IRRODL

THE INTERNATIONAL REVIEW OF RESEARCH IN OPEN AND DISTRIBUTED LEARNING

Vol 26 No 4

November 2025

**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Editorial – Volume 26, Issue 4

**Adnan Qayyum, PhD**
*Associate Editor, Athabasca University*

## Do Open and Distance Learning institutions need more (or better supported) centres for teaching and learning?

The scope of research among the articles in this issue is broad. The article topics range from Arabic nursing students using artificial intelligence (AI) to teaching opera online in China to using mobile learning to teach mathematics. What they share is a common interest in initiating or understanding the use of open, digital, or distance education to support or improve learning.

**Aydemir and Kır** researched AI-generated tests. The study assessed how learners and experts responded to AI-generated multiple-choice questions, fill-in-the-blank exercises, and true-false activities in an online course. Students were more likely to view and then complete instructor-created content than AI-generated content. Students seemed to have more superficial engagement with AI-generated content, as they struggled to clearly understand instructions from AI. Therefore, the authors strongly advocate for the role of human oversight with AI content.

As AI becomes used in various educational courses and programs, it is important to assess its uptake and use. **Alenazi** studied the psychometric properties of an Arabic version of the well-known UTAUT model of technology adoption (Unified Theory of Acceptance and Use of Technology). Was the model useful for assessing nursing students' use of AI? The analysis supported the construct validity of the nine UTAUT constructs: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, habit, behavioral intention, and use behavior. The results support the validity of the UTAUT model in the Arabic cultural context. The study also offers insights into the factors influencing nursing students' acceptance and use of AI in healthcare education.

**Shen, Chang, & Yang** discuss learning effectiveness and deep learning in blended learning formats. Poor integration of online and in-person learning can lead to superficial student engagement. Using the Community of Inquiry-based blended learning model, the authors conducted a quasi-experimental design to study learning effectiveness and deep learning. They found that the blended learning group demonstrated superior learning effectiveness and reported higher deep learning perceptions compared to the traditional learning group. The results are useful, especially to educators who want to design blended learning that fosters deep learning.

**Wang, Zhan & Song** researched HyFlex courses, courses where students can choose to attend classes in person, participate synchronously online, or engage asynchronously through recorded materials. Hyflex allows students to combine multiple modes to tailor their learning experience. Though the multiple modes could pose challenges for team collaboration, the study found that effective communication significantly enhanced team performance. Creating a positive team atmosphere moderated this relationship. But an overly positive atmosphere may hinder constructive critique and diminish performance. They suggest that educators foster communication strategies that encourage open dialogue and critical thinking, while maintaining a supportive yet not overly positive team environment**.**

**Islam and Mahmud** were interested in the experience of important, often understudied educators in ODL: tutors. Over 80 tutors at eight tutorial centres of the Bangladesh Open University were asked to identify areas requiring improvement and provide suggestions for enhancing the academic quality of the programs. Major findings from the study include: the importance of tutor professional development; offering and requiring attendance of tutorial sessions to develop pedagogical competence; ensuring curriculum and modules are revised and updated; developing dedicated physical resources for students (computer labs, library and multimedia support); the crucial role of coordinators as essential in aligning tutorial schedules, resolving administrative challenges, and serving as communication bridges between institutional leadership and ground-level tutors. By hearing the voices of people doing the work, this study contributes to a more inclusive model of academic development.

**Li** writes about teaching opera online. Li compared the strengths of specific master classes at Juilliard Extension and the Royal College of Music, vs online seminars at Living Opera and Angel's Music Academy. The study finds that individuals learning to sing opera through master classes reported experiencing improvements in memory and problem-solving skills. Students who attended the online seminars reported improvements in concentration and memory. The study shows how ODL can be used for learning a broad range of skills, including vocal skills.

**Meylani** conducted a systematic review of how mobile learning in math affected students' learning attitudes, motivation, and performance. The research found that mobile learning enhances students' essential thinking skills and higher-order thinking skills. Mobile fostered interactivity and inquiry-based learning. Uneven access to resources and technology was still a barrier. Importantly, Meylani contends institutions should try to improve digital competency, engage faculty members in collaborative platforms to enhance their experience with mobile platforms and ensure data protection.

**ElGamal and Zawacki-Richter** conducted an umbrella review of flipped learning in higher education. An umbrella review is a second-order systematic review: it consolidates systematic reviews. They analyze 23 systematic reviews on flipped teaching and learning in open, distance and digital education. Their synthesis reveals several themes, including: how effective are flipped learning and teaching interventions, how flipped classrooms have been applied in various fields, learning design considerations, a scarcity of theoretical frameworks, and limited research on the pedagogical challenges of flipped learning models.

**Martin** reviews the timely book Resisting the Dehumanization of Refugees. Written by Abu Laban, Frishkopf, Hasmath, and Kirova, the book explores the struggle refugees face in creating a sense of

belonging during a time of immense backlash against refugees and immigration. As Martin points out, the thoughtful chapters are important for educators who can either exacerbate dehumanization or mitigate this.

**Boulhrir** reviews the book Brave New Worlds by Salman Khan. This book is aimed at parents, teachers or education administrators who are new to AI. It is unsurprising that, as the CEO of the eponymous Khan Academy, Khan is very enthusiastic about AI in education. However, as Boulhrir states, Khan's enthusiasm for AI technology in education outpaces a deeper engagement with its long-term social and pedagogical implications.

The world is changing quickly, creating much uncertainty. Open and Distributed Learning is changing quickly as well. Without a doubt, AI is the elephant in the room for ODL and all education. We know the pace is accelerating for using AI and other technologies in learning and teaching. There are many other important (non-AI) ODL practices that are ongoing, from hyflex to online opera. This brings many challenges for planning, designing, and supporting ODL on the ground.

As learning models and learning become more flexible, there is a demand for more, not less, human interaction and sophisticated instructional design. As the Shen article on blended learning suggests, the response to "superficial student engagement" is not in the *blend*, but in the *design*.

Perhaps more or better-supported teaching and learning centres can act as intermediaries. They can allow for the sandboxing of ODL initiatives on the ground, whether it be AI, Hyflex, online opera, or math on mobile. Education leaders may not necessarily know what to do with some of these initiatives. Perhaps leadership in ODL institutions do not really know what is happening with AI and other ed tech initiatives. And maybe they don't need to know all of it. Teaching and learning centres can be resources and repositories. They can help support the work of students, tutors, teachers, and perhaps even researchers.

AI transparency statement: No GenAI apps were used for writing this editorial.


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Assessment of Learner Engagement and Expert Evaluations of AI-Generated Versus Human-Created Interactive Content in an Online Course

**Hamza Aydemir[1] and Şeyda Kır[2]**
*[1]Faculty of Communication, Kahramanmