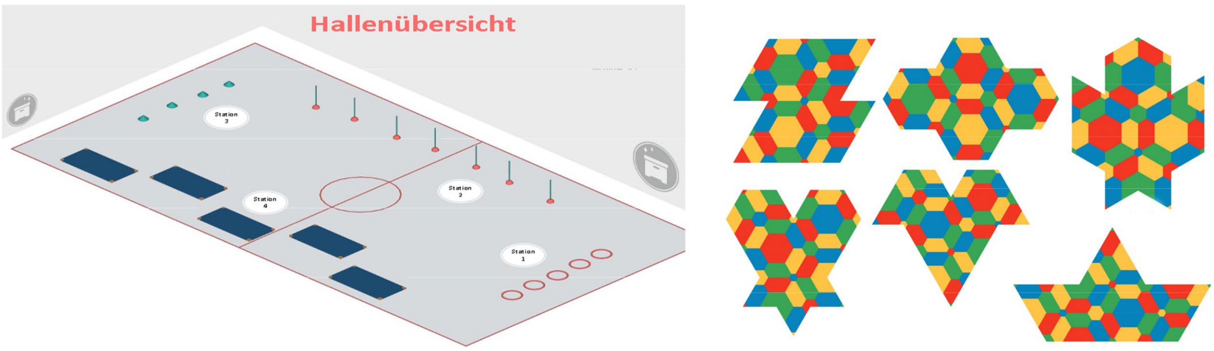


FIGURE 5
Overview (German: "Hallenübersicht") regarding the obstacle course and "Tangram" tasks in W2 (Saxon and Stettner, 2019).

(25-30 min), and a final Evaluation and Feedback phase in the plenum (10-15 min). Transitional phases for the assembly and disassembly of equipment were calculated with a duration of ten to 15 min each.

2.5.1. Course of W2 "poly-universe in physical education"

In the beginning, the students were informed about the goals and contents of the lesson. Then, to generally warm up the students, the movement game "Treasure hunt" was introduced. After the Warm-up Phase was over, the students were gathered into six groups. After group selection, the next section of the lesson "Tangram" was explained (Saxon and Stettner, 2019). Therefore, the previously created teams were given organization cards to introduce the subsequent set-up phase. Together, an obstacle course was set up, the extent of which was oriented along a volleyball court. Within the volleyball court, each team was given a small workspace at a distance of about two meters from the other team's workspaces as well as from the obstacle course. Finally, one triangle set of the game was placed on each of the individual workspaces. Before the game began, the students were asked to sit in pairs in their assigned workspace. The teacher then explained the features and rules of the game "Tangram." The goal was for the students to work together to build a given geometric figure using the PolyUni tiles. The level of difficulty could be adapted to the students' performance and experience.

In addition to the "Tangram"-puzzle (Saxon and Stettner, 2019), team members took turns completing the previously constructed obstacle course. The number of laps to be completed was depending on the performance level of the students (e.g., two or three rounds). Before the start, each team received a small card on which the figures to be created were shown. If a team succeeded in solving a figure, another shape had to be solved. After a predefined period (20-25 min) the game ended with an agreed signal. The winner was the team that was able to solve the most puzzles together. At the end of the game, all materials and equipment were put away by the previously assigned students. Finally, there was a reflective discussion in the plenary, where the students were given the opportunity for feedback and open questions about the lesson. The teacher also referred again to the main points of the lesson mentioned at the beginning and asked the students content-related questions.

2.6. Task and course design: W3 "poly-universe in digital education"

The workshop "Poly-Universe in Digital Education" was about algorithmic dancing using codes, mapped with individual PolyUni elements. Application examples of this workshop content are in addition to the subject DGE, interdisciplinary also the subject PE, or computer science (CS). Within this teaching unit, students learned playfully and dance-wise with algorithms. Students could clearly name and describe instructions for action and execution. Aim of W3 was that all students learn the concept of programming using simple dance movements. After the learning unit, the participants knew how to dance the shown movements and could guide them using the PolyUni elements. At the end of two teaching units, all students mastered the movements of the PolyUni dance programming and their own choreography. Furthermore, they could create, read, and understand PolyUni codes as dance instructions, and could thus apply simple CS concepts such as algorithms.

The two teaching units were divided into the following phases (Figure 6): "Explanatory Phase" (10–20 min), "Practice Phase" (30–40), "Development Phase" (20–30 min), "Performances" (20 min), "Evaluation and Feedback" (5–10 min). In the "Explanatory Phase," the game and its rules, the worksheet with the dance (see Figure 6) and algorithms in general were explained. In the "Practice Phase," the students first practice the individual movements within groups. At the end, the given dances were danced in the plenum as a class. In the "Development Phase," new code and dances were created within six small groups, which were presented to the class in the "Performances" phase. At the end of the workshop, in the "Evaluation and Feedback," students' opinions on the game and regarding the tasks, and their fun-factor are evaluated and written down, using the grid. During W3, the research team took pictures and notes for the final evaluation grid.

2.6.1. Course of W3 "poly-universe in digital education"

The individual phases went as planned. The explanatory phase and the introduction to the topic "algorithms" took longer (25 min), but the students could already dance the codes shown much faster than planned (20 min), and read the dance introduction in the plenum. Furthermore, they were able to gather in small groups in the first hour

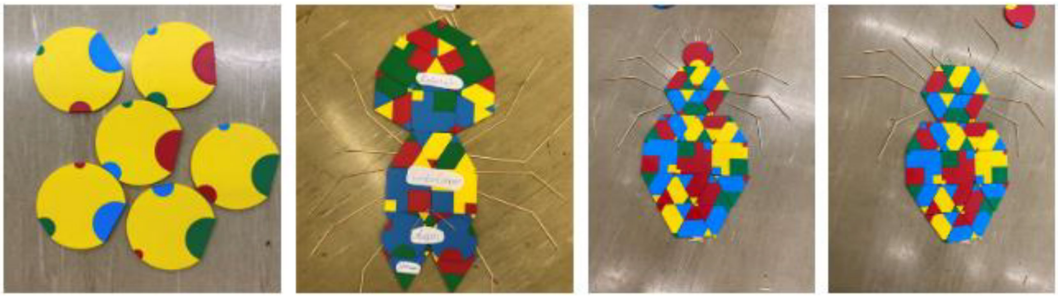


FIGURE 6
Students' solutions to represent the eggs of a cockchafer (left), spider morphology (middle left), and the first (middle right) and final (right) approach of a spider morphology.

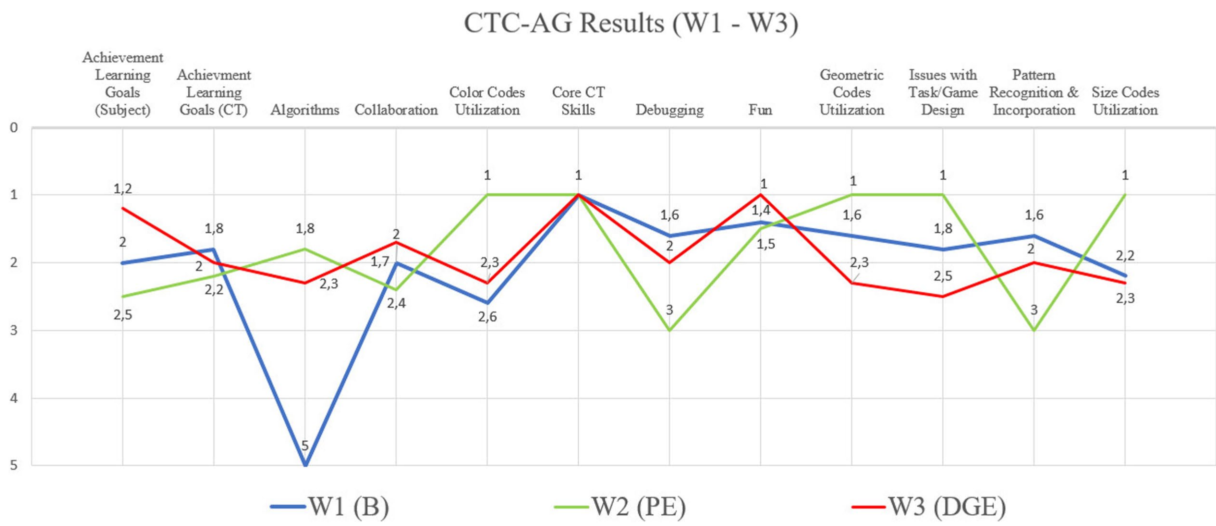


FIGURE 7
Results of the comparison of all participating groups from workshops 1, 2 and 3. The Y-axis shows the grades 1–5, according to the researcher, and the X-axis is showing excerpts of the CTC-AG items.

to develop their own dances and codes with the Poly-Universe parts. The second session ran as prepared in advance.

3. Results

Figure 7 shows the comparison of the observers' opinions of all participating groups from all three workshops (W1–3). The results of the individual groups were evaluated using the Austrian school grading system (1–5) and in the table, the average value (arithmetic mean) of the findings is given. All authors manually filled in the grid and the average value of the results of all answers was taken as the final result. The concordance relating to the comparison of the individual workshops among the research team according to filling out the CTC-AG for each workshop was 89%. These data indicate that the results obtained are valid. A complete table containing all groups and answers according to the research team, is in the Appendix (There, the groups are described as numbers from "1 to 6″ in the corresponding row and column: "1″ stands for group number 1, "2" for group number 2, and so on. Further, a list with all 21 questions (Q1–21) of the CTC-AG is also listed).

In the first category "Task" of the grid regarding the task and game design, all groups achieved an average score (school grade) of 1.7 in W1, 1.1 in W2, and 2.1 in W3. Thus, in the second workshop, there were the fewest issues in terms of task understanding, and explanations of the game or the exercises, especially in W3, shown in Figure 7 ("Issues with task/Game Design"). Overall, there were hardly any problems with task understanding in all W1-2.

Regarding the curriculum, the predefined teaching/learning objectives in the subjects B, Pe, DGE were also achieved above average in all workshops, especially in W1 (1.2). Concerning the achievement of the learning goals regarding CT, all participating groups of students achieved near-top grades.

Looking at the usage of the game PolyUni in relation to "Core CT Skills," according to the research team, the goals are achieved beyond measure in all workshops (1): all students used CT skills (such as problem-solving, pattern recognition, dividing problems into sub-problems, tinkering) in order to solve complex subject-specific

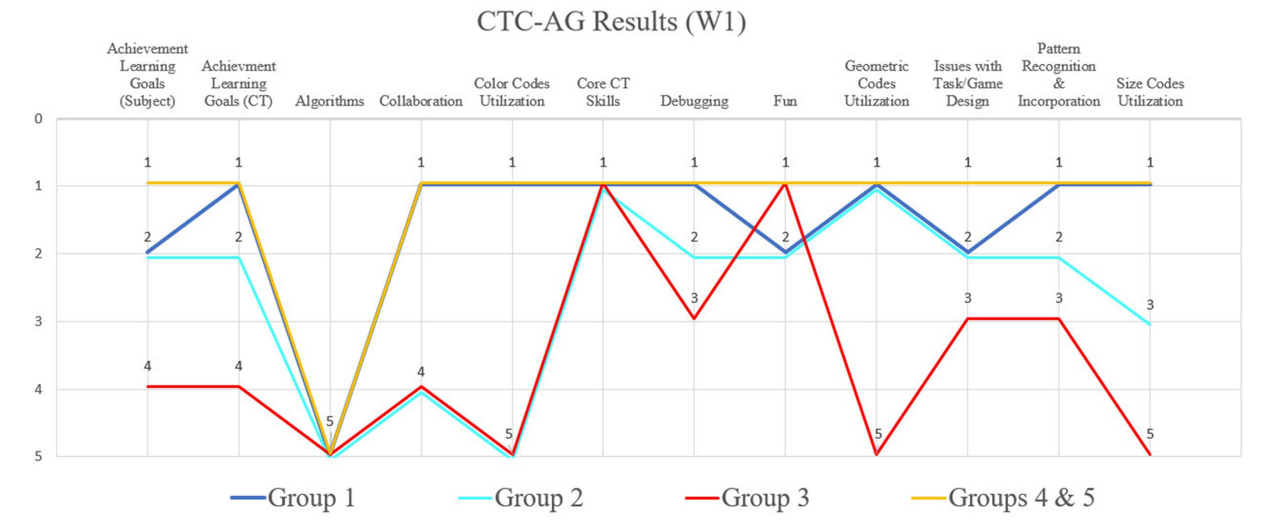


FIGURE 8
Results of the comparison of all five participating groups from workshop 1. The Y-axis shows the grades 1–5, according to the researcher, and the X-axis is showing excerpts of the CTC-AG items.

problems using the tiles (1.0). In addition, all the children were able to present the illustrations of the animals in a simplified manner (abstraction) in order to generally represent the most important characteristics of, for example, insects and spiders (generalization). Furthermore, they had to In addition, strikingly, students did not use algorithms in W1, but very well in W2 (1.8) and W3 (2.3).

In summary, according to the research team, it can be seen that in all workshops the previously defined teaching/learning goals were achieved, and that the students also had a lot of fun in the lessons (W1 = 1.4; W2 = 1.5; W3 = 1), and worked well together in peers to solve the tasks, especially in W2 (1.7). In the following, the results of the individual workshops are described in more detail and illustrated with additional image material.

3.1. Results: W1 "poly-universe in biological education"

Regarding the task and game design in W1, all groups understood the rules of the game and the tasks "fully met in the essential areas." Furthermore, all groups could start solving the tasks with only a little help from the research team. Group 3 needed the most help and groups 4 and 5 the least (Figure 8). It is significant that all five groups enjoyed the use of PolyUni in biology lessons "far beyond the essentials" or "beyond the essentials." This result indicates that the initial problems with the task and the explanation of the game had no impact on the fun in the workshop, and, furthermore, PolyUni creates enjoyment in biology class. Concerning the results of tasks in reference to the Austrian biology curriculum, the findings show that groups 4 and 5 could solve the tasks perfectly. Figure 9 shows group 4's solution to task 2, presenting the life cycle of a cockchafer.

Using the tiles of the game, the students were able to represent the different stages (egg, larva (younger and older), pupa, adult animal) in a biologically correct manner and label them correctly using the cards. Groups 4 and 5 had no problems in understanding and presenting previous and new learning content. Two groups (1, 2) were partly able to understand and present previously learned and new biological content with the game. It is evident that group 3 struggled the most in biological correctly solving the tasks. After comparing the solutions of the individual groups, it turned out that 4 out of 5 groups had achieved the specified teaching and learning goals in biology, but group 3 only "largely fulfilled" the tasks "in the essential areas." In contrast, groups 1, 2, 4, and 5 achieved the teaching and learning goals set in advance "far beyond the essentials" or at least "beyond the essentials."

These results indicate that PolyUni can be successfully used to reach the required teaching and learning goals in biology education. Regarding CT, all students were tinkering, playing, and experimenting with different tiles of PolyUni to solve the tasks. All groups could break down a complex biological problem into sub-problems and thus solve the required tasks. Furthermore, all students used PolyUni to identify similarities, differences, and patterns to solve complex biological problems. In addition, all five groups could hide unimportant details so that the participating students could focus on the essential aspects of the problem and thus solve the task.

After photo analysis, the findings show that groups 1, 4, and 5 used color, size, and (geometric) shape combinations as codes to correctly display the biological content and to solve the different tasks. In Figure 6 groups 1 and 5's solutions are presented. Group 1 used the same sizes, the same-sized semicircles pointing in the same direction, same shapes (circles) and color (yellow) to represent the eggs of a cockchafer. Same size, color, and shape codes are also used in group 5's solution. Two green triangles of equal size are used for the spider's jaw claws and circles for the eyes. As seen on the prosoma two of the same size semi-squares are connected vertically with each other. Group 2 used shape combinations as codes "far beyond the essentials" and the size combination utilization was "fully met in the essential areas." Only group number 3 hardly used geometric shape codes, and furthermore, no color and size codes. Concerning the debugging, at least all groups

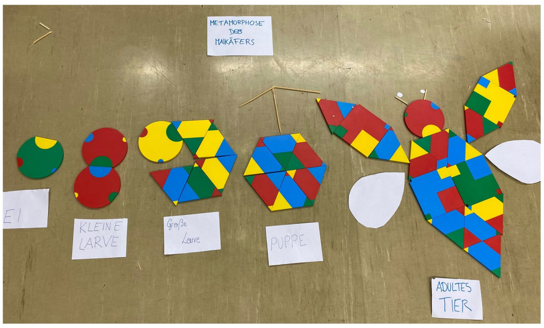


FIGURE 9
Solution of task 2, representing the life cycle of a cockchafer.

"fully met in the essential areas." From the observer's point of view, all students found mistakes and improved them on their own, recognized patterns and codes from the other groups, and further incorporated them into their own codes to improve their solutions, especially in groups 1, 4, and 5. As seen in Figure 6, group 4's first approach was that Araneae have a caput, a thorax, and an abdomen, i.e., a body structure like an insect. After comparison with the other groups, their first approach was improved, and the final solution showed the prosoma and the opisthosoma.

After analyzing the photos, the findings indicate that with the utilization of the game the previously defined teaching and learning goals for CT were fulfilled "far beyond the essentials" for groups 1, 4 and 5, "beyond the essentials" for groups 2 and 3 only "largely fulfilled" the tasks "in the essential areas." These results regarding the learning and teaching goals in CT and biology appear verified after comparing the intensity of collaboration within the different groups: The students of groups 1, 4 and 5 collaborated in peers "far beyond the essentials" to solve all problems and tasks together, but groups 2 and 3 only "fully met in the essential areas" in this category. In summary, with the results concerning CT, it can be assumed that the game PolyUni can be successfully used in biology classes to promote CT in digital literacy in secondary school.

3.2. Results: W2 "poly-universe in physical education"

Results related to the game and task design suggest that all groups understood both the task and the function of the game after the teacher explained it. The same results could be found regarding task comprehension. Furthermore, it was shown that all six groups were able to initiate the game without additional explanatory aids after the teacher's introduction. It also became apparent that the game was particularly enjoyable "far beyond the essentials" for most of the participant groups (1, 3, 4, and 6). As for the results concerning the Austrian physical education curriculum, the data should be interpreted with caution. Since PolyUni was embedded in the lesson structure only as a cognitive subsection, student experiences with similar tasks were included in the analysis. With this information in mind, the data provide preliminary evidence to suggest that with the help of PolyUni, most of the students may appropriately apply and convey prior information regarding previous contents of PE lessons. Considering that the task "Tangram" (Saxon and Stettner, 2019) was already the content of the physical education of all group participants in its original variation, it was unsurprising that the game allowed the students to accurately define and explain prior knowledge. Minor difficulties were encountered in the transition from the original game idea "Tangram" to the variant with the game's elements. Only students from groups 1 and 3 were fully capable of accurately comprehending and presenting. and explaining new material, whereas the remaining groups needed additional explanations in the introductory part of the lessons (Figure 10).

A similar picture emerged during the learning goal orientation, whereby group 2 was the only group that could not fulfill the predefined teaching goals from the observers' point of view. At the same time, results regarding CT yielded some interesting findings. First, the data showed that all students were able "far beyond the essentials" to complete the objective by exploring, playing, and/or tinkering. Additionally, every group seemed to be able to divide a difficult problem into smaller ones and complete the assignment. To address challenging issues, all students also seemed to utilize PolyUni to find patterns, similarities, and contrasts. Moreover, all students were able to find patterns, similarities, and contrasts to tackle complex problems. Differences between the individual groups only became clear in the way the tasks were processed. It was found that groups 1 and 3, and to a lesser extent group 4, were particularly good at blocking out unnecessary details and avoiding attentional distractions to solve the task. Group 5 and 6, on the other hand, managed the task with some difficulty regarding task focus, whereas group 2 frequently excelled due to distractions from various group members. Furthermore, possibly due to the basic idea of the game "Tangram," the data showed that all groups were able to adequately use geometric figures as well as color and size combinations to accomplish tasks.

However, major differences were revealed in the detection and processing of faulty work steps. Whereas groups 1 and 3 identified and corrected errors independently, the remaining groups sought help from the teacher to varying degrees to solve the task. A similar trend was again reflected in the ability to use algorithms in a task-specific manner. This showed that groups 1 and 3 stood out particularly positively, while group 2 required assistance more frequently. However, commonalities emerged in the collaboration between individual group members of each team, as a high proportion of teamwork was observed in all groups. Finally, it became apparent that by using PolyUni, the teacher was able to complete the teaching goals and learning objectives for CT and digital literacy to a full extent in groups 1 and 3. In groups 4, 5 and 6, the task only partially met the predefined learning objectives and for group 2, the task was found to be unsuitable from the observers' point of view (Figure 11).

3.3. Results: W3 "poly-universe in digital education"

In W3, all students of the six groups at least "fully met the essential areas" by understanding the tasks and the rules of the game. All groups only needed a little support by the research team but had a lot of fun with the dance programming (Figure 12).

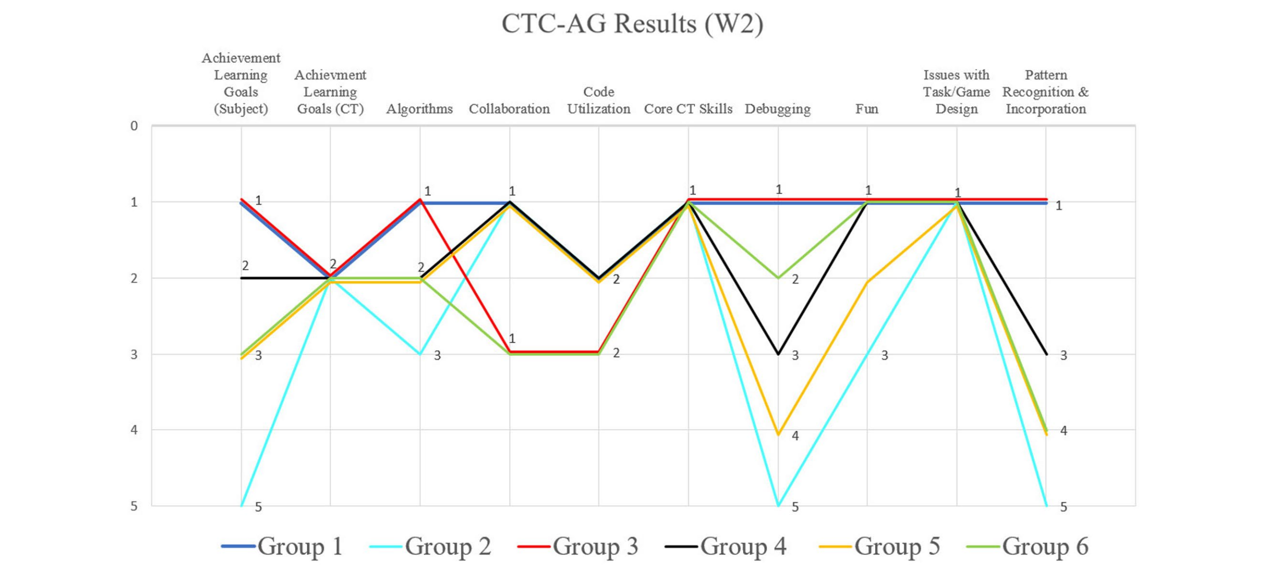


FIGURE 10
Results of the comparison of all six participating groups from workshop 2. The Y-axis shows the grades 1–5, according to the researcher, and the X-axis is showing excerpts of the CTC-AG items.

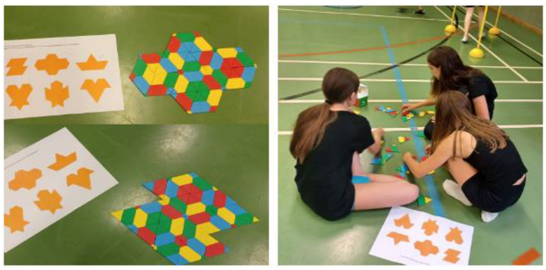


FIGURE 11
Solutions of "Tangram" (Saxon and Stettner, 2019) tasks in W2.

With regard to the "Digital Education" curriculum, all students also achieved at least "beyond the essentials" the previously defined teaching objectives. Neither in the category "Task" nor "DGE Curriculum" did one group stand out particularly prominently in one direction or the other. All participating children of W3 were on a very similar level and there were hardly any differences. This is also evident in the "CT" category. All groups played and experimented with the PolyUni parts and created new codes for individual dance moves. Particularly many and different codes for dance movements (between 5 and 13 different codes; group1: $f = 13$) were created by groups 1, 2, 4 and 5, but over all, all participants developed new movement sequences, as shown in Figure 13.

After the photo analysis, it is striking that all groups found mistakes on their own (de-bugging) and could improve their dance algorithms on their own, without any additional help from the research team. Furthermore, all participating students could recognize patterns and (dance) codes with the PolyUni tiles from other groups, and incorporate them into their own codes and representations to improve their final approach. Therefore, all predefined objectives were achieved at least "beyond the essentials" with regard to CT.

4. Discussion

Even if at first glance the connection between CT and such diverse subjects, especially B and PE, seems impossible, previous studies show that this is not (Hammrich et al., 2021; Galoyan et al., 2022). However, in order to be able to make general statements more scientific research must be done in this field in the future. The game Poly-Universe clearly shows positive effects on the participating students throughout several studies, especially relating visual perception and attitude change toward STEAM-related tasks. It can therefore be assumed that Poly-Universe can also be usefully applied to other subjects and school levels (Dardai, 2018; Hoffmann, 2020). Based on these conclusions and preliminary results, new and adapted teaching materials with the game were invented for the three Austrian subjects, B, PE, DGE, in order to not only teach the participating secondary school students the required teaching material in a creative, innovative, fun, and game-based way, but also to promote CT.

In further considerations and conclusions it must be included, that the majority of the 80 students were girls, and there is a risk in using the participating observation method that the objectivity can suffer due to the complete participation (DeWalt et al., 1998). Previous studies showed similar findings, that game-based learning promotes enjoyment in class and it can be assumed that it positively affects students' collaboration (Vandercruysse et al., 2012; Hamari and Koivisto, 2015; Crocco et al., 2016). However, it must be mentioned that the three classes were deliberately chosen for these three workshops (W1-3). On the one hand, because three of the authors actively teach as in-service biology, physical education, and digital education teachers in these classes. Further, the BRG Steyr is a partner institution of the research team. To reduce researchers' bias in reporting and analyzing the data, participants were additionally observed by authors who did not create the tasks themselves and were not actively teaching in those classes. The observations and written records were then evaluated by all authors. On the other hand, because

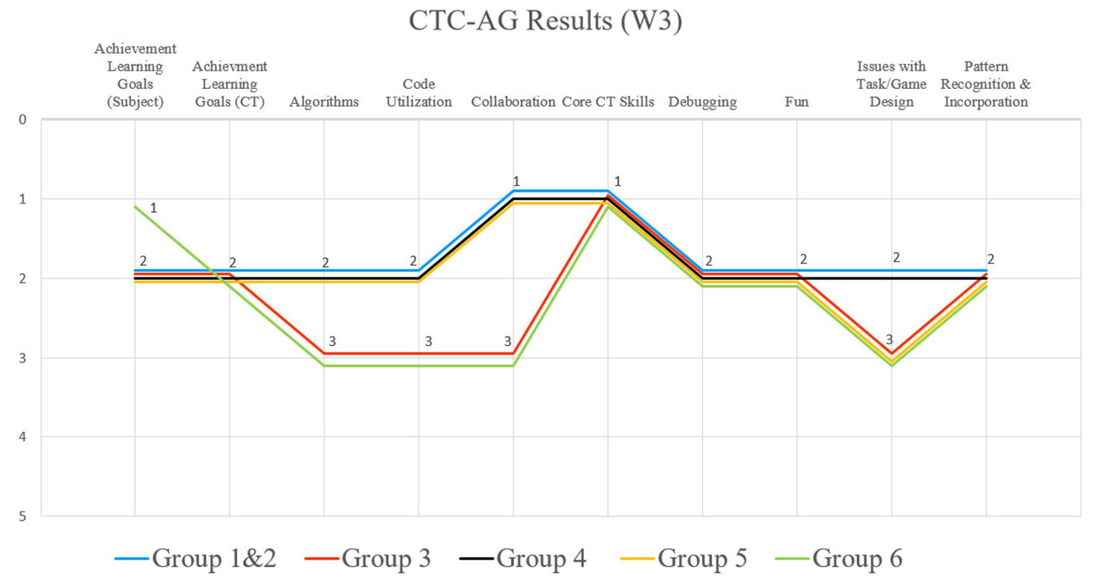


FIGURE 12
Results of the comparison of all six participating groups from workshop 3. The Y-axis shows the grades 1–5, according to the researcher, and the X-axis is showing excerpts of the CTC-AG items.

FIGURE 13
Students in W3 are dancing their own dance programming (left) and creating new codes for every movement using the PolyUni tiles (right).

the majority of the classes studied are girls, the research team was also keen to promote interest in CT among young people, especially girls, from an early age. In addition, it should also be noted that three workshops with 80 children has no general significance. However, a positive trend towards the use of PolyUni as an educational game to promote CT in secondary school can already be seen in this study, but has to be further researched in the future.

This study surveyed which CT concepts were positively influenced and promoted in regards to PolyUni. In summary, the following concepts and skills were be promoted using the PolyUni game within all secondary school workshops: the students had to simplify and abstract the illustration of animals (W1: e.g., an eye of a spider is represented by a square shaped PolyUni tile), geometric shapes (W2: represent geometric figures in squares, triangles and circles), and dance programs (W3: one color represents one movement) (Abstraction), recognize similarities and differences between insects and spiders (W1: e.g., number of legs; insects have three-part bodies, spiders have two-part bodies), geometric tasks (W2: e.g., number of squares), and dance moves (W3) (Pattern recognition), and had to break down complex tasks into subtasks and sub-problems (W1-3: e.g., dance program consists of different parts, repetitions) (decomposition). Regarding solutions on how to generalize animals and their characteristics (W1), a geometric shape (W2), or a dance program (W3) (Generalization), the participants had to use a wide variety of color, size, or geometric form combinations and algorithms that stood for certain movements (W3: e.g., red color stands for raising your hand up), shapes (W2: only using triangle-shaped tiles), or animal characteristics (W1: square-shaped tiles are used for legs, yellow circle-shaped PolyUni elements are used for eggs; W1-3 create an algorithm for a specific shape or dance movement) (Coding, Algorithmic Thinking). The codes of the dance steps (W3) can generally be recorded and automatically reproduced by programs (automation). All of the participants had to experiment and play with the PolyUni elements (Experimentation & Tinkering), and were able to change their approaches at any time through the solutions of others, recognize mistakes and patterns of others and thus optimize their final solutions (Debugging) in all workshops (W1-3: group discussion, group work).

The majority of the children were able to present the illustrations of their tasks (e.g., animals, dance codes) in a simplified manner (Abstraction) in order to represent the most important characteristics of, for example, insects and spiders (Generalization). Furthermore, the results concerning the breaking down of complex problems also indicate that the tasks were selected to be child- and age-appropriate and that the individual elements of PolyUni obviously invite students to try them out, and enjoyably play with their peers.

It can be assumed that a great collaboration and enjoyment during a lesson can very well have a positive effect on a student's learning success whilst using PolyUni in an educational context. Furthermore, findings suggest that PolyUni can be used successfully outside of the math curriculum, as seen in the study by Hoffmann (2020), as well in this case in biology, digital, and physical education classes. Above all, dance programming is an innovative alternative to teach and learn CT and basic IT concepts (e.g., algorithms) in a creative way. In terms of fun, there were some differences in the second workshop (PE): a possible explanation for a slightly different result in the second workshop is that the students could not see the approaches and results of the other groups very well while working out. From this finding, it can be concluded that if the focus in a teaching unit is on finding errors (de-bugging) or code (pattern) recognition in other groups, no strenuous movement exercises, such as running, should take place at the same time.

Finally, this study also surveyed how the educational game affects the learning success in the respective subjects. Overall, from the point of view of the observers, most of the students were able to achieve the previously defined teaching and learning goals for CT and digital literacy with PolyUni, especially in the biological and digital education workshops. The results indicate that also the strenuous endurance run during the exercise probably had a negative impact on the CT results (pattern recognition and incorporation). On this point, further studies with more moderate sports exercises or shorter running phases must be carried out in the future to confirm the assumptions.

5. Conclusion

In this research, three workshops were held at an Austrian school by implementing Poly-Universe into the courses to examine whether the game can be used successfully in biological, digital, and physical education in secondary school, to teach the required curricula, and further, promote CT at the same time. To explore these assumptions, the participating students were observed, the photos of the presented results and during W1-3 were analyzed for their correctness, and the results were recorded in a self-designed evaluation grid, which was then evaluated. Regarding the positive influence of fun on learning outcomes, the results vary between the three workshops in this study. In W2, findings indicate that the participating students who had less fun with the exercise than the rest of the students showed poorer results in the previously defined teaching and learning goals and had more issues correctly understanding, presenting, describing, and explaining new content. Further findings indicate that most groups in W1-3 had no problems understanding the game itself and were able to start the exercises without much additional help or further explanation. The results show that the tasks and explanations were designed to be age- and student-appropriate and that the students also understood the basics of PolyUni.

The majority of participants met their teaching and learning objectives. We examined not only the predefined teaching and learning goals of the curricula, but also other positive effects on students: During the lectures, most of the participants of all three workshops collaborated well with their peers. Furthermore, they enjoyed using the game during the lectures, especially whilst creating new codes for their dance programming.

All groups were able to use CT skills, such as abstraction, generalization, problem-solving, and the ability to break complex problems into subproblems, as well as using the files to tinker, play, or experiment with the PolyUni game. Additionally, all students were able to identify similarities, differences, and recognize patterns to solve complex problems. Most of the participating students used codes of geometric shapes, sizes, and color combinations, to correctly present new or previous learning biological content or to solve the physical education tasks. It is also striking that, especially in the first workshop, most of the students recognized their own mistakes in their approaches and codes and were able to solve them well by comparing them with others. Overall, from the point of view of the observers, most of the students were able to achieve the previously defined teaching and learning goals for CT PolyUni, especially in the first and last workshops.

Therefore, it can be assumed that the game is a great educational tool in various subjects in the secondary school promoting CT skills. The original learning benefit for which this game was designed can also be extended to teach the required curriculum in DGE and promote CT in B, DGE, and PE. In addition to the examples given in the article, in the "PUSE Methodology" material collection for primary and secondary school students, there are numerous teaching examples for various STEAM subjects available online as OER. These range from simple mathematical tasks (e.g., recreating given geometric shapes) up to more advanced exercises (e.g., Biology: characteristics of flowering plants).

6. Outlook

The participating students in this study showed increased collaboration and enjoyment during the workshops, therefore, further studies with adapted B, PE, DGE, and dance exercises in Austrian elementary and primary schools will take place in winter and summer 2023. Additional qualitative research will also take place in 2023 and 2024. In addition, further workshops for secondary education, not only for B, PE, and DGE, but also for other subjects, are planned on this promising topic in 2023.

Data availability statement

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

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/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

ES, MScha, and MSchm contributed to conception, evaluation and design of the study, further organized the database and performed the statistical analysis and wrote the first draft of the manuscript. BA, BS, and ZL contributed in the conception. ES, MScha, MSchm, and

SH wrote sections of the manuscript. CH and MR contributed especially in revision. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the project Poly-Universe in Teacher Training Education, ID: KA203-DE88CF6B.

Acknowledgments

The workshops and materials were created in cooperation with the Erasmus+ project PUNTE.

Conflict of interest

Author SH was employed by Dynatrace.

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

Publisher's note

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

References

Arastoopour Irgens, G., Dabholkar, S., and Bain, C. (2020). Modeling and measuring high school students' computational thinking practices in science. *J. Sci. Educ. Technol.* 29, 137–161. doi: 10.1007/s10956-020-09811-1

Barr, V., and Stephenson, C. (2011). Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? *ACM Inroads.* 2, 48–54. doi: 10.1145/1929887.1929905

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (1956). "Taxonomy of educational objectives: the classification of educational goals" in *Vol. Handbook I: Cognitive Domain* (New York: David McKay Company)

BMBWF. (2022c). Digital Education (Digitale Grundbildung). Available at: https://www.bmbwf.gv.at/Themen/schule/zrp/dibi/dgb.html (Accessed December 6, 2022).

BMBWF. (2022a). RIS-LBVO. Available at: https://www.ris.bka.gv.at/geltendefassung.wxe?abfrage=bundesnormen&gesetzesnummer=10009375&fassungvom=2020-02-01&ShowPrintPreview=True (Accessed December 6, 2022).

BMBWF. (2022b). RIS-Lehrpläne AHS. Available at: https://ris.bka.gv.at/geltendefassung.wxe?abfrage=bundesnormen%20&gesetzesnummer=10008568&ShowPrintPreview=True (Accessed December 6, 2022).

Crocco, F., Offenholley, K., and Hernandez, C. (2016). A proof-of-concept study of game-based learning in higher education. *Simul. Gaming* 47, 403–422. doi: 10.1177/1046878116632484

Curzon, P., and McOwan, P. W. (2018). *Computational Thinking: Die Welt des Algorithmischen Denkens in Spielen, Zaubertricks und Rätseln*. Heidelberg, Berlin: Springer.

Dardai, Z. (2018). Poly-Universe in School Education. Erasmus+ PUSE Project. Available at: http://www.poly-universe.com/wp-content/uploads/PUSE_STUDY_EN.pdf (Accessed December 6, 2022).

De Gérando, J. M. (2016). *Considérations sur les Diverses Méthodes à Suivre Dans L'observation des Peuples Sauvages* (*Ed. 1800*). Paris/France: Hachette Livre BNF

De Vasconcelos Martins, I., Bidarra, G., and Rebelo, B. V. (2022). Looking for the good shape in the poly-universe material. *Int. J. Dev. Educ. Psychol. Rev. INFAD Psicol.* 1, 341–350. doi: 10.17060/ijodaep.2022.n1.v1.2390

DeWalt, K. M., DeWalt, B. R., and Wayland, C. B. (1998). "Participant observation" in *Handbook of Methods in Cultural Anthropology*. ed. H. R. Bernard (Walnut Creek/California: Alta Mira Press), 259–299.

Fritz, C. D. (2022). Integrating Computing through Physical Activity in Elementary School Classrooms (Doctoral Dissertation). University of Texas at Austin. doi: 10.26153/tsw/44017

Galoyan, T., Barany, A., Donaldson, J. P., Ward, N., and Hammrich, P. (2022). Connecting science, design thinking, and computational thinking through sports. *Int. J. Instr.* 15, 601–618. doi: 10.29333/iji.2022.15134a

GeoGebra. (2022). Available at: https://www.geogebra.org/ (Accessed December 6, 2022).

Goldberg, D. S., Grunwald, D., Lewis, C., Feld, J. A., and Hug, S. (2012). "Engaging computer science in traditional education: the ECSITE project" in *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE' 12)*, vol. *17* (New York, NY, USA: Association for Computing Machinery), 351–356.

Hamari, J., and Koivisto, J. (2015). Why do people use gamification services? *Int. J. Inf. Manag.* 35, 419–431. doi: 10.1016/j.ijinfomgt.2015.04.006

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B. D., Asbell-Clarke, J., and Edwards, T. (2016). Challenging games help students learn: an empirical study on engagement, flow and immersion in game-based learning. *Comput. Hum. Behav.* 54, 170–179. doi: 10.1016/j.chb.2015.07.045

Hammrich, P., Richardson, G., Green, T., and Livingston, B., B. (2021). Sisters in Science: Using Sports as a Vehicle for Science Learning. Available at: https://eric.ed.gov/?id=ED452094 (Accessed December 6, 2022).

Hoffmann, M. (2020). Discovering information visualization through poly-universe. *Symmetry Cult. Sci.* 31, 15–22. doi: 10.26830/symmetry_2020_1_015

Hong, Q. (2009). "Teaching computational thinking through bioinformatics to biology students" in *Proceedings of the 40th ACM Technical Symposium on Computer Science Education (SIGCSE '09)* (New York, NY, USA: Association for Computing Machinery), 188–191.

Hooshyar, D., Pedaste, M., Yang, Y., Malva, L., Hwang, G.-J., Wang, M., et al. (2021). From gaming to computational thinking: an adaptive educational computer game-based learning approach. *J. Educ. Comput. Res.* 59, 383–409. doi: 10.1177/0735633120965919

Hutchins, N. M., Biswas, G., Zhang, N., Snyder, C., Lédeczi, Á., and Maróti, M. (2020). Domain-specific modeling languages in computer-based learning environments: a systematic approach to support science learning through computational modeling. *Int. J. Artif. Intell. Educ.* 30, 537–580. doi: 10.1007/s40593-020-00209-z

Israel-Fishelson, R., Hershkovitz, A., Eguíluz, A., Garaizar, P., and Guenaga, M. (2021). The associations between computational thinking and creativity: the role of personal characteristics. *J. Educ. Comput. Res.* 58, 1415–1447. doi: 10.1177/0735633120940954

JKU COOL Lab Materialbörse. (2022). Available at: https://www.cool-lab.net/materialboerse/ (Accessed December 6, 2022).

Kennedy, T. J., and Sundberg, C. W. (2020). "21st century skills" in *Science Education in Theory and Practice. Springer Texts in Education*. eds. B. Akpan and T. J. Kennedy (Cham: Springer)

Likert, R. A. (1932). A Technique for the Measurement of Attitudes. *Arch Psychol.* 22, 5–55.

Naulakha, S., and Bar-Joseph, Z. (2011). Algorithms in nature: the convergence of systems biology and computational thinking. *Mol. Syst. Biol.* 7:546. doi: 10.1038/msb.2011.78

Piaget, J., (1971). *Psychology and Epistemology: Towards a Theory of Knowledge*. New York: Grossman.

Qian, M., and Clark, K. R. (2016). Game-based learning and 21st century skills: a review of recent research. *Comput. Hum. Behav.* 63, 50–58. doi: 10.1016/j.chb.2016.05.023

Sabitzer, B. (2013). A Neurodidactial Approach to Cross-Curricular Open Learning. Cool Informatics. Habilitation. Alpen-Adria-University Klagenfurt.

Sabitzer, B., and Stefan Pasterk, S. (2013). Informatics is cool. Cross-curricular concepts for computer-supported open learning in secondary schools. *Proce. ICERI Conf.* 6, 2739–2749.

Saxon, J. S. (2018). Poly-universe in school education. PUSE. Available at: http://www.poly-universe.com/ (Accessed December 6, 2022).

Saxon, J. S., and Stettner, E. (2019). *PUSE (Poly-Universe in School Education) METHODOLOGY–Visual Experience Based Mathematics Education*. Szokolya: Poly-Universe Ltd.

Selby, C., and Woollard, J. (2013). Computational Thinking: The Developing Definition. University of Southampton (E-prints). 1–6. Available at: https://eprints.soton.ac.uk/356481/ (Accessed December 6, 2022).

Settle, A., Franke, B., Hansen, R., Spaltro, F., Jurisson, C., Rennert-May, C., et al. (2012). Infusing Computational Thinking Into the Middle- and High-School Curriculum. Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education. (ITiCSE '12). Association for Computing Machinery, 17, 22–27.

Vandercruysse, S., Vandewaetere, M., and Clarebout, G. (2012). "Game-based learning: a review on the effectiveness of educational games" in *Handbook of Research on Serious Games as Educational, Business and Research ToolsChapter: 32* (IGI GlobalEditors: Maria Manuela Cruz-Cunha), 628–647.

Weisl, H., and Schermaier, A. (2021). bio@school 2 NEU. Schulbuch. Veritas-Verlag. 2. Auflage. Available at: https://www.veritas.at/bioatschool-2-neu-schulbuch-978-3-7101-3198-1

Appendix

CTC-AG curriculum assessment grid (BMBWF, 2022a,b).

Group number:––––––––––––––––– Name of class: ––––––– Date: ––––––––
Number of participants: ––––––––– Age:–––––––––– Location:–––––––––
Female:––––––––––––– Male:–––––––––– Name of the school: ––––––––––

	Please have the teacher tick where applicable during and after the lesson	1 Far beyond the essentials	2 Beyond the essentials	3 Fully met in the essential areas	4 Largely fulfilled in the essential areas	5 Not fulfilled in the essential areas/not rateable
	Task					
1	After the teacher has explained the game Poly-Universe, the students have understood the content of the game and its process and function					
2	After the teacher has explained the tasks, the students have understood the content of the worksheets and tasks with Poly-Universe					
3	After the teacher explained the tasks, the students could start with the Poly-Universe work tasks without additional help from the teacher					
4	The students enjoy using Poly-Universe in the lecture					
	Curriculum Biology/Physical Education/Digital Education					
5	The students can use and present previous knowledge correctly with the game Poly-Universe					
6	The students can correctly describe and explain previous knowledge with the game Poly-Universe					
7	The students can correctly understand and present new learning/teaching content with the game Poly-Universe					
8	The students can correctly describe and explain new learning/teaching content with the game Poly-Universe					
9	With the Poly-Universe game, the teacher has achieved the teaching/learning goals set in advance					
	Computational thinking					
10	The students tinker, play and/or experiment to solve the problem and task					
11	The students can break down a complex biological problem into sub-problems and thus solve the tasks					
12	Students can use Poly-Universe to identify similarities, differences, and patterns to solve complex subject-specific problems					
13	The students can hide unimportant details so that they can focus on the essential aspects of the problem and thus solve the task					
14	The students use new/identical geometric figures, shapes and combinations as codes in order to correctly present (subject-specific) content, to solve the task					
15	The students use new and the same color combinations as codes to correctly display content and subject-specific content and to solve the task					
16	The students use new and identical size combinations as codes in order to correctly display content and subject-specific content and to solve the task					
17	The students find mistakes on their own and can improve them on their own					
18	Students can recognize patterns and codes from other groups and incorporate them into their own codes and representations to improve them					
19	The students use algorithms and algorithmic thinking to correctly present content and subject-specific content and to solve the task					
20	The students work together to solve the problem and the task together					
21	With Poly-Universe, the teacher has achieved the previously defined teaching/learning goals for computational thinking and digital education					

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 April 2023
DOI 10.3389/fpsyg.2023.1131765

OPEN ACCESS

EDITED BY
Heng Luo,
Central China Normal University,
China

REVIEWED BY
Satyanarayana Parayitam,
University of Massachusetts Dartmouth,
United States
Yunpeng Ma,
Zhoukou Normal University,
China
Tianjiao Chen,
Central China Normal University,
China

*CORRESPONDENCE
Shu Ching Yang
✉ shyang@mail.nsysu.edu.tw
Seokmin Kang
✉ seokmin.kang@utrgv.edu

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

RECEIVED 26 December 2022
ACCEPTED 13 March 2023
PUBLISHED 12 April 2023

CITATION
Luo YF, Kang S, Yang SC and Lu CM (2023) The relationships among Taiwanese youth's polychronicity, multitasking behavior and perceived learning performance in online learning.
Front. Psychol. 14:1131765.
doi: 10.3389/fpsyg.2023.1131765

The relationships among Taiwanese youth's polychronicity, multitasking behavior and perceived learning performance in online learning

Yi Fang Luo[1], Seokmin Kang[2]*, Shu Ching Yang[3]* and Chia Mei Lu[4]

[1]Graduate School of Human Sexuality, Shu-Te University, Kaohsiung, Kaohsiung County, Taiwan, [2]Department of Teaching and Learning, College of Education & P-16 Integration, The University of Texas Rio Grande Valley, Edinburg, TX, United States, [3]Institute of Education, National Sun Yat-sen University, Kaohsiung, Taiwan, [4]Department of Digital Multimedia Design, Cheng Shiu University, Kaohsiung, Taiwan

Background: The advancement of digital technology implies the importance of polychronic learning. Since polychronicity is not equivalent to multitasking behavior, they need to be considered separately. However, less research has been explored on how polychronicity is related to multitasking behavior in the educational field.

Objective: To explore the relationships among polychronicity, multitasking behavior and learning performance (including knowledge acquisition and learning satisfaction) in an online learning environment.

Methods: The relationship among variables was analyzed from 865 responses obtained from a questionnaire survey, and independent sample *t* tests and SEM analysis were used to examine the research hypotheses.

Results: College students showed a higher frequency of multitasking behavior, time tangibility and scheduling preference, and learning satisfaction in multitasking online learning environments than high school students. Additionally, college students were different from high school students on the paths of involvement with people to multitasking behavior ($\Delta\chi^2$=5.42, p=0.02) and scheduling preference to learning satisfaction ($\Delta\chi^2$=9.54, p=0.002).

Conclusion: The relationship among polychronicity, multitasking behavior and perceived learning performance in an online learning environment varies by student educational stage.

KEYWORDS

polychronicity, multitasking, learning performance, online learning, knowledge acquisition, learning satisfaction

1. Introduction

COVID-19 has forced educational institutions to promptly respond to this emergency and take a big step forward from traditional face-to-face learning to online learning to ensure that education continues (Gumede and Badriparsad, 2022). Previous studies have pointed out the advantages of online learning, including time and cost effectiveness (Elida et al., 2012;

Almahasees et al., 2021), flexible use of time and space (Lemay et al., 2021; Maqableh and Alia, 2021; Turan et al., 2022), and the repetitive learning features of online resources (Bdair, 2021). These online resources are considered especially important for students' self-directed learning, but research also points out that students' online self-directed learning is challenging because students report that they are easily distracted and have limited attention (Lemay et al., 2021; Maqableh and Alia, 2021), which in turn affects their academic performance (Kim et al., 2022).

Distraction and inattention may be exacerbated by learners' multitasking behavior (MB; Cebollero-Salinas et al., 2022). However, time-limited multitasking seems to be different from the current online learning ethos that emphasizes flexibility and self-direction. The recent education reform in Taiwan requires the cultivation of interdisciplinary, independent, and autonomous talent, prompting self-directed online learning to become the focus of education in Taiwan and emphasizing students' personalized learning processes rather than learning within a given time (Chen and Li, 2021). Given the lack of research on MB in online learning, one of the study's purposes is to explore MBs in online learning and students' perceived learning performance.

Multitasking behavior is closely related to polychronicity. Polychronicity is an individual's natural tendency or preference for constructing time (Sanderson, 2012; Capdeferro et al., 2014), specifically defined as an individual's preference to participate in two or more tasks or events simultaneously and the belief that one's preference is the best way to do things (Bluedorn et al., 1999). Although studies of polychronicity are rare in educational settings, previous studies have shown differences in multitasking behavior among students at different levels of education (e.g., Jeong et al., 2010). However, polychronicity needs to be discussed more in online education. Based on the relationship between multitasking behavior and learning outcomes, as well as the relationship between multitasking behavior and polychronicity, this study explores the relationship between polychronicity, multitasking behavior, and learning outcomes by considering differences in educational stages.

2. Literature review

2.1. Online multitasking and learning performance

Distraction and inattention are often seen as a result of a learner's multitasking behavior (Cebollero-Salinas et al., 2022). Learning multitasking behavior is defined as distraction and nonsequential task switching of ambiguous tasks performed in a learning environment (Chen and Yan, 2016), and multitasking behavior seems to be more likely to arise in an online learning environment because information technology (IT) combined with the internet has brought changes to the working patterns in time allocation (Lee and Perry, 2001). IT makes learning no longer limited by time and space, thus changing traditional classroom experiences to being intermittent, multidirectional, and nonsequential and easily disrupting the traditional view of time and space in learning activities (Capdeferro et al., 2014). This also means that IT provides the opportunity for learners to multitask. Sun and Zhong (2020) pointed out that media users are more likely to engage in multitasking behavior and behavioral responses to new information and communication technology applications in today's mobile media era. Media multitasking behavior can be defined as using more than two types of media at the same time (Rideout et al., 2010) or quickly switching between tasks on the same media, such as working with multiple browsers or using several types of software simultaneously (Cardoso-Leite et al., 2015).

Multitasking behavior is often used to explore student learning outcomes. However, previous research has been administered mainly in the context of multitasking learning in a given time period to find the relationship between learning outcomes and multitasking behaviors, revealing that multitasking learning behaviors lead to anxiety (e.g., Seddon et al., 2021), poorer academic performance (e.g., Loh et al., 2016), or perceived lower learning performance (e.g., Fried, 2008).

2.2. Online multitasking and polychronicity

Szumowska et al. (2018) noted the difference between behavior and preference and argued that multitasking behavioral performance does not represent an individual's preference for multitasking. Preference for multitasking is nearly equivalent to polychronicity (Szumowska et al., 2018). As a natural trend or preference for constructing time (Sanderson, 2012; Capdeferro et al., 2014), polychronicity is defined as the preference for simultaneous involvement in two or more tasks or events and the belief that one's preference is the best way to do things (Bluedorn et al., 1999). This definition of general polychronicity was adapted to the computing context and called computer polychronicity (Davis et al., 2009).

According to Palmer and Schoorman (1999), there are three distinct dimensions that are typically associated with polychronic structure: time-use preference, context, and time tangibility. These three dimensions were again applied to individual-level polychronicity and called scheduling preference (SP), involvement with people (IP) and time tangibility (TT; Capdeferro et al., 2014). Polychrons refer to those who have polychronic character. First, they value interpersonal interaction and involvement with other people (high IP). Therefore, they cross the boundaries between work and nonwork domains, such as social and leisure activities. Second, polychrons favor simultaneous activities (high SP). They prefer doing many things at the same time, show tolerance to multitasks at a given time, and effectively address interruptions and unpredictability. Finally, polychrons have time-use preferences and believe that time is not tangible and cannot be managed but is the background for an event. Since time is not a tangible resource for them, people in high-context cultures do not think they have to finish their work in a certain period of time or think that it is problematic to leave tasks unfinished (Hecht and Allen, 2005; Capdeferro et al., 2014). However, Luo et al. (2021) found that polychrons in an IT-supported learning environment prefer to work in a timely, time-saving and schedule-based manner (high TT). As polychrons are more likely to undertake multiple tasks, they may be able to achieve more goals than monochrons under the same work conditions. Thus, time management enables polychrons to complete tasks on time and even complete more tasks. Given this, this study proposes the following hypothesis:

> H_1: Students tend to be more polychromic (including time tangibility, involvement with people, and scheduling preference) and their multitasking behavior is more frequent.

2.3. Polychronicity and learning performance

Capdeferro et al. (2014) argued that we may better understand students' learning behaviors in IT environments with the concept of polychronicity because various learning activities under IT environments occur in a polychronic manner, such as web browsing, interacting with peers on discussion boards, and responding to a teacher. Therefore, compared to monochronic learners, polychronic learners may feel more comfortable with a more flexible timeframe and more interactive environments. This can be explained by the person–environment fit (P–E fit) theory based on the interactionist theory of behavior (Chuang et al., 2016). The P–E fit theory argues that individuals' unique behaviors must be understood in the specific situations in which they occur (Payne et al., 1982). Interactionism regulates the interaction between personal characteristics and situations and argues that only when situational cues related to individual characteristics exist, that is, situation-trait relevance, can these individual-related characteristics produce behaviors consistent with the characteristics (Geukes et al., 2012). Based on this assumption, interactionists believes that certain environmental conditions can be adapted to individual characteristics, which can lead to better individual performance and higher satisfaction (Chuang et al., 2016). The existing research has also shown that matching personal characteristics with environmental characteristics is an effective predictor of overall work satisfaction (Hardin and Donaldson, 2014).

> H_2: Students tend to be more polychronic (including time tangibility, involvement with people, and scheduling preference) and their perceived learning performance is better.

Predictably, polychronicity is the most important indicator of MB and leads to such behavior (König et al., 2010; Kirchberg et al., 2015; Lepp et al., 2019). For example, Lepp et al. (2019) found that polychronicity is positively associated with multitasking during online learning activities. Polychrons may be more willing to challenge multiple tasks in a given time period and are satisfied with work that requires multitasking (Sanderson, 2012), and vice versa. Madjar and Oldham's (2006) research indicates that polychrons experience higher time pressure in a work environment where tasks are completed sequentially than in a work environment where tasks are alternated; for monochrons, the opposite is true. However, existing polychronicity studies have mainly focused on the concept of time orientation in the workplace (Manrai and Manrai, 1995; Hecht and Allen, 2005; Kirchberg et al., 2015; Bhattacharyya et al., 2018). The time factor in IT integration has been largely ignored in learning research, and polychronicity studies are rarely found in educational settings, although learners' time orientation is an extremely important variable in the educational environment, especially in an online learning environment, which has great potential for polychronicity (Barbera et al., 2012; Capdeferro et al., 2014). The learning process involved in IT makes events and tasks increasingly occur in a polychronic manner, and students are usually expected to engage in their learning activities accordingly (Lee and Perry, 2001; Capdeferro et al., 2014). However, not every learner feels comfortable engaging in this IT learning environment because of the differences in learners' time perceptions (for example, polychronicity and monochronicity; Capdeferro et al., 2014). Given this, this study proposes the following hypothesis:

> H_3: Students' multitasking behavior is more frequent, and their perceived learning performance is better.

2.4. The influence of education stage

Teenage students are known as digital natives, millennials, or i-Gen groups (Akçayır et al., 2016), which means that these students are taken for granted that they are good at using technology. However, although these teenage students were seen as a tech-savvy group, IT experience may have played a role in the differences. Multitasking behavior is malleable, and task-switching behavior can increase with usage (Cardoso-Leite et al., 2015; Moisala et al., 2016). Jeong et al. (2010) pointed out that college students have more media use time and exhibit more frequent media MBs than high school students. This shows that adolescent students may have differences in media MB across age groups due to different daily lifestyles. Given this, this study proposes the following hypothesis:

> H_4: The relationships among polychronicity (including time tangibility, involvement with people, and scheduling preference), multitasking behavior and perceived learning performance differ between high school students and college students.

3. Materials and methods

3.1. Participants

This cross-sectional study was administered among Taiwanese youths using convenience sampling methods. A total of 939 students from four high schools and two colleges responded to the questionnaire. After we excluded 66 incomplete and invalid questionnaires, 873 valid questionnaires remained. Furthermore, eight questionnaires were excluded because the participants self-reported that they had no experience using technology (such as mobile phones, computers or tablets) for learning. Ultimately, 865 questionnaires were included in the analysis. The background of the 865 respondents indicated that 51.21% were high school students and 48.79% were college students.

3.2. Instruments

3.2.1. The scale of MB in online learning

This scale was mainly used to evaluate students' participation in multiple tasks in the process of using digital technology, such as mobile phones and computers, to learn. These items were developed based upon related literature (Kaufman et al., 1991; Bluedorn et al., 1999; Poposki and Oswald, 2010; Haase et al., 2016). It is a 6-point

Likert scale ranging from "never" (1) to "always" (6). The exploratory factor analysis (EFA) showed that the Kaiser–Meyer–Olkin (KMO) test value was 0.76 (χ^2 = 1127.40, p < 0.001), the factor loadings ranged from 0.68 to 0.86, and the total explained variance was 62.49% (α = 0.80). The final scale contained 4 items (e.g., "When I use digital technology such as mobile phones and computers to learn, I do many things at the same time") attributed to one factor.

3.2.2. The scale of polychronicity in online learning

To measure the participants' polychronicity in online learning, we used the Polychronicity in IT-Supported Learning Scale (Luo et al., 2021). The scale consisted of three subscales in the context of using digital technology such as mobile phones and computers to learn: time tangibility (TT), involvement with people (IP), and scheduling preference (SP). In their study, the subscales were reliable (α = 0.84 to 0.89), and the model fit of the exploratory factor analysis results was ideal (CMIN/DF = 1.84, RMSEA = 0.05, TLI = 0.97, CFI = 0.97, SRMR = 0.043). In this study, the reliability of each aspect of the scale is: TT (α = 0.90), IP (α = 0.87), and SP (α = 0.85).

3.2.3. The scale of perceived online learning performance

This scale was mainly used to evaluate students' perceptions of the effectiveness of online learning, and these items were developed based upon related literature (e.g., Hsia and Tseng, 2005; Liu and Wu, 2017). Responses were given on a 6-point Likert scale ranging from 1 (definitely disagree) to 6 (definitely agree). The EFA showed that the KMO test value was 0.92 (χ^2 = 6384.59, p < 0.001), the factor loadings ranged from 0.71 to 0.84, and the total explained variance was 78.48%. The scale was divided into two aspects: knowledge acquisition (KA) contained 4 items (α = 0.92; e.g., "Online learning allows me to learn more knowledge") and learning satisfaction (LS) contained 5 items (α = 0.92; e.g., "I am satisfied that online learning has improved my confidence in learning").

3.3. Research ethics

All participation was voluntary, anonymous and confidential. We did not collect any information that could be provided to the participants, and the participants had the right to refuse to participate in the study at any time without any penalty. The analysis results are also presented in a holistic manner.

4. Results

4.1. Hypothetical SEM test

SEM for the total sample showed that $\chi^2_{(285)} = 1351.84, p < 0.001$. The test failed to obtain nonsignificant results, likely because the χ^2 value is sensitive to the number of cases (Bergh, 2015). Therefore, in large samples, the χ^2 value may not be an appropriate indicator, and alternative indicators will be needed (Luo et al., 2021). Other indicators used in this study showed that χ^2/df = 4.74, TLI = 0.92, CFI = 0.93, and RMSEA = 0.07, all satisfying the following criteria: χ^2/df < 5 (Jöreskog and Sörbom, 1993), TLI > 0.90 (Bentler and Bonett, 1980), CFI > 0.90 (Li, 2006), and RMSEA < 0.08 (McDonald and Ho, 2002), indicating that the fit between the model and the observed data was good.

4.2. Path relationships between multitasking, polychronicity, and perceived learning performance

As shown in Figure 1 and Table 1, the results of the structural model assessment revealed that the 11 main paths in the whole sample are significant except the path of "involvement with people" to "knowledge acquisition." Specifically, most of the research hypotheses H_1 to H_3 have gained statistical support.

4.3. Comparison the differences between education levels

The independent sample t test analysis results in Table 2 show no significant difference between the high school and college students in IP or KA. However, the college students scored significantly higher than the high school students in MB, TT, SP, and LS.

The paths of the SEM were examined according to education levels. The goodness-of-fit indices for the high school student sample were $\chi^2_{(285)}$ = 766.37, p < 0.001, χ^2/df = 2.69, TLI = 0.93, CFI = 0.93, and RMSEA = 0.06; the goodness-of-fit indices for the college student sample were $\chi^2_{(285)}$ = 1057.83, p < 0.001, χ^2/df = 3.71, TLI = 0.89, CFI = 0.90, and RMSEA = 0.08. In general, the criteria were very close to or higher than the standard, indicating acceptable goodness-of-fit for the sample data and justifying further analysis.

Table 3 shows the estimates of the hypothetical SEM for the two subsamples. In the high school student sample, the path relationship between IP and KA was nonsignificant. In addition, SP had no significant effect on either LS or KA. In the college student sample, IP had no significant effect on MB, LS, or KA.

Furthermore, multigroup analysis was performed, which allowed analysis of the coefficient of the different paths of the two groups. As shown in Table 3, there were significant differences between the groups on the paths between IP to MB and SP to LS. The research hypotheses were partially supported for H_4 (Figures 2, 3).

5. Discussion

In the predictive relationships of the entire sample, "scheduling preference" is an important predictor of "multitasking behavior" and perceived learning performance. This result is consistent with the general intuitive assumption that when learners prefer to do multiple things in a given time period and believe that this method is the best way to do things, they are more likely to adopt "multitasking behavior" (König et al., 2010); furthermore, a multitasking learning environment is in line with their personal interests and needs, which increases their satisfaction (Sanderson, 2012; Chuang et al., 2016). In addition, "time tangibility" can also predict "multitasking behavior" and perceived learning performance. That is, people who prefer to work in a timely, timesaving, and schedule-based manner actually participate in more tasks simultaneously and experience higher satisfaction in the process of using digital technology to learn. Luo et al. (2021) and Nonis et al.

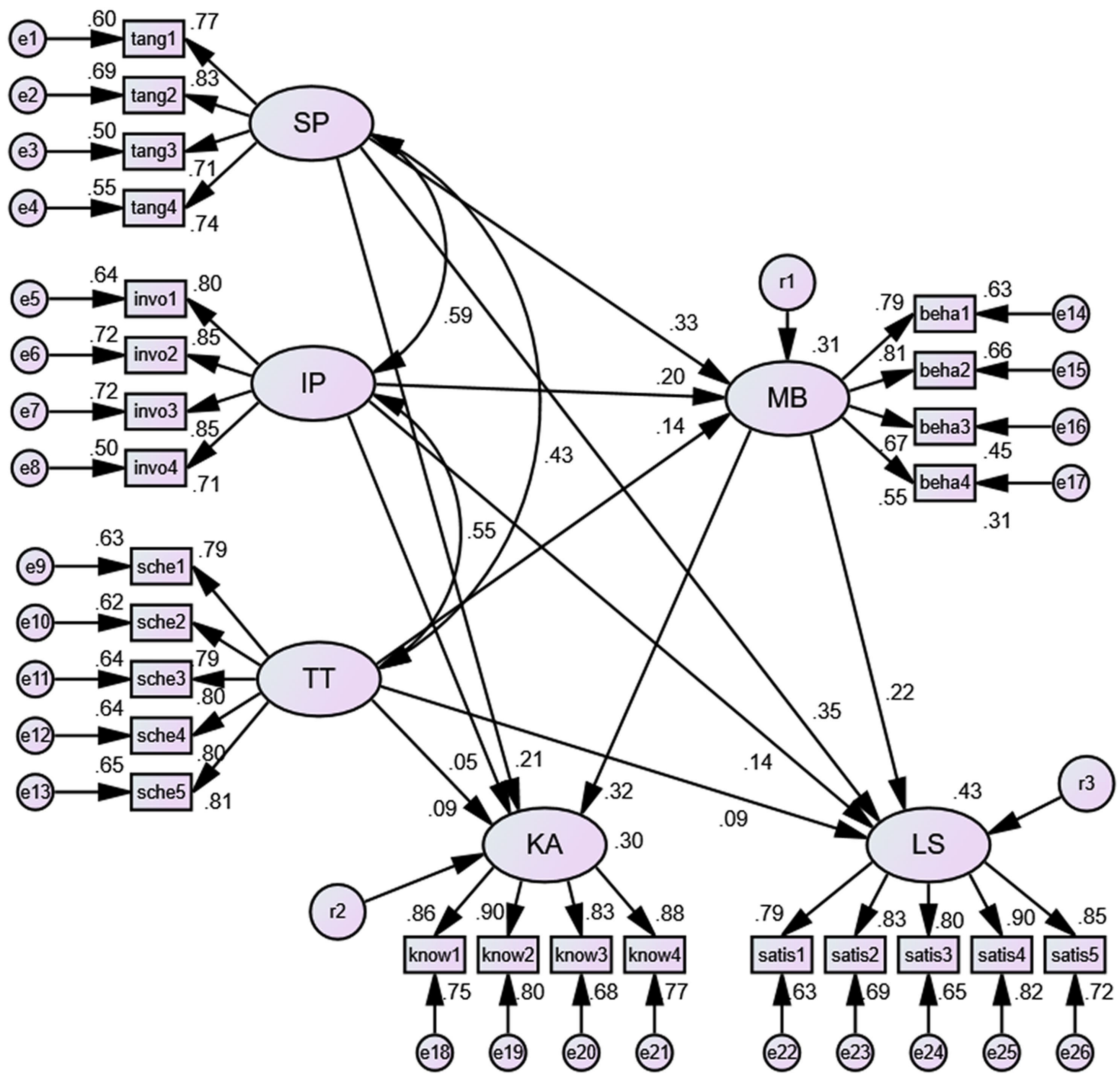


FIGURE 1
The standardization coefficient pattern of the SEM. SP=Scheduling preference, IP=Involvement with people, TT=Time tangibility, MB=Multitasking behavior, KA=Knowledge acquisition, LS=Learning satisfaction.

(2005) pointed out that "time tangibility," which is closely related to time management, enables people to complete tasks on time or even deal with more tasks, and it is in line with the online learning environment that requires learners to perform learning tasks in a polychronic manner (Lee and Perry, 2001; Capdeferro et al., 2014; Lepp et al., 2019). Finally, the "involvement with people" aspect of polychronicity in online learning did not predict "knowledge acquisition." It is possible that although emerging IT has multiple interpersonal interaction functions, Taiwanese students still seldom use these interactive functions in IT-supported learning, possibly because obedience and silence in classroom learning have long been aspects of the learning culture of Chinese students (Luo and Yang, 2018), which may extend to virtual online learning environments. Therefore, "involvement with people" may not reflect knowledge acquisition in the online learning environment.

Jeong et al. (2010) found that the frequency of media multitasking behaviors differed between high school students and college students because of the differences in their daily life schedules. This study aligns with a previous study (Jeong et al., 2010) and further finds that college students' frequency of "multitasking behavior," "scheduling preference" and "time tangibility" and their "learning satisfaction" in multitasking online learning environments were significantly higher than those of high school students. This may be because college students have more experience in using technology to learn than high school students (Jeong et al., 2010), which is evidenced in studies revealing more time overall in internet usage (Taiwan Network Information Center, 2018) and a higher percentage of internet addiction than high school students (Writer, 2022). Personal possession of technological media increases individuals' use of technological media. Accessibility provides individuals with more opportunities for media MBs, thereby improving their self-efficacy in the use of technological media and increasing their preference for "multitasking behavior" in technological media (Srivastava et al., 2016).

Furthermore, this study found that, for students at different educational stages, different aspects of polychronicity had different predictive relationships on online learning "multitasking behavior" and

TABLE 1 Path relationships among multitasking behavior, polychronicity, and perceived learning performance (n=865).

Path	β	S.E	C.R.	p
H_1 Polychronicity to Multitasking behavior				
Scheduling preference → Multitasking behavior	0.33***	0.05	6.68	<0.001
Involvement with people → Multitasking behavior	0.20***	0.05	3.91	<0.001
Time tangibility → Multitasking behavior	0.14***	0.04	3.19	<0.001
H_2 Polychronicity to perceived learning performance				
Scheduling preference → Learning satisfaction	0.35***	0.05	7.81	<0.001
Scheduling preference → Knowledge acquisition	0.22***	0.05	4.53	<0.001
Involvement with people → Learning satisfaction	0.14***	0.05	3.21	<0.001
Involvement with people → Knowledge acquisition	0.05	0.05	1.00	0.32
Time tangibility → Learning satisfaction	0.10*	0.04	2.49	0.013
Time tangibility → Knowledge acquisition	0.09*	0.04	2.13	0.034
H_3 Multitasking behavior to perceived learning performance				
Multitasking behavior → Learning satisfaction	0.22***	0.04	5.51	<0.001
Multitasking behavior → Knowledge acquisition	0.33***	0.04	7.37	<0.001

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 2 T test of multitasking, polychronicity, and perceived learning performance across educational levels.

	High school students		College students		t Value	p Value
	Mean	SD	Mean	SD		
Multitasking behavior	3.61	1.02	3.87	0.91	−3.91***	<0.001
Polychronicity						
Time tangibility	3.71	1.16	3.85	0.98	−2.00*	0.045
Involvement with people	3.68	1.15	3.67	1.10	0.04	0.97
Scheduling preference	3.41	1.13	3.80	1.03	−5.40***	<0.001
Perceived learning performance						
Knowledge acquisition	4.54	0.99	4.43	1.00	1.58	0.11
Learning satisfaction	3.92	1.12	4.11	0.99	−2.60**	0.009

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 3 Cross-group path coefficient invariance test results.

Path	Default model		Moderator model		$\Delta\chi^2$	p	Path coefficient	
	χ^2	df	χ^2	df			High school	College
SP → MB	1824.22	570	1825.73	571	1.51	0.22	0.26***	0.37***
IP → MB	1824.22	570	1829.64	571	5.42	0.02	0.30***	0.08
TT → MB	1824.22	570	1824.40	571	0.18	0.67	0.14*	0.15*
SP → LS	1824.22	570	1824.89	571	0.67	0.41	0.36***	0.28***
SP → KA	1824.22	570	1824.23	571	0.002	0.96	0.25***	0.21**
IP → LS	1824.22	570	1825.41	571	1.19	0.28	0.17*	0.09
IP → KA	1824.22	570	1824.34	571	0.12	0.73	0.00	0.04
TT → LS	1824.22	570	1833.76	571	9.54	0.002	0.02	0.25***
TT → KA	1824.22	570	1827.08	571	2.85	0.09	0.05	0.17*
MB → LS	1824.22	570	1824.68	571	0.46	0.50	0.19***	0.26***
MB → KA	1824.22	570	1824.64	571	0.42	0.52	0.33***	0.34***

SP = Scheduling preference, IP = Involvement with people, TT = Time tangibility, MB = Multitasking behavior, KA = Knowledge acquisition, LS = Learning satisfaction.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

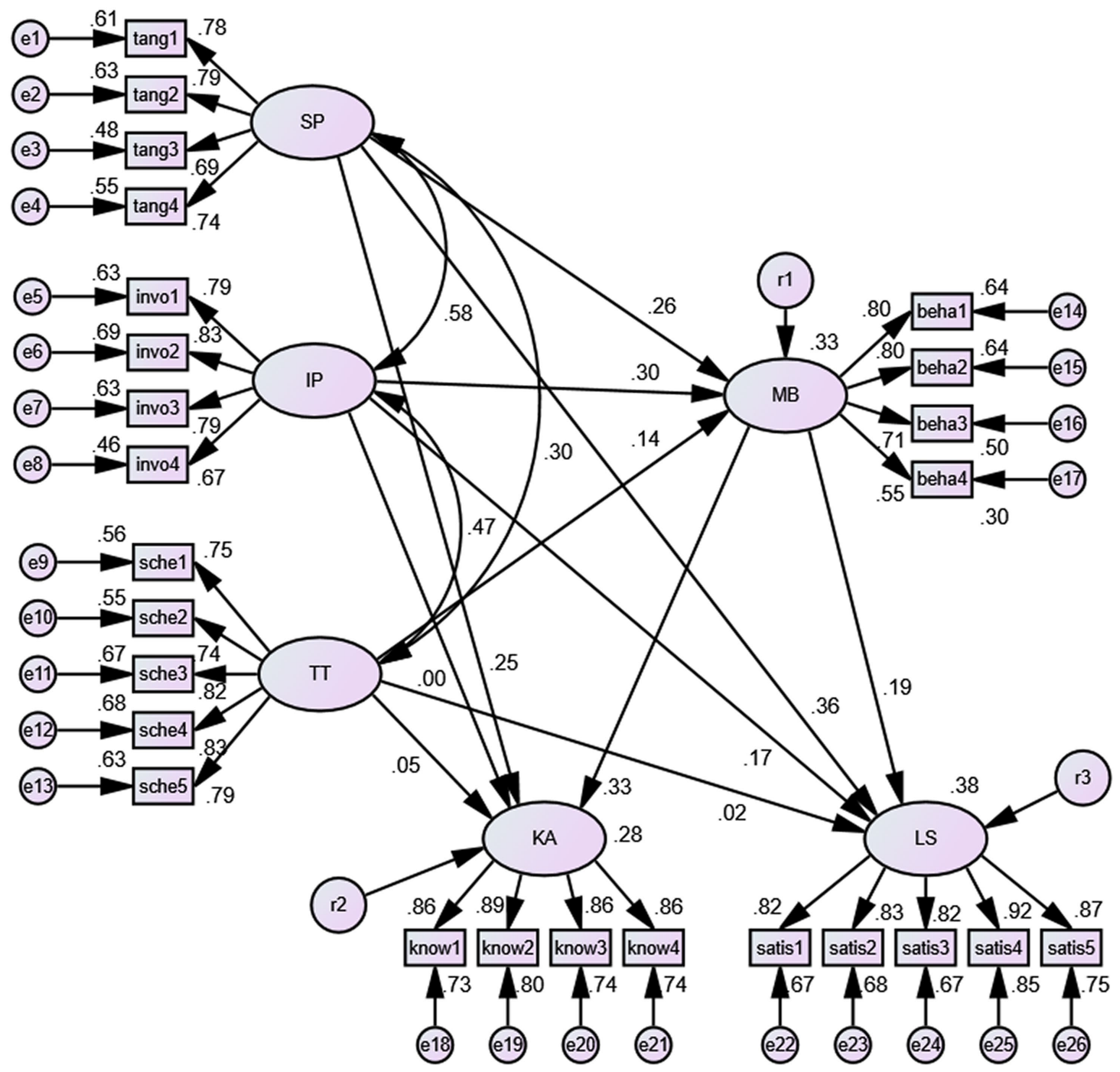


FIGURE 2
The standardization coefficient pattern of the SEM for the high school student sample. SP=Scheduling preference, IP=Involvement with people, TT=Time tangibility, MB=Multitasking behavior, KA=Knowledge acquisition, LS=Learning satisfaction.

perceived learning performance. "Scheduling preference" positively predicted "multitasking behavior" and perceived learning performance for both high school and college students. This result seems not surprising. As mentioned above, students who prefer and believe that multitasking is an ideal method are more likely to prefer multitasking behavior (König et al., 2010) and are more satisfied with the online learning environment, which may be potentially full of multitasking (Sanderson, 2012; Chuang et al., 2016). The statistical results of different education stages further explain that this phenomenon applies to many young people.

"Time tangibility" had a significant impact on the learning performance (including "knowledge acquisition" and "learning satisfaction") of college students but not on that of high school students. Especially in "time tangibility" and "learning satisfaction," the path coefficients of the two samples were significantly different. This may be because students with "time tangibility" feel that they can complete tasks on time or even complete more tasks through time management, resulting in higher academic performance (Nonis et al., 2005; Luo et al., 2021). In addition, "time tangibility," which is related to time management, relates to students' successful self-regulated learning. Liu et al. (2014) showed that time management was associated with self-regulated learning and that both positively predicted learning engagement. Wu (2017) found that students' media multitasking indirectly predicted course grades through perceived attention and self-management strategies. Students' learning performance on media multitasking can be improved by using attention regulation strategies, while students at higher stages of education have a clearer understanding of their learning patterns and are better able to regulate their learning in specific situations (Song and Vermunt, 2021).

High school students were more likely to be involved with people and had a higher frequency of multitasking behaviors, suggesting that media multitaskers were more likely to value online socializing (Zhong et al., 2011). However, this relationship did not apply to college students, possibly due to the aforementioned regulation strategies. College students can master multitasking more effectively than high school students through self-regulation strategies without being affected by interpersonal relationships. The positive effect of IP on "learning satisfaction" was also shown among high school students

10.3389/fpsyg.2023.1131765

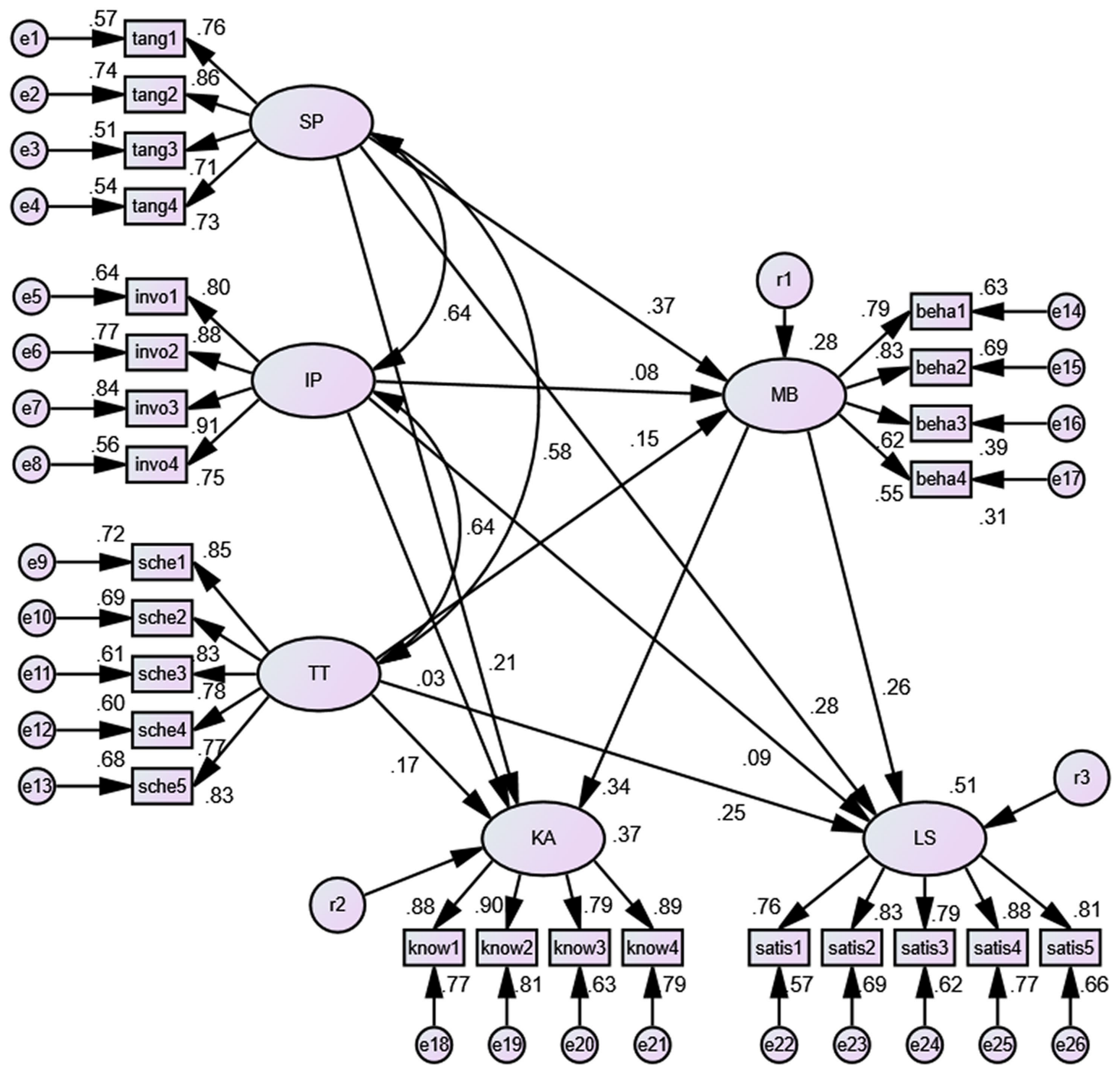


FIGURE 3
The standardization coefficient pattern of the SEM for the college student sample. SP=Scheduling preference, IP=Involvement with people, TT=Time tangibility, MB=Multitasking behavior, KA=Knowledge acquisition, LS=Learning satisfaction.

but not among college students, highlighting the importance of online socialization for high school students' learning. For high school students who are primarily in adolescence, a desire to "fit in" with peers is characteristic (Wang and Chen, 2018), and "involvement with people" satisfies their desires, leading to positive "learning satisfaction." However, "involvement with people" did not significantly predict "knowledge acquisition" in either college or high school students. Previous research has shown that interpersonal engagement in the classroom, such as collaborative learning and classroom dialog, must be a part of effective teacher scaffolding to produce good learning benefits (Muhonen et al., 2016; Chan, 2020). Therefore, online "involvement with people" may lack scaffolding guidance and effective collaboration and dialog, leading to reduced "knowledge acquisition."

Interpretation of our findings should be made with caution due to the following limitations. First, due to individual subjectivity, students' self-reports may not reflect their actual multitasking behaviors. Second, although we found that polychronicity is related to multitasking behavior and perceived learning performance, these variables are possibly not causally related in a practical sense. To clarify this issue, other research methods, such as experiments or observation methods, should be used in future research to deeply examine polychronicity, multitasking behavior and learning performance in an online learning environment. Finally, the "knowledge acquisition" investigated in this study is the result of students' subjective perceptions, lacking objective data on academic knowledge and academic performance. Future research should consider collecting multiple academic achievements. Although there are limitations in this study, it is still helpful for developing a preliminary understanding of the polychronicity, multitasking and learning performance of online learning, as well as the differences across educational stages, and can be regarded as a basis for subsequent research development.

6. Conclusion

Media multitasking behavior has become a response to the application of digital technology, which also reflects that online learning multitasking behavior suitable for digital media ecology results in higher learning performance for both high school and college students. However, students' mastery of self-regulation strategies at different educational stages caused by learning experiences may influence digital learning multitasking behavior and learning performance. In addition, the quality of involvement with people may be one of the antecedent factors affecting students' online learning outcomes. Therefore, it is suggested that future research consider the moderating effects of self-regulation strategies and the quality of cooperative scaffolding and classroom dialog on polychronicity, multitasking behavior and learning performance in an online learning environment.

Future studies can also further examine what motivates heavy polychronic, or molychronic, learners to perform online multitasking behaviors in their learning engagement and how those learners activate and regulate cognitive capacity in their processing and consuming different media tasks at hand (cognitive/attentional control, executive functioning, media type attention, superficial/deep level of processing) and their cognitive/affective/physiological emotions (positive feelings accomplishments or negative stress) associated with their online multitasking and polychronicity activities (the perceived feeling of control, cognitive overload, perceived ability to process all information). Finally, future studies can further examine behavioral and neural as well as inhibiting or stimulating indicators related to the cognitive/affective/social mechanism of students' engagement in polychronicity and multitasking behavior in an online learning environment.

Data availability statement

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

Ethics statement

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

Author contributions

YL and SY contributed to conception and design of the study. CL collected the data and organized the database. YL performed the statistical analysis and wrote the first draft of the manuscript. YL, SK, SY, and CL wrote sections of the manuscript. 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.

References

Akçayır, M., Dündar, H., and Akçayır, G. (2016). What makes you a digital native? Is it enough to be born after 1980? *Comput. Hum. Behav.* 60, 435–440. doi: 10.1016/j.chb.2016.02.089

Almahasees, Z., Mohsen, K., and Amin, M. O. (2021). Faculty's and students' perceptions of online learning during Covid-19. *Front. Educ.* 6. doi: 10.3389/feduc.2021.638470

Barbera, E., Gros, B., and Kirschner, P. (2012). Temporal issues in e-learning research: a literature review. *Br. J. Educ. Technol.* 43, E53–E55. doi: 10.1111/j.1467-8535.2011.01255.x

Bdair, I. A. (2021). Nursing students' and faculty members' perspectives about online learning during COVID-19 pandemic: a qualitative study. *Teach. Learn. Nurs.* 16, 220–226. doi: 10.1016/j.teln.2021.02.008

Bentler, P. M., and Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* 88, 588–606. doi: 10.1037/0033-2909.88.3.588

Bergh, D. (2015). "Sample Size and Chi-Squared Test of Fit—A Comparison Between a Random Sample Approach and a Chi-Square Value Adjustment Method Using Swedish Adolescent Data." in *Pacific Rim Objective Measurement Symposium (PROMS) 2014 Conference Proceedings*. eds. Zhang, Q., and Yang, H. Berlin, Heidelberg: Springer.

Bhattacharyya, A., Suresh, A., and Selvaraj, J. T. (2018). Contextualizing polychronicity: integrating work and non-work domains. *Indian J. Health Well-being* 9, 19–24.

Bluedorn, A. C., Kalliath, T. J., Strube, M. J., and Martin, G. D. (1999). Polychronicity and the inventory of Polychronic values (IPV): the development of an instrument to measure a fundamental dimension of organizational culture. *J. Manag. Psychol.* 14, 205–231. doi: 10.1108/02683949910263747

Capdeferro, N., Romero, M., and Barberàa, E. (2014). Polychronicity: review of the literature and a new configuration for the study of this hidden dimension of online learning. *Distance Educ.* 35, 294–310. doi: 10.1080/01587919.2015.955249

Cardoso-Leite, P., Green, C. S., and Bavelier, D. (2015). On the impact of new technologies on multitasking. *Dev. Rev.* 35, 98–112. doi: 10.1016/j.dr.2014.12.001

Chan, M. (2020). A multilevel SEM study of classroom talk on cooperative learning and academic achievement: Does cooperative scaffolding matter?. *International Journal of Educational Research,* 101, 101564. doi: 10.1016/j.ijer.2020.101564

Cebollero-Salinas, A., Cano-Escoriaza, J., and Orejudo, S. (2022). Are emotional e-competencies a protective factor against habitual digital behaviors (media multitasking, cybergossip, phubbing) in Spanish students of secondary education? *Comput. Educ.* 181:104464. doi: 10.1016/j.compedu.2022.104464

Chen, K. Z., and Li, S. C. (2021). Sequential, typological, and academic dynamics of self-regulated learners: learning analytics of an undergraduate chemistry online course. *Comput. Educ. Artif. Intell.* 2:100024. doi: 10.1016/j.caeai.2021.100024

Chen, Q., and Yan, Z. (2016). Does multitasking with mobile phones affect learning? A review. *Comput. Hum. Behav.* 54, 34–42. doi: 10.1016/j.chb.2015.07.047

Chuang, A., Shen, C. T., and Judge, T. A. (2016). Development of a multidimensional instrument of person–environment fit: the perceived person–environment fit scale (PPEFS). *Appl. Psychol.* 65, 66–98. doi: 10.1111/apps.12036

Davis, J. M., Lee, L. S., and Yi, M. Y. (2009). Time-user preference and technology acceptance: measure development of computer polychronicity. *Am. J. Bus.* 24, 23–32. doi: 10.1108/19355181200900008

Elida, T., Nugroho, W., and Suyudi, I. (2012). Cost effectiveness of web based learning. *Procedia. Soc. Behav. Sci.* 65, 1071–1076. doi: 10.1016/j.sbspro.2012.11.373

Fried, C. (2008). In-class laptop use and its effects on student learning. *Computers & Education*, 50, 906–914.

Geukes, K., Mesagno, C., Hanrahan, S. J., and Kellmann, M. (2012). Testing an interactionist perspective on the relationship between personality traits and performance under public pressure. *Psychol. Sport Exerc.* 13, 243–250. doi: 10.1016/j.psychsport.2011.12.004

Gumede, L., and Badriparsad, N. (2022). Online teaching and learning through the students' eyes – uncertainty through the COVID-19 lockdown: a qualitative case study in Gauteng province, South Africa. *Radiography* 28, 193–198. doi: 10.1016/j.radi.2021.10.018

Haase, R. F., Ferreira, J. A., Fernandes, R. I., Santos, E. J. R., and Jome, M. L. (2016). Development and validation of a revised measure of individual capacities for tolerating information overload in occupational settings. *J. Career Assess.* 24, 130–144. doi: 10.1177/1069072714565615

Hardin, E. E., and Donaldson, J. R. III (2014). Predicting job satisfaction: a new perspective on person–environment fit. *J. Couns. Psychol.* 61, 634–640. doi: 10.1037/cou0000039

Hecht, T. D., and Allen, N. J. (2005). Exploring links between polychronicity and well-being from the perspective of person-job fit: does it matter if you prefer to do only one thing at a time? *Organ. Behav. Hum. Decis. Process.* 98, 155–178. doi: 10.1016/j.obhdp.2005.07.004

Hsia, J. W., and Tseng, A. H. (2005). The effects of characteristics of e-learning on employee satisfaction: the mediating role of perception of e-learning. Chung Hua. *J. Manag.* S_2, 43–52.

Jeong, S. H., Zhang, W., Fishbein, M., Davis, E., Bleakley, A., Jordan, A., et al. (2010). "Multiple media use and multitasking with media among high school and college students: a diary method" in *The Interrelationship of Business and Communication*. ed. M. B. Hinner (Frankfurt am Main, Germany: Peter Lang GmbH), 409–434.

Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*. Chicago, IL; Hillsdale, NJ: Scientific Software International; Lawrence Erlbaum Associates, Inc.

Kaufman, C. F., Lane, P. M., and Lindquist, J. D. (1991). Exploring more than 24 hours a day: a preliminary investigation of polychronic time use. *J. Consum. Res.* 18, 392–401. doi: 10.1086/209268

Kim, S., Jeong, S. H., Kim, H. S., and Jeong, Y. J. (2022). Academic success of online learning in undergraduate nursing education programs in the COVID-19 pandemic era. *J. Prof. Nurs.* 38, 6–16. doi: 10.1016/j.profnurs.2021.10.005

Kirchberg, D. M., Roe, R. A., and van Eerde, W. (2015). Polychronicity and multitasking: a diary study at work. *Hum. Perform.* 28, 112–136. doi: 10.1080/08959285.2014.976706

König, C. J., Oberlarcher, L., and Kleinmann, M. (2010). Personal and situational determinants of multitasking at work. *J. Pers. Psychol.* 9, 99–103. doi: 10.1027/1866-5888/a000008

Lee, H., and Perry, M. (2001). Contextualising virtuality: polychronicity and multipresence. In *International Conference On Spacing And Timing: Rethinking Globalization And Standardization*. Palermo, Italy. Available at: https://www.academia.edu/17147050/Contextualising_virtuality_polychronicity_and_multipresence

Lemay, D. J., Bazelais, P., and Doleck, T. (2021). Transition to online learning during the COVID-19 pandemic. *Comput. Hum. Behav. Reports* 4:100130. doi: 10.1016/j.chbr.2021.100130

Lepp, A., Barkley, J. E., Karpinski, A. C., and Singh, S. (2019). College students' multitasking behavior in online versus face-to-face courses. *SAGE Open* 9:215824401882450. doi: 10.1177/2158244018824505

Li, M. N. (2006). *An Introduction to Graphic Amos and Its Uses in Research*. Taipei, Taiwan: Wunan.

Liu, J., Liu, Y., and Yang, J. (2014). Impact of learning adaptability and time management disposition on study engagement among Chinese baccalaureate nursing students. *J. Prof. Nurs.* 30, 502–510. doi: 10.1016/j.profnurs.2014.05.002

Liu, T. Z., and Wu, C. S. (2017). A study on the influencing factors of continuance using intention of food and beverage management students after studying a certified e-learning course. *J. Hosp. Tour.* 14, 19–35.

Loh, K. K., Tan, B. Z. H., and Lim, S. W. (2016). Media multitasking predicts video-recorded lecture learning performance through mind wandering tendencies. *Comput. Hum. Behav.* 63, 943–947. doi: 10.1016/j.chb.2016.06.030

Luo, Y. F., and Yang, S. C. (2018). "Ubiquitous learning: the effectiveness of mobile application import learning" in *New Frontiers in Electronic Commerce and e-Society 2018*. ed. T. P. Liang (New Taipei City, Taiwan: Future Career Publishing Corporation), 268–281.

Luo, Y. F., Yang, S. C., and Lu, C. M. (2021). Development of a multidimensional polychronicity scale for information technology learning. *J. Comput. Assist. Learn.* 37, 387–395. doi: 10.1111/jcal.12496

Madjar, N., and Oldham, G. R. (2006). Task rotation and polychronicity: effects on individuals' creativity. *Hum. Perform.* 19, 117–131. doi: 10.1207/s15327043hup1902_2

Manrai, L. A., and Manrai, A. K. (1995). Effects of cultural-context, gender, and acculturation on perceptions of work versus social/leisure time usage. *J. Bus. Res.* 32, 115–128. doi: 10.1016/0148-2963(94)00034-C

Maqableh, M., and Alia, M. (2021). Evaluation online learning of undergraduate students under lockdown amidst COVID-19 pandemic: the online learning experience and students' satisfaction. *Child Youth Serv. Rev.* 128:106160. doi: 10.1016/j.childyouth.2021.106160

McDonald, R. P., and Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychol. Med.* 7, 64–82. doi: 10.1037/1082-989X.7.1.64

Moisala, M., Salmela, V., Hietajärvi, L., Salo, E., Carlson, S., Salonen, O., et al. (2016). Media multitasking is associated with distractibility and increased prefrontal activity in adolescents and young adults. *NeuroImage* 134, 113–121. doi: 10.1016/j.neuroimage.2016.04.011

Muhonen, H., Rasku-Puttonen, H., Pakarinen, E., and Poikkeus, A. M. (2016). Scaffolding through dialogic teaching in early school classrooms. *Teaching and Teacher Education* 55, 143–154. doi: 10.1016/j.tate.2016.01.007

Nonis, S. A., Teng, J. K., and Ford, C. W. (2005). A cross-cultural investigation of time management practices and job outcomes. *Int. J. Intercult. Relat.* 29, 409–428. doi: 10.1016/j.ijintrel.2005.05.002

Palmer, D., and Schoorman, F. D. (1999). Unpackaging the multiple aspects of time in polychronicity. *J. Bus. Psychol.* 14, 323–345. doi: 10.1108/02683949910263918

Payne, R. L., Fineman, S., and Jackson, P. R. (1982). An interactionist approach to measuring anxiety at work. *J. Occup. Organ. Psychol.* 55, 13–25. doi: 10.1111/j.2044-8325.1982.tb00074.x

Poposki, E. M., and Oswald, F. L. (2010). The multitasking preference inventory: toward an improved measure of individual differences in polychronicity. *Hum. Perform.* 23, 247–264. doi: 10.1080/08959285.2010.487843

Rideout, V. J., Foehr, U. G., and Roberts, D. F. (2010). *Generation M2: Media in the Lives of 8-to 18-year-olds*. San Francisco, CA: Henry J. Kaiser Family Foundation.

Sanderson, K. R. (2012). Time orientation in organizations: Polychronicity and multitasking doctoral dissertation. Available at: http://digitalcommons.fiu.edu/etd/738

Seddon, A. L., Law, A. S., Adams, A. M., and Simmons, F. R. (2021). Individual differences in media multitasking ability: the importance of cognitive flexibility. *Comput. Hum. Behav. Reports* 3:100068. doi: 10.1016/j.chbr.2021.100068

Song, Y., and Vermunt, J. D. (2021). A comparative study of learning patterns of secondary school, high school and college students. *Stud. Sci. Educ.* 68:100958. doi: 10.1016/j.stueduc.2020.100958

Srivastava, J., Nakazawa, M., and Chen, Y.-W. (2016). Online, mixed, and offline media multitasking: role of cultural, socio-demographic, and media factors. *Comput. Hum. Behav.* 62, 720–729. doi: 10.1016/j.chb.2016.04.040

Sun, T., and Zhong, B. (2020). Multitasking as multisensory behavior: revisiting media multitasking in the perspective of media ecology theory. *Comput. Hum. Behav.* 104:106151. doi: 10.1016/j.chb.2019.09.027

Szumowska, E., Popławska-Boruc, A., and Kossowska, M. (2018). How many things do you (like to) do at once? The relationship between need for closure and multitasking preference and behavior. *Pers. Individ. Differ.* 134, 222–231. doi: 10.1016/j.paid.2018.06.023

Taiwan Network Information Center (2018). Taiwan Internet Report 2018. Retrieved December 25, 2022, from https://report.twnic.tw/2018/TrendAnalysis_internetUsage_en.html.

Turan, Z., Kucuk, S., and Cilligol Karabey, S. (2022). The university students' self-regulated effort, flexibility and satisfaction in distance education. *Int. J. Educ. Technol. High. Educ.* 19:35. doi: 10.1186/s41239-022-00342-w

Wang, C. H., and Chen, K. Y. (2018). *Structural Equation Modeling: Analysis Using AMOS*. Taipei, Taiwan: Wunan.

Writer, S. (2022). Internet use increasing among older age groups. Taipei Times.https://www.taipeitimes.com/News/taiwan/archives/2022/03/27/2003775541

Wu, J.-Y. (2017). The indirect relationship of media multitasking self-efficacy on learning performance within the personal learning environment: implications from the mechanism of perceived attention problems and self-regulation strategies. *Comput. Educ.* 106, 56–72. doi: 10.1016/j.compedu.2016.10.010

Zhong, B., Hardin, M., and Sun, T. (2011). Less effortful thinking leads to more social networking? The associations between the use of social network sites and personality traits. *Comput. Hum. Behav.* 27, 1265–1271. doi: 10.1016/j.chb.2011.01.008

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 23 May 2023
DOI 10.3389/fpsyg.2023.1071985

OPEN ACCESS

EDITED BY
Zhongling Pi,
Shaanxi Normal University, China

REVIEWED BY
Erinc Ercag,
University of Kyrenia, Cyprus
Yi Zhang,
Central China Normal University, China

*CORRESPONDENCE
Cheng Fei
✉ feicheng@bnu.edu.cn

RECEIVED 17 October 2022
ACCEPTED 20 April 2023
PUBLISHED 23 May 2023

CITATION
Tlili A, Sun T, Denden M, Kinshuk, Graf S, Fei C and Wang H (2023) Impact of personality traits on learners' navigational behavior patterns in an online course: a lag sequential analysis approach.
Front. Psychol. 14:1071985.
doi: 10.3389/fpsyg.2023.1071985

Impact of personality traits on learners' navigational behavior patterns in an online course: a lag sequential analysis approach

Ahmed Tlili[1], Tianyue Sun[2], Mouna Denden[3,4], Kinshuk[5], Sabine Graf[6], Cheng Fei[1]* and Huanhuan Wang[1]

[1]Smart Learning Institute of Beijing Normal University, Beijing, China, [2]Department of Health and Behavior Studies, Teachers College, Columbia University, New York, NY, United States, [3]University Polytechnique Hauts-de-France, LAMIH, CNRS, UMR 8201, Valenciennes, France, [4]INSA Hauts-de-France, Valenciennes, France, [5]University of North Texas, Denton, TX, United States, [6]School of Computing and Information Systems, Athabasca University, Edmonton, AB, Canada

Personality is considered as the internal factor that defines a person's behavior. Therefore, providing adaptive features and personalized support in online learning by considering learners' personalities can improve their learning experiences and outcomes. In this context, several research studies have investigated the impact of personality differences in online learning. However, little is known about how personality differences affect learners' behavior while learning. To fill this gap, this study applies a lag sequential analysis (LSA) approach to understand learners' navigational behavior patterns in an online three-months course of 65 learners based on their personalities. In this context, the five factor model (FFM) model was used to identify learners' personalities. The findings revealed that learners with different personalities use different strategies to learn and navigate within the course. For instance, learners high in extraversion tend to be extrinsically motivated. They therefore significantly navigated between viewing the course module and their personal achievements. The findings of this study can contribute to the adaptive learning field by providing insights about which personalization features can help learners with different personalities. The findings can also contribute to the field of automatic modeling of personality by providing information about differences in navigational behavior based on learners' personalities.

KEYWORDS

personality, online learning, navigational behaviors, adaptive systems, distance education, lag sequential analysis

1. Introduction

Taking into consideration learners' individual differences in computer-based learning is important as these differences can define how a given learner behave in a learning environment (Tlili et al., 2016). For instance, learners' individual differences can affect the learning process, where some learners might find it easy to learn a particular course, whereas others find the same course difficult (Jonassen and Grabowski, 1993). Personality is widely identified as an important indicator of individual differences (Irani et al., 2003; Essalmi et al., 2017). It can affect several important predictors of academic performance, such as learning approaches, effective learning and self-regulation strategies, cognitive abilities, and academic motivation (Barrick and Mount, 1996; Diseth, 2003; Zhang, 2003; Bidjerano and Yun Dai, 2007; Clark and Schroth, 2010; Swanberg and Martinsen, 2010).

In a comprehensive review, Tlili et al. (2016) highlighted the importance of understanding learners' behaviors based on their personalities to provide adaptive computer-based learning experiences accordingly. Fatahi (2019) also reported that adaptive e-learning environments based on personality improved learners' performance. Lai et al. (2019) further emphasized the importance of adaptive e-learning based on learners' personality traits, which facilitated learning efficiency and met learners' demands; thus, learners may understand the learning materials better. Denden et al. (2021) found that different personality traits prefer different game elements in gamification. They therefore recommended providing adaptive design of gamified online learning systems based on learners' personality traits. It is therefore important to investigate how learners with different personalities use and navigate through online courses. Navigational behavior refers to how learners navigate through the course and in which order they visit different kinds of learning objects and activities (Graf et al., 2010). Adaptive navigation support, in terms of recommending learners a suitable way through learning materials and activities, is one of the two main ways for adding adaptive functionality to learning systems (Brusilovsky, 2001). Hence, learning about how learners with different personalities navigate through an online course can help to provide adaptive learning features and provide personalized support, which enhance their learning experiences and outcomes.

While there is much agreement on the effects of personality on learning, limited empirical findings are found related to leaners' navigational behavior patterns in online courses based on their personalities. Therefore, to fill this gap, this study analyzes the learners' log files on the learning management system Moodle using lag sequential analysis (LSA) to identify their navigational behavior patterns based on their personalities. LSA allows conducting in-depth investigation on learning behaviors or event chains that occur at frequencies greater than chance (Sackett, 1978). In education specifically, LSA takes transitional relationships into consideration to identify temporal differences in learning behaviors (Chen et al., 2017). For example, Graf et al. (2010) applied LSA to investigate the impact of learners' learning styles on their navigational behaviors in an online course. Cheng et al. (2019) applied LSA to explore the process of co-construction of knowledge where 24 sixth-grade learners are competing in an augmented reality mathematic game. Tlili et al. (2021) also applied LSA to examine the behavioral pattern differences among learners from either China, Tunisia or Serbia who enrolled in an online six-week course. Wang et al. (2022) further applied LSA to investigate how gender might moderate learners' online learning behavioral patterns.

To identify the learners' personalities, this study relies on the five-factor model (FFM), which is one of the most common psychological models (Franić et al., 2014) and is frequently used in education (Tlili et al., 2016). It attributes five personality dimensions, namely openness, conscientiousness, extraversion, agreeableness, neuroticism, often abbreviated as OCEAN. Each of these dimensions is discussed in the next section.

2. Theoretical background

Online learning differs from traditional face-to-face learning in the way that it does not require learners to present themselves in an actual classroom setting (Wang et al., 2013). Learners who enroll in online courses have greater flexibility in their learning process as they decide when, where and how to navigate the learning materials (Wang et al., 2013; Hong et al., 2021). With the rapid evolution of technology, online learning research has gained an increasing attention as through technology it can enhance learners' learning engagement and achievement (Wengrowicz et al., 2018). In this context, several studies called for more investigation on the factors that could affect the process of online learning, such as personality (Sun et al., 2020; Hong et al., 2021).

2.1. Personality

Personality is defined as the internal factor that makes a person's behavior consistent over time (Child, 1968). It accounts for the individual differences in emotional, interpersonal, motivational and other aspects (McCrae and John, 1992; Gustavsson et al., 2003). Numerous personality theories and models exist in the literature, such as five factor model (FFM; McCrae and John, 1992), Myer Briggs types (Myers et al., 1998) and Han Eysenck's model (Boeree, 2006).

This study uses the Five Factor Model (FFM) of personality, which is one of the most accepted personality models in the literature (Trull and Sher, 1994; Barrick et al., 2003; Costa and McCrae, 2009) to describe learners' personality traits. It is validated across various countries and cultures (John et al., 2008; Gurven et al., 2013; Novikova and Vorobyeva, 2019; Murphy et al., 2021). FFM is derived from common language descriptors (DeYoung et al., 2007; Ackerman, 2020). It is an accurate personality model and it is easy to be reused in different contexts (DeYoung et al., 2007). FFM consists of five personality dimensions, namely (John and Srivastava, 1999): (1) extraversion focuses on a person's sociability, activeness and enthusiasm; (2) agreeableness emphasizes a person's compliance, altruism and generosity; (3) conscientiousness relates to a person's self-discipline, achievement-striving and responsibility; (4) neuroticism is concerned with a person's emotional stability, hostility and impulsivity; and (5) openness refers to a person's interest in new experience, curiosity and imagination.

2.2. Effects of personality on online learning

Personality has been proved essential to create an adaptive online learning environment. Many studies have emphasized the importance of providing adaptive computer-assisted learning environments based on personality, as personality affects individual learning preferences and learning processes (Tlili et al., 2016; Fatahi, 2019; Lai et al., 2019). For instance, Fatahi (2019) designed an adaptive online learning environment by gathering introverts' learning preferences. The adaptive e-learning system helped learners perform better with a higher grade. Harrington and Loffredo (2010) claimed that compared to learners high in extraversion who preferred the traditional in-person learning, learners low in extraversion found online learning more comfortable since they did not need to do face-to-face communication with their peers. In terms of online learning adoption, while agreeableness predicted the lowest adoption value, openness and conscientiousness were positively correlated with online learning adoption (Haron and Sahar, 2010; Harrington and Loffredo, 2010).

Furthermore, Arockiam and Selvaraj (2013) proved that personality plays an important role in learners' preferences of the design of online learning interfaces. Specifically, learners high in extraversion found it easier to recall information colored in blue with "Times" font style, whereas learners high in neuroticism found it easier to recall information colored in green with "Times" font style. Personality can also mediate the learning process (Al-Dujaily et al., 2013). For instance, learners high in extraversion were prone to critical thinking learning approaches in online learning environments (Zhang, 2003); learners high in neuroticism preferred highly structured learning environments (Furnham, 1992); and learners high in conscientiousness preferred organizing learning approaches and advanced time management (Moldasheva and Mahmood, 2014). Furthermore, learners with diverse personality types significantly engaged in learning activities differently. For instance, Lee and Lee (2006) found that compared to learners low in extraversion, learners high in extraversion were more social and interactive. They, therefore, posted more messages in the web-based discussion forums. Yu (2021) found that learners high in agreeableness, conscientiousness, and openness personality traits outperformed those high in extraversion and neuroticism personality traits in online learning outcomes during the covid-19. Finally, Denden et al. (2018) and Tlili et al. (2019) relied on the FFM personality model and revealed that learners' personalities affect the way learners engage in different learning environments, including educational games and Moodle.

2.3. Research gap and the purpose of the study

Research on online learning effectiveness has experienced a shift towards focusing on learner characteristics or differences like personality traits (Chai et al., 2022). However, only a few studies focused on the relationship between personality and online learning behaviors, and these studies focused on analyzing single behaviors, such as note taking or discussing course related topics with peers (Lee and Lee, 2006; Wu and Hou, 2015). Shang et al. (2020) mentioned that such single behaviors cannot reflect the learners' cognitive engagement and learning behaviors characteristics in details. Yang et al. (2016) further mentioned that investigating the behavior transformation sequence can deeply explain how learners engaged in a given course and their cognitive behaviors. As such, it would be important to do further studies that consider more complex behaviors and investigate their relationship to personality traits.

Additionally, the existing studies aimed at drawing connections between personality and online learning behaviors focusing on a specific personality trait, such as procrastination (Hong et al., 2021). To the best of our knowledge, no research has examined the effects of a personality model, such as the five-factor model (with five personality traits), on navigational behavior patterns in an online course.

To cover this gap, this study complements the available body of research by analyzing the learners' personality traits based on the FFM model (considering its five dimensions: extraversion, conscientiousness, agreeableness, neuroticism and openness) and their navigational behavior patterns in an online course. Specifically, this study answers the following research question: *How do personality traits of the Five Factor Model affect navigation behavioral patterns of learners in an online course?* To answer this research question, this study applies lag sequential analysis (LSA) to investigate the impact of learners' personality on their navigational behavior patterns in a three-months online course in a public university. LSA was used in this study because it can reveal knowledge-construction behaviors' temporal dynamics (Zhang et al., 2017; Sun et al., 2021). It estimates the probability of a given behavior to occur, as well as its successive behavior (Bakeman and Gottman, 1997). This can help researchers examine behavior patterns (e.g., Yang et al., 2015; Kucuk and Sisman, 2017; Wang et al., 2022) and understand how a given user might behave in a given context. Therefore, LSA was used in this study to understand how learners with different personalities might behave in an online course.

3. Method

3.1. Study context

Data from a three-months (the length of the semester) Basic Software (BS) course was used in this study. The course aims to help learners learn computer architecture and compilation, operating systems, and assembly language. It was chosen because it is part of the Computer Science curriculum at a public Tunisian university, and it was taught online. All the learners who participated in this experiment were already enrolled in the course. The online course system was run on the learning management system (LMS) Moodle, a free and open-source system.

Weekly learning materials in various forms, such as videos, texts, PowerPoint presentations, external links for online resources and mental break items (e.g., pictures) were uploaded by the teacher. For each course module, the learners had to read different learning materials uploaded by the teacher, as well as finish different assignments. These assignments are quizzes to be answered on Moodle or also a particular exercise that learners had to finish on a separate Word document, and then upload it in Moodle. They also had the chance to update their uploaded assignments, if needed (before the given deadline). The learners were rewarded with a digital badge for each completed course module. Additionally, they could freely use the course forum to post their questions and to communicate with their peers (i.e., not mandatory task). The teacher was more as a facilitator by grading and providing written feedback on the uploaded assignments online. She also encouraged and helped the learners by answering questions and joining their online forum discussions, when needed. Figure 1 presents the whole learning activities on Moodle.

A summary of each learner's course achievements was displayed on their profiles. Learners had also the possibility to see their detailed course achievement report (grades, completion rate, collected badges, etc.), as well as the course achievements of their peers. For instance, as shown in Figure 2A, the learners' profiles show a summary of their course achievements, such as their collected badges, which is visible to all learners. Additionally, the learners had the possibility to see the list of their course peers, and view their profiles (see Figure 2B).

FIGURE 1
Learning activities on Moodle.

Finally, the learners can see the status of their assignments (grades or feedback given by the teacher) (see Figure 2C).

3.2. Participants and instrument

Participants were 92 undergraduate learners (66% of them are males and 34% are females) majoring in computer science and aged between 18 and 23. At the beginning of the semester, the learners' personality traits were identified using the big five inventory (BFI). BFI is validated and widely used in the literature to identify individuals' personalities (John and Srivastava, 1999). It is a five-point Likert-type questionnaire, with answers ranging from 1 (strongly disagree) to 5 (strongly agree). It consists of 44 items which cover the five personality dimensions in the FFM, such as "I am someone who is helpful and unselfish with others" for the agreeableness dimension and "I am someone who is talkative" for the extraversion dimension.

At the end of the semester (after three months), the learners' log data was collected. To ensure an accurate analysis and findings with more representative behaviors, learners who dropped-out from the course ($n=27$) were excluded. Therefore, the study had 65 participants. Since there is no guidance in the BFI scoring for determining whether an individual has a high or low personality trait (e.g., high extraversion or low extraversion) (Codish and Ravid, 2014), the standard z-score was computed, as suggested in several studies (Bidjerano and Yun Dai, 2007; Codish and Ravid, 2014). It provides information on how far a data point is from the mean. In this context, learners with $z>0$ were considered as having a high value on the respective personality dimension (e.g., high in extraversion, high in openness, etc.), while learners with $z<0$ were considered to have a low value on the respective personality dimension (e.g., low in extraversion, low in openness, etc.). In this present study, no learners were found with $z=0$. Table 1 presents the mean and standard deviation of each personality dimension. Since the agreeableness personality dimension had an unbalanced number of learners (see Table 1), it was excluded from this study. Therefore, this study investigated the learners' navigational behavior patterns of the remaining four personality traits, namely extraversion, openness, neuroticism and conscientiousness.

In our study, for each personality trait, the reliability, mean and standard deviation were calculated. As shown in Table 1, the results yielded an alpha of 0.7 or higher, which means that all the personality traits produced acceptable reliabilities.

3.3. Data coding and analysis

The learners' navigational behaviors were automatically captured and stored by Moodle online. Specifically, after data cleaning, this study collected 15,869 log data from the 65 learners. These log data described 12 online learning behaviors (see Table 2), which are considered significant to the representation of learners' navigational behaviors on learning management systems (Wang, 2017; Tlili et al., 2019). The frequency distribution of each of these online learning behaviors according to the four personality dimensions is presented in Appendix.

To identify the navigational behavior patterns of each personality trait based on the learning behaviors described above, LSA was applied using GSEQ version 5.1 software (Bakeman and Quera, 2011). The motivation behind using LSA in this study is because it is a common statistical technique in behavioral science and well situated for analyzing the interaction data collected through log files (Pohl et al., 2016). LSA has been widely used to examine the behavioral

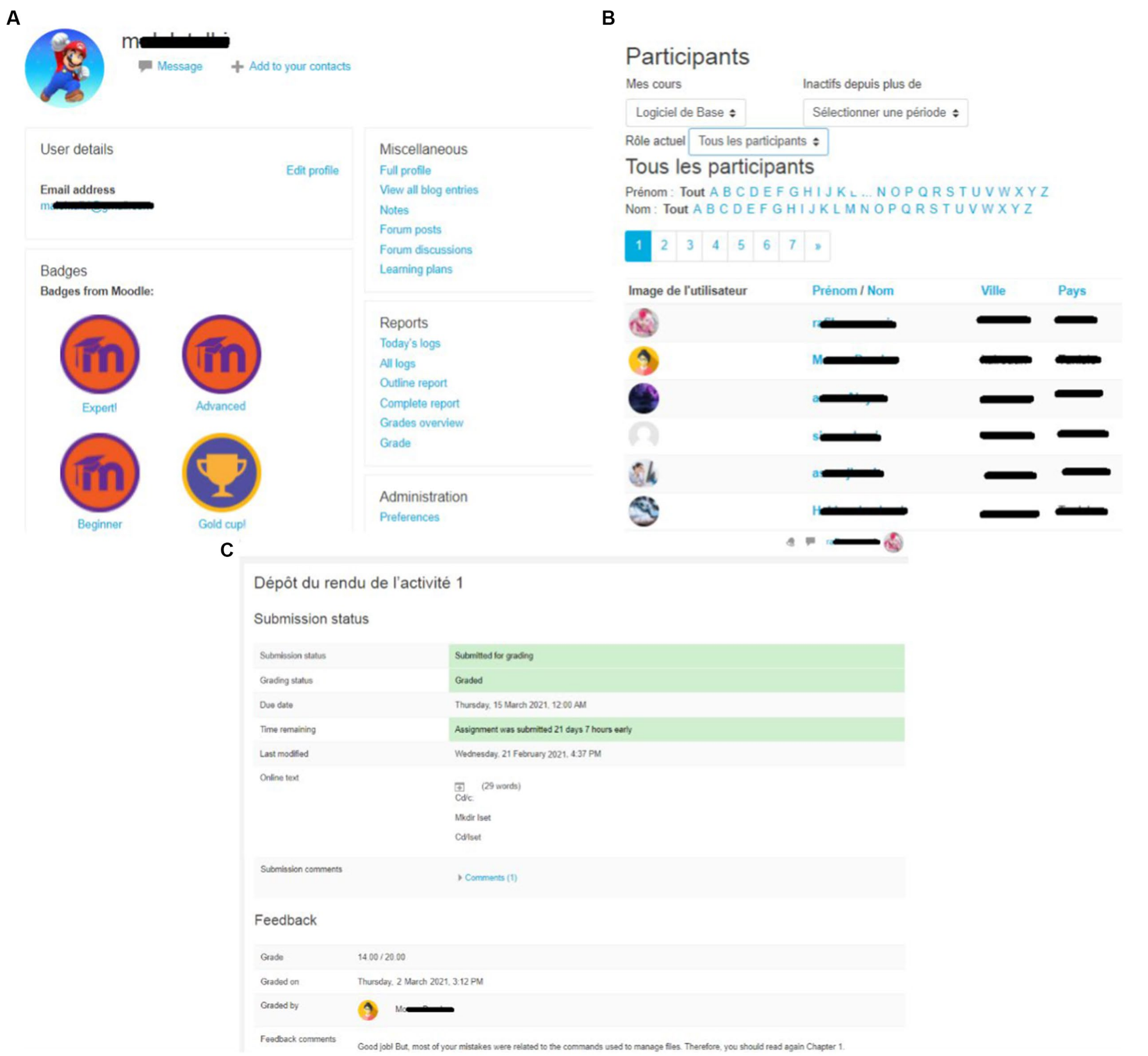


FIGURE 2
The Moodle system: **(A)** learner profile with a summary of their course achievements, **(B)** list of peers with the access of their profiles, and **(C)** examples of an assignment status.

TABLE 1 Distribution of personality traits.

Personality dimensions	Extraversion α =0.88		Agreeableness α =0.82		Conscientiousness α =0.84		Neuroticism α =0.78		Openness α =0.86	
Level	High	Low	High	Low	High	Low	High	Low	High	Low
Number of learners	35	30	46	19	31	34	31	34	35	30
Mean	3.75	2.79	4.27	3.37	4.17	3.18	3.22	2.06	4.05	3.34
SD	0.37	0.38	0.22	0.45	0.31	0.41	0.51	0.44	0.27	0.22

transitions which happen at frequencies greater than chance (Sackett, 1978). Subsequently, a series of behavioral transitions or navigational patterns (i.e., the order in which learners go through different activities) could be identified (Graf et al., 2010).

To conduct LSA, learners' behaviors were coded in the chronological order of their occurrences. For example, after logging into the system, a learner viewed a course module (CA1), uploaded an assignment (AS2), and then saw her achievements (AR1); this series of behaviors was thus coded as "CA1 AS2 AR1." The z-score value of each connection between each sequence was calculated to determine if that connection reached the statistical significance. Bakeman and Quera (1995) stated that a z-score greater than 1.96 indicates that a specific sequence has reached the level of significance ($p < 0.05$). Wampold (1992) further mentioned that, as cited in

TABLE 2 Coding of the learning behaviors.

Learning behaviors	Learning activity	Description	Code
Course activity (CA)	Course module viewed	A learner has viewed a particular course module	CA1
	Course module completion	A learner has completed a particular course module	CA2
Assignment submission (AS)	Assignment form viewed	A learner has viewed a particular submission form (i.e., assignment, deadline, and time line)	AS1
	Assignment uploaded	A learner has uploaded a particular assignment	AS2
	Assignment updated	A learner has updated a particular assignment	AS3
	Assignment status viewed	A learner has viewed the status of a particular assignment (i.e., grades or feedback given by the teacher)	AS4
Discussion (D)	Discussion viewed	A learner has viewed the discussion on the forum	D1
	Discussion made	A learner has been involved in the discussion on the forum (i.e., write a post or reply to a discussion)	D2
Achievement result (AR)	Personal achievement viewed	A learner has viewed his/her personal course achievements	AR1
	Peers' achievement viewed	A learner has viewed his/her peers' course achievements	AR2
Peers (P)	The list of peers viewed	A learner has viewed the list of his/her peers taking the course	P1
	Peers' profile viewed	A learner has viewed a specific profile of his/her peers	P2

McComas et al. (2009), a z-score does not indicate the degree to which a pattern is present. Based on this, this study therefore used the z-score in conjunction with a strength of association measurement, specifically Yule's Q (Pohl et al., 2016). Yule's Q is a transformation of the odds ratio to a $[-1 \ldots +1]$ range. Therefore, in this study, a transition from one code (behavior) to another was then only considered significant if the z-score was above the 1.96 level (the critical value assuming a normal distribution and a significance level of 0.05) and the Q-value was at least 0.30 (a moderate association).

The transitional probabilities, which is the conditional probability of a transition type (Cheng et al., 2017), were also calculated using GSEQ version 5.1. A transitional probability indicates the likelihood that an initial behavior follows a subsequent or same behavior.

Furthermore, prior to conducting LSA, a Pearson chi-square test was also conducted on the behavior frequency table of all the learners to determine if a significant dependence between rows and columns exist. Rows contain the initial behaviors, while columns contain the successive behaviors after conducting the initial ones.

4. Results and discussion

A behavior transition diagram was drawn for each personality trait (low and high), as shown in Figures 3–6, showing those sequences which reached a significant effect. Each transition in Figures 3–6 has both significance and probability values represented on each line as follow: *Significance (Probability)*. The effect size was highlighted based on the probability of each transition, where the thicker the lines, the higher the probability of each transition. Orange and green colors were used to highlight the unique navigational behavior patterns of learners with high and low levels in each personality trait, respectively. Black color shows that the pattern was significant for low and high levels of the respective personality trait.

4.1. Extraversion

The obtained chi-square test results confirmed that there is a significant relation between the rows and columns of the tallied frequencies ($\chi^2 = 7224.71$, $df = 121$, $p < 0.001$ for learners high in extraversion; $\chi^2 = 5726.87$, $df = 121$, $p < 0.001$ for learners low in extraversion). Figure 3 shows the navigational behavior patterns of learners high and low in extraversion.

As shown in Figure 3, learners high in extraversion had some unique navigational behavior patterns compared to those low in extraversion. For instance, they significantly navigated between viewing the course module and their personal achievements, where their grades, points and badges were displayed (CA1 → AR1). This could be because learners high in extraversion tend to be extrinsically motivated (Moldasheva and Mahmood, 2014), therefore they kept navigating to view their earned badges and points to motivate themselves. Furthermore, unlike learners low in extraversion, learners high in extraversion significantly navigated between seeing their peers' profiles and the list of all peers (P2 → P1), meaning that they are not just looking up one person and then moving somewhere else, but seem to be looking up the profiles of multiple other learners. This could be explained with people high in extraversion are socializers and want to know new persons (Zhang, 2002; Tlili et al., 2019). Therefore, learners high in extraversion used the course as a place to build friendships by seeing the list of peers as well as their associated profiles.

As shown in Figure 3, learners low in extraversion, on the other hand, did not significantly focus on their personal

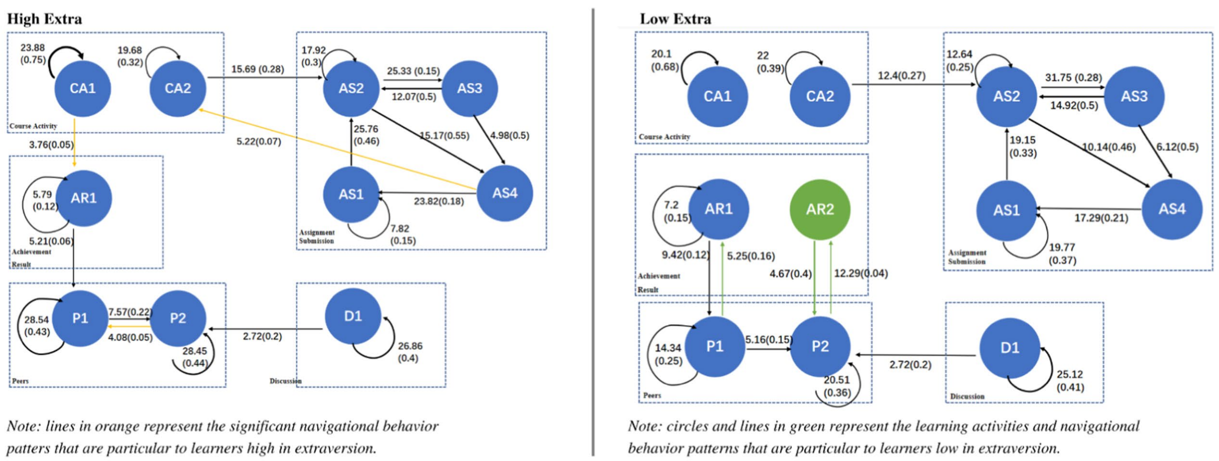


FIGURE 3
Navigational behavior patterns of learners high and low in extraversion.

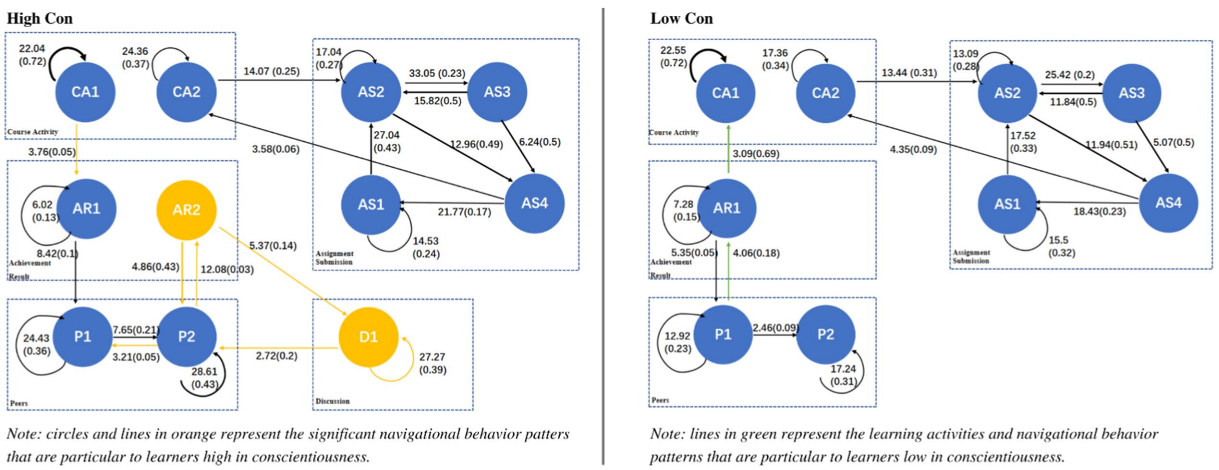


FIGURE 4
Navigational behavior patterns of learners high and low in conscientiousness.

achievements like their peers high in extraversion. However, they compared their achievements with that of their peers. Specifically, learners low in extraversion significantly navigated between their achievements and the list of peers, through which they could easily explore others' profiles and achievements (AR1 ↔ P1 then P1 → P2 followed by P2 ↔ AR2). These navigational behavior patterns could be because learners low in extraversion are high achievement-driven (Mcclelland et al., 1953; Farley, 1966).

Based on the obtained findings, it is possible to enhance the course navigation experience for learners based on their extraversion personality. For instance, it is possible to make the rewards (e.g., points or badges) earned by learners within the course more visible on their profiles for those high in extraversion since they are more extrinsically motived. On the other hand, to motivate learners low in extraversion, it is better to make their progress more visible compared to their peers, for instance, through the use of leaderboard and progress bar. Denden et al. (2021) confirmed that compared to learners high in extraversion, learners low in extraversion found progress bar more useful.

4.2. Conscientiousness

The obtained chi-square test results confirmed that there is a significant relation between the rows and columns of the tallied frequencies (for high conscientiousness, $\chi^2 = 7558.36$, $df = 121$, $p < 0.001$; for low conscientiousness, $\chi^2 = 5085.49$, $df = 121$, $p < 0.001$). Figure 4 shows the navigational behavior patterns of learners high and low in conscientiousness.

As shown in Figure 4, unlike learners low in conscientiousness, those high in conscientiousness significantly navigated between viewing the course module and their achievements (CA1 → AR1). They then navigated to see their peers' achievements (AR1 → P1; P1 → P2; P2 → AR2), where they kept checking their peers' profiles and achievements (AR2 ↔ P2). This could be explained with learners high in conscientiousness are achievement-driven (Tlili et al., 2019), they therefore kept checking their course achievements and comparing them to their peers' achievements. It is also found that learners high in conscientiousness significantly navigated to reading their peers' discussions after seeing their achievements (AR2 → D1), but they were

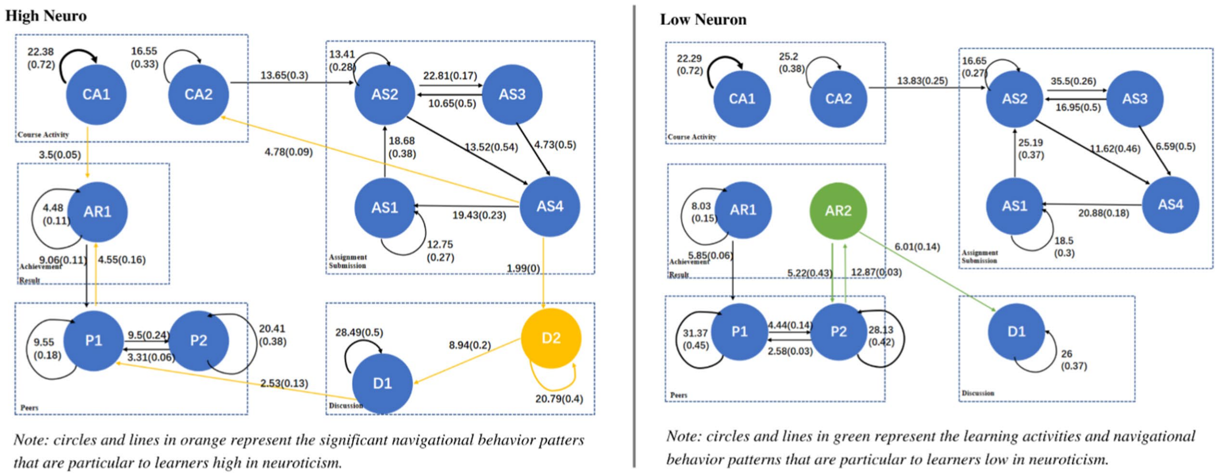


FIGURE 5
Navigational behavior patterns of learners high and low in neuroticism.

not involved in discussions (i.e., they did not write any comments, no D2). This could be explained with conscientious individuals are cautious about their public sayings in social networking websites, they therefore prefer sending private messages instead (Olson and Suls, 2000; Muscanell and Guadagno, 2012).

Interestingly, it is found that learners low in conscientiousness were also achievement-driven, however in a different way. Specifically, unlike learners high in conscientiousness, they mainly focused on their own achievements (AR1 → CA1) without going through others' achievements (i.e., no AR2) (see Figure 4).

Based on the obtained findings, it is possible to enhance the course navigation experience for learners with conscientiousness personality. For instance, it is recommended that the developed online course should not mainly use public communication channels, such as forums, but also provide private ones, where learners high in conscientiousness will be more willing to involve in online social activities within the course. For example, software designers could develop a forum function that allows learners to write private messages to the persons involved in the discussion. On the other hand, it is recommended to develop some functionalities, such as badges and points, to make the course achievements for learners low in conscientiousness more visible, hence be more engaged while learning.

It should be noted that both learners high in conscientiousness and low in conscientiousness had unique navigational patterns (P2 → P1 and D1 → P2 for learners high in conscientiousness; P1 → AR1 for learners low in conscientiousness) that no explanation was found and further investigations need to be done to explain them.

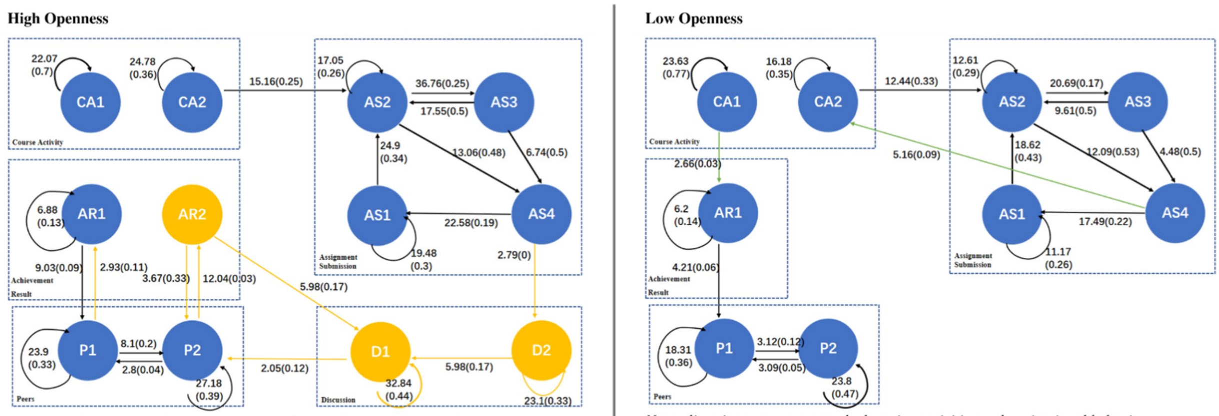
4.3. Neuroticism

The chi-square test results revealed that the relationship between rows and columns of the tallied frequency is significant ($\chi^2 = 4948.42$, $df = 121$, $p < 0.001$ for learners high in neuroticism; $\chi^2 = 7634.56$, $df = 121$, $p < 0.001$ for learners low in neuroticism). Figure 5 shows the navigational behavior patterns of learners high and low in neuroticism.

As shown in Figure 5, unlike learners low in neuroticism, learners high in neuroticism went directly to see their associated personal achievements after they viewed the course/ module (CA1 → AR1). They also went to see the list of their peers and peers' profiles where some of achievements had been displayed (P1 ↔ P2), and then went back to see their own achievements and compare (P1 → AR1). These two behavior sequences are explained with learners high in neuroticism have high level of anxiety and stress, and they always feel unsafe (Watjatrakul, 2016; Denden et al., 2021), and this was reflected in their behaviors where they significantly kept going back to see their course achievements and compared them to their peers, to know if they are doing well or not within the course. De Feyter et al. (2012) mentioned that neurotic learners cope with their anxiety about academic failure by intensifying their efforts in trying to prevent failure, which is seen in their online behavioral patterns of significantly checking their achievements and comparing them to their peers' achievements.

Additionally, Figure 5 shows that, unlike learners low in neuroticism, learners high in neuroticism involved in discussions within the course forum (D2). Specifically, they significantly involved in discussions after they saw their assignment status, where teachers' feedback and grades had been displayed (AS4 → D2) and they kept reading discussions (D2 → D1) and answering them (D2 ↔ D2). This could be explained with learners high in neuroticism used forums to communicate with their peers, as one of the ways to reduce their anxiety and stress, after seeing the feedback given by their teachers. Bhagat et al. (2019) also suggested that learners high in neuroticism should have the possibility to freely ask questions and talk with their peers to reduce their anxiety level and increase their chances of success.

On the other hand, Figure 5 shows that learners low in neuroticism, unlike learners high in neuroticism, significantly navigated between seeing their peers' achievements and then specifically seeing their peers' specific profiles (AR2 ↔ P2). This means that learners low in neuroticism visited the profiles of those they were interested in based their course achievements. Furthermore, they also went to the forum to view discussions possibly posted by their peers who they found interested in based on the course achievements (AR2 → D1). These behavioral sequences of keeping viewing peers' profiles, achievements and discussions could be explained with persons low in neuroticism tend to be more self-confident and want to build connections within their environments (Baluku et al., 2016).



Note: circles and lines in orange represent the significant navigational behavior patters that are particular to learners high in openness.

Note: lines in green represent the learning activities and navigational behavior patterns that are particular to learners low in openness.

FIGURE 6
Navigational behavior patterns of learners high and low in openness.

Moldasheva and Mahmood (2014) pointed out that learners high in neuroticism can get easily anxious. Therefore, to reduce their anxiety level within a given course and make them feel that they did well when learning online, it is possible to make the learning environment customizable where rewards, such as digital badges and points, can be configured to be visible for learners. Consequently, learners with high anxiety can see their course achievements and feel much better while learning. On the other hand, as learners low in neuroticism showed their willingness to discover people they were interested in, it would be helpful to develop and adopt more communication functionalities (i.e., public and private messaging) in the online learning environment, where learners are able to send messages privately to their peers to get to know them.

It should be noted that learners high in neuroticism had a unique navigational pattern, namely D1 → P1 that no explanation was found and further investigations need to be done to explain them.

4.4. Openness

The chi-square test results highlighted that there is a significant relationship between rows and columns of the behavior frequency tables of learners high ($\chi^2 = 8527.14$, $df = 121$, $p < 0.001$) or low ($\chi^2 = 3556.29$, $df = 121$, $p < 0.001$) in openness. Figure 6 shows the navigational behavior patterns of learners high and low in openness.

As shown in Figure 6, unlike those low in openness, learners high in openness conducted a lot of peers-related activities. Specifically, they were highly involved in discussions. For example, they engaged in discussions within the course forum after viewing their assignment status, where the grades and teachers' feedback had been displayed (AS4 → D2). This behavior pattern could be explained by their willingness to know others' opinions/experiences after reading the feedback given by the teacher. Furthermore, learners high in openness, unlike those low in openness, not only kept posting discussions (D2 → D2) but also continuously kept reading them (D2 → D1 and D1 → D1). These behavior sequences could possibly be the result of their inclination to exchange opinions with others. In this context, several studies pointed out that learners high in openness tend to be curious, imaginative, and creative. Moreover, they continuously seek out new experiences and are willing to take part in peer learning (Moldasheva and Mahmood, 2014; Tlili et al., 2019). The obtained findings about the discussion behavior patterns of learners high in openness are in line with previous research which indicates that openness is a significant predictor for the use of social networking sites (Banczyk et al., 2008; Krishnan and Atkin, 2014; Huang, 2019).

Furthermore, it is seen that learners high in openness checked their peers' profiles and then saw their peers' achievements following they viewed discussions (P2 ↔ AR2 and AR2 → D1). This is possibly because people high in openness are more likely to form a larger friendship network (Lang et al., 1998; Zhu et al., 2013). Specifically, they could consider the course forum as a place to meet new friends and further explored them by their profiles and achievements. Besides, it is interesting to see that after viewing peers' achievement, learners high in openness could also go back to review the comments posted by the persons they were interested in (AR2 → D1).

Learners low in openness, on the other hand, were less interested in others' experiences and mainly focused on their own experiences (see Figure 6). This can be seen when they significantly navigated between the course and their own achievements (CA1 → AR1). Interestingly, it can be seen that no significant behavioral pattern was found related to peers' achievements (AR2) or discussion involvement (D1 or D2), unlike learners high in openness. This confirms that learners low in openness were not much interested in their peers' experiences or achievements.

Based on the obtained findings, since learners high in openness are highly engaged in communication with others, it is recommended to provide them a learning environment which fosters various communication channels, such as forum or instant messages, to keep their learning engagement high. On the other hand, learners low in openness are less likely to get involved in discussions even after viewing their peers' posts in the course forum, and they were more interested in their own course achievements. Therefore, it is recommended that learning environments should be designed with functionalities (e.g., dashboards or reports) that automatically generate the progress of learners in a given course to help them keep up with their progress and achievements.

5. Conclusions, implications and future directions

This study analyzed the impact of learners' personality differences on their online learning navigation behavior patterns along four dimensions, namely: extraversion, conscientiousness, neuroticism, and openness. To the best of our knowledge, this study is one of the earliest studies that attempted to investigate how learners' personalities might impact their navigational behavior patterns in an online course. The findings revealed that the learners' extraversion level was found to affect their course navigation experience. Learners high in extraversion were more likely to have navigational behavior patterns, which were primarily influenced by extrinsic motivation. They navigated significantly between viewing peer profiles and all peer lists due to a greater desire to meet new people; conversely, learners low in extraversion navigated significantly between their achievements and peer lists due to their high achievement driven. The study also found that high levels of conscientiousness influenced learners' online navigation behavior patterns due to being driven by achievement. Learners high in conscientiousness navigate significantly between viewing course modules and their achievements, and they tend to view their peers' achievements and thus check their peers' profiles and achievements; in contrast, learners low in conscientiousness were more likely to focus on their own achievements. Related to the neuroticism dimension, our study showed that learners with high neuroticism are usually accompanied by high levels of anxiety and stress, and they tend to look at their personal achievements directly after viewing the course/module, and they also compared their profiles and achievements with those of their peers; on the contrary, learners with low neuroticism are usually accompanied by self-confidence, and they tended to visit the profiles of people they are interested in based on their course grades. In addition, this present study showed that the level of openness also influences learners' performance in peer-related activities. Those with high openness are highly engaged in discussions and learning and they expect to learn with their peers; conversely, those with low openness are focused on their own experiences and are not significant in terms of peer achievement and discussion participation.

This study can contribute to the literature in several ways. From a theoretical perspective, it can contribute to the educational psychology field by providing empirical evidences on how learners with different personalities tend to behave in an online course, hence better understand each personality trait and its related features in education. From a practical perspective, this study can contribute to the human-computer interaction and adaptivity fields by providing to various stakeholders (e.g., educators, designers, psychologists) several recommendations to enhance online course design and adaptivity with respect to each learner's personality. It can also contribute to the learner modeling field by providing different behavioral patterns that could help to identify learners' personalities. Finally, several studies have been conducted to automatically identify the learners' personalities from their behaviors in an online course, these studies, however, focused on single behaviors, such as frequency of visiting a course or forum (Denden et al., 2018; Tlili et al., 2019). Therefore, the use of navigation behavioral patterns, from this present study, could be another data source for getting information about learners, hence making the automatic identification of their personalities more accurate.

It should be noted that this study has several limitations that should be acknowledged and further researched. For instance, the agreeableness personality trait was not investigated. Additionally, this study mainly investigated single personalities (e.g., learners high in openness vs. learners low in openness) and did not investigate combined personalities (e.g., learners high in openness and extraversion vs. learners low in openness and extraversion), which might reveal more information on learners' online learning behaviors based on their personalities. Future research directions could focus on investigating these limitations, as well as designing an adaptive learning system based on personality, which takes into consideration the obtained findings of this study.

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 Smart Learning Institute of Beijing Normal University, China. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1071985/full#supplementary-material

References

Ackerman, C. E. (2020). Big Fiver personality traits: the OCEAN model explained. Available at: https://positivepsychologyprogram.com/big- five-personality-theory/ (Accessed November 07, 2020).

Al-Dujaily, A., Kim, J., and Ryu, H. (2013). Am I extravert or introvert? Considering the personality effect toward e-learning system. *Educ. Technol. Soc.* 16, 14–27.

Arockiam, L., and Selvaraj, J. C. (2013). User interface design for effective e-learning based on personality traits. *Int. J. Comput. Appl.* 61, 28–32. doi: 10.5120/9998-4861

Bakeman, R., and Gottman, J. M. (1997). *Observing interaction: an introduction to sequential analysis*. New York: Cambridge University Press.

Bakeman, R., and Quera, V. (1995). *Analyzing interaction: sequential analysis with SDIS and GSEQ*. UK: Cambridge University Press.

Bakeman, R., and Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge, UK: Cambridge University Press

Baluku, M. M., Kikooma, J. F., and Kibanja, G. M. (2016). Does personality of owners of micro enterprises matter for the relationship between startup capital and entrepreneurial success? *Afr. J. Bus. Manag.* 10, 13–23. doi: 10.5897/AJBM2015.7738

Banczyk, B., Krämer, N. C., and Senokozlieva, M. (2008). The wurst "meets, fatless" in MySpace. The relationship between personality, nationality and self- presentation in an online community. In Paper Presented at the Annual Conference of the International Communication Association, Montreal, Canada.

Barrick, M., and Mount, M. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *J. Appl. Psychol.* 81, 261–272. doi: 10.1037/0021-9010.81.3.261

Barrick, M. R., Mount, M. K., and Gupta, R. (2003). Meta-analysis of the relationship between the five-factor model of personality and Holland's occupational types. *Pers. Psychol.* 56, 45–74. doi: 10.1111/j.1744-6570.2003.tb00143.x

Bhagat, K. K., Wu, L. Y., and Chang, C. Y. (2019). The impact of personality on students' perceptions towards online learning. *Australas. J. Educ. Technol.* 35, 98–108. doi: 10.14742/ajet.4162

Bidjerano, T., and Yun Dai, D. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learn. Individ. Differ.* 17, 69–81. doi: 10.1016/j.lindif.2007.02.001

Boeree, G. C. D. (2006). *Personality theories.* USA: Psychology Department Shippensburg University.

Brusilovsky, P. (2001). Adaptive hypermedia. *User Model. User-Adap. Inter.* 11, 87–110. doi: 10.1023/A:1011143116306

Chai, H., Hu, T., and Niu, G. (2022). How proactive personality promotes online learning performance? Mediating role of multidimensional learning engagement. *Educ. Inf. Technol.* 28, 4795–4817. doi: 10.1007/s10639-022-11319-7

Chen, B., Resendes, M., Chai, C. S., and Hong, H. Y. (2017). Two tales of time: uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interact. Learn. Environ.* 25, 162–175. doi: 10.1080/10494820.2016.1276081

Cheng, H. N., Liu, Z., Sun, J., Liu, S., and Yang, Z. (2017). Unfolding online learning behavioral patterns and their temporal changes of college students in SPOCs. *Interact. Learn. Environ.* 25, 176–188. doi: 10.1080/10494820.2016.1276082

Cheng, Y., Wang, Y., Cheng, I., and Chen, N. (2019). An in-depth analysis of the interaction transitions in a collaborative augmented reality-based mathematic game. *Interact. Learn. Environ.* 27, 782–796. doi: 10.1080/10494820.2019.1610448

Child, I. L. (1968). "Personality in culture" in *Handbook of personality theory and research*. eds. E. Borgatta and W. W. Lambert (Chicago, IL: Rand McNally), 80–101.

Clark, M. H., and Schroth, C. A. (2010). Examining relationships between academic motivation and personality among college students. *Learn. Individ. Differ.* 20, 19–24. doi: 10.1016/j.lindif.2009.10.002

Codish, D., and Ravid, G. (2014). Personality based gamification—educational gamification for extroverts and introverts. In Proceedings of the 9th Chais Conference for the Study of Innovation and Learning Technologies: Learning in the Technological Era. 36–44).

Costa, P. T. Jr, and McCrae, R. R. (2009). "The Five-Factor Model and the NEO Inventories," in Oxford handbook of personality assessment. Ed. J. N. Butcher (United States: Oxford University Press), 299–322.

De Feyter, T., Caers, R., Vigna, C., and Berings, D. (2012). Unraveling the impact of the Big Five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learn. Individ. Differ.* 22, 439–448. doi: 10.1016/j.lindif.2012.03.013

Denden, M., Tlili, A., Essalmi, F., and Jemni, M. (2018). Implicit modeling of learners' personalities in a game-based learning environment using their gaming behaviors. *Smart Learn. Environ.* 5, 1–19. doi: 10.1186/s40561-018-0078-6

Denden, M., Tlili, A., Essalmi, F., Jemni, M., Chen, N., and Burgos, D. (2021). Effects of gender and personality differences on students' perceptions of game design elements in educational gamification. *Int. J. Hum. Comput. Stud.* 154:102674. doi: 10.1016/j.ijhcs.2021.102674

DeYoung, C. G., Quilty, L. C., and Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *J. Pers. Soc. Psychol.* 93, 880–896. doi: 10.1037/0022-3514.93.5.880

Diseth, A. (2003). Personality and approaches to learning as predictors of academic achievement. *Eur. J. Personal.* 17, 143–155. doi: 10.1002/per.469

Essalmi, F., Tlili, A., Ayed, L. J. B., and Jemni, M. (2017). Toward modeling the learner's personality using educational games. *Int. J. Distance Educ. Technol.* 15, 21–38. doi: 10.4018/IJDET.2017100102

Farley, F. H. (1966). Introversion and achievement motivation. *Psychol. Rep.* 19:112. doi: 10.2466/pr0.1966.19.1.112

Fatahi, S. (2019). An experimental study on an adaptive e-learning environment based on learner's personality and emotion. *Educ. Inf. Technol.* 24, 2225–2241. doi: 10.1007/s10639-019-09868-5

Franić, S., Borsboom, D., Dolan, C. V., and Boomsma, D. I. (2014). The Big Five personality traits: psychological entities or statistical constructs? *Behav. Genet.* 44, 591–604. doi: 10.1007/s10519-013-9625-7

Furnham, A. (1992). Personality and learning style: a study of three instruments. *Personal. Individ. Differ.* 13, 429–438. doi: 10.1016/0191-8869(92)90071-V

Graf, S., Liu, T. C., and Kinshuk, (2010). Analysis of learners' navigational behaviour and their learning styles in an online course. *J. Comput. Assist. Learn.* 26, 116–131. doi: 10.1111/j.1365-2729.2009.00336.x

Gurven, M., Rueden, C., and Massenkoff, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager-farmers in the Bolivian Amazon. *J. Pers. Soc. Psychol.* 104, 354–370. doi: 10.1037/a0030841

Gustavsson, J. P., Jönsson, E. G., Linder, J., and Weinryb, R. M. (2003). The HP5 inventory: definition and assessment of five health-relevant personality traits from a five-factor model perspective. *Personal. Individ. Differ.* 35, 69–89. doi: 10.1016/S0191-8869(02)00142-3

Haron, H., and Sahar, S. (2010). An investigation on predictors of e-learning adoption among Malaysian e-learners. In. International Conference on Science and Social Research (CSSR), 927–932.

Harrington, R., and Loffredo, D. A. (2010). MBTI personality type and other factors that relate to preference for online versus face-to-face instruction. *Internet High. Educ.* 13, 89–95. doi: 10.1016/j.iheduc.2009.11.006

Hong, J., Lee, Y., and Ye, J. (2021). Procrastination predicts online self-regulated learning and online learning ineffectiveness during the coronavirus lockdown. *Personal. Individ. Differ.* 174:110673. doi: 10.1016/j.paid.2021.110673

Huang, C. (2019). Social network site use and Big Five personality traits: a meta-analysis. *Comput. Hum. Behav.* 97, 280–290. doi: 10.1016/j.chb.2019.03.009

Irani, T., Telg, R., Scherler, C., and Harrington, M. (2003). Personality type and its relationship to distance education students' course perceptions and performance. *Q. Rev. Distance Educ.* 4:445.

John, O. P., Robins, R. W., and Pervin, L. A. (2008). *Handbook of personality: theory and research*. New York, NY: The Guilford Press.

John, O. P., and Srivastava, S. (1999). "The big-five trait taxonomy: history, measurement, and theoretical perspectives" in *Handbook of personality: theory and research*. eds. L. A. Pervin and O. P. John (New York: Guilford Press), 102–138.

Jonassen, D. H., and Grabowski, B. L. (1993). "Handbook of individual differences" in *Learning & Instruction* (Hillsdale, NJ: Lawrence Erlbaum Associates)

Krishnan, A., and Atkin, D. (2014). Individual differences in social networking site users: the interplay between antecedents and consequential effect on level of activity. *Comput. Hum. Behav.* 40, 111–118. doi: 10.1016/j.chb.2014.07.045

Kucuk, S., and Sisman, B. (2017). Behavioral patterns of elementary students and teachers in one-to-one robotics instruction. *Comput. Educ.* 111, 31–43. doi: 10.1016/j.compedu.2017.04.002

Lai, S., Sun, B., Wu, F., and Xiao, R. (2019). Automatic personality identification using students' online learning behavior. *IEEE Trans. Learn. Technol.* 13, 26–37. doi: 10.1109/TLT.2019.2924223

Lang, F. R., Staudinger, U. M., and Carstensen, L. L. (1998). Perspectives on socioemotional selectivity in late life: How personality and social context do (and do not) make a difference. *J. Gerontol. B* 53B, 21–30. doi: 10.1093/geronb/53B.1.P21

Lee, J. M., and Lee, Y. (2006). Personality types and learners' interaction in web-based threaded discussion. *Q. Rev. Dist. Learn.* 7, 83–94.

Mcclelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McComas, J., Moore, T., Dahl, N., Hartman, E., Hoch, J., and Symons, F. (2009). Calculating contingencies in natural environments: issues in the application of sequential analysis. *J. Appl. Behav. Anal.* 42, 413–423. doi: 10.1901/jaba.2009.42-413

McCrae, R. R., and John, O. P. (1992). An introduction to the five-factor model and its applications. *Personality* 60, 175–215. doi: 10.1111/j.1467-6494.1992.tb00970.x

Moldasheva, G., and Mahmood, M. (2014). Personality, learning strategies, and academic performance: evidence from post-Soviet Kazakhstan. *Educ. Train.* 56, 343–359. doi: 10.1108/ET-10-2012-0101

Murphy, S. A., Fisher, P. A., and Robie, C. (2021). International comparison of gender differences in the five-factor model of personality: an investigation across 105 countries. *J. Res. Pers.* 90:104047. doi: 10.1016/j.jrp.2020.104047

Muscanell, N. L., and Guadagno, R. E. (2012). Make new friends or keep the old: gender and personality differences in social networking use. *Comput. Hum. Behav.* 28, 107–112. doi: 10.1016/j.chb.2011.08.016

Myers, I. B., McCaulley, M. H., Quenk, N. L., and Hammer, A. L. (1998). *MBTI manual: a guide to the development and use of the Myers–Briggs type indicator* (*3*). Palo Alto: Consulting Psychologists Press.

Novikova, I. A., and Vorobyeva, A. A. (2019). "The five-factor model: contemporary personality theory," in *Cross-cultural psychology: contemporary themes and perspectives*. Ed. K. D. Keith, United States, 685–706.

Olson, B. D., and Suls, J. (2000). Self-, other-, and ideal-judgments of risk and caution as a function of the five-factor model of personality. *Personal. Individ. Differ.* 28, 425–436. doi: 10.1016/S0191-8869(99)00105-1

Pohl, M., Wallner, G., and Kriglstein, S. (2016). Using lag-sequential analysis for understanding interaction sequences in visualizations. *Int. J. Hum. Comput. Stud.* 96, 54–66. doi: 10.1016/j.ijhcs.2016.07.006

Sackett, G. P. (1978). *Observing behavior: I. Theory and applications in mental retardation*. USA: University Park.

Shang, J., Xiao, R., and Zhang, Y. (2020). A sequential analysis on the online learning behaviors of Chinese adult learners: take the KGC learning platform as an example. In International Conference on Blended Learning. 61–76). Thailand: Springer.

Sun, Z., Lin, C. H., Lv, K., and Song, J. (2021). Knowledge-construction behaviors in a mobile learning environment: a lag-sequential analysis of group differences. *Educ. Technol. Res. Dev.* 69, 533–551. doi: 10.1007/s11423-021-09938-x

Sun, T., Yu, Q., Guo, J., Ma, Y., and Zhang, Y. (2020). A preliminary study of online learning: the influence of the class approaches and the personality of students. Advances in Social Science, Education and Humanities Research, 496, 382–391.

Swanberg, A. B., and Martinsen, L. (2010). Personality, approaches to learning and achievement. *Educ. Psychol.* 30, 75–88. doi: 10.1080/01443410903410474

Tlili, A., Denden, M., Essalmi, F., Jemni, M., Chang, M., Kinshuk, , et al. (2019). Automatic modeling learner's personality using learning analytics approach in an intelligent Moodle learning platform. *Interact. Learn. Environ.* doi: 10.1080/10494820.2019.1636084

Tlili, A., Essalmi, F., Jemni, M., and Chen, N. S. (2016). Role of personality in computer based learning. *Comput. Hum. Behav.* 64, 805–813. doi: 10.1016/j.chb.2016.07.043

Tlili, A., Wang, H., Gao, B., Shi, Y., Zhiying, N., Looi, C. K., et al. (2021). Impact of cultural diversity on students' learning behavioral patterns in open and online courses: A lag sequential analysis approach. *Interact. Learn. Environ.* 1–20. doi: 10.1080/10494820.2021.1946565

Trull, T. J., and Sher, K. J. (1994). Relationship between the five-factor model of personality and Axis I disorders in a nonclinical sample. *J. Abnorm. Psychol.* 103:350. doi: 10.1037/0021-843X.103.2.350

Wampold, B. (1992). "The intensive examination of social interactions" in *Single-Case Research Design and Analysis: New Directions for Psychology and Education*. eds. T. R. Kratochwill and J. R. Levin (Hillsdale, NJ: Lawrence Erlbaum Associates), 93–131.

Wang, F. H. (2017). An exploration of online behavioral engagement and achievement in flipped classroom supported by learning management system. *Comput. Educ.* 114, 79–91. doi: 10.1016/j.compedu.2017.06.012

Wang, C., Shannon, D. M., and Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Educ.* 34, 302–323. doi: 10.1080/01587919.2013.835779

Wang, H., Tlili, A., Lämsä, J., Cai, Z., Zhong, X., and Huang, R. (2022). Temporal perspective on the gender-related differences in online learning behaviour. *Behav. Inform. Technol.* 1–15. doi: 10.1080/0144929X.2022.2039769

Watjatrakul, B. (2016). Online learning adoption: effects of neuroticism, openness to experience, and perceived values. *Interact. Technol. Smart Educ.* 13, 229–243. doi: 10.1108/ITSE-06-2016-0017

Wengrowicz, N., Swart, W., Paul, R., Macleod, K., Dori, D., and Dori, Y. J. (2018). Students' collaborative learning attitudes and their satisfaction with online collaborative case-based courses. *Am. J. Dist. Educ.* 32, 283–300. doi: 10.1080/08923647.2018.1511509

Wu, S., and Hou, H. (2015). How cognitive styles affect the learning behaviors of online problem-solving based discussion activity: A lag sequential analysis. *J. Educ. Comput.* 52, 277–298. doi: 10.1177/0735633115571307

Yang, X., Li, J., Guo, X., and Li, X. (2015). Group interactive network and behavioral patterns in online English-to-Chinese cooperative translation activity. *Internet High. Educ.* 25, 28–36. doi: 10.1016/j.iheduc.2014.12.003

Yang, X., Wang, H., and Li, J. (2016). The application of lag sequential analysis method in analyzing learning behavior. *China Educ. Technol.* 2, 17–23.

Yu, Z. (2021). The effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic. *Int. J. Educ. Technol. High. Educ.* 18:14. doi: 10.1186/s41239-021-00252-3

Zhang, L. (2002). Thinking styles and the big five personality traits. *Educ. Psychol.* 22, 17–31. doi: 10.1080/01443410120101224

Zhang, L. (2003). Does the big five predict learning approaches? *Personal. Individ. Differ.* 34, 1431–1446. doi: 10.1016/S0191-8869(02)00125-3

Zhang, S., Liu, Q., Chen, W., Wang, Q., and Huang, Z. (2017). Interactive networks and social knowledge construction behavioral patterns in primary school teachers' online collaborative learning activities. *Comput. Educ.* 104, 1–17. doi: 10.1016/j.compedu.2016.10.011

Zhu, X., Woo, S. E., Porter, C., and Brzezinski, M. (2013). Pathways to happiness: from personality to social networks and perceived support. *Soc. Networks* 35, 382–393. doi: 10.1016/j.socnet.2013.04.005

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 →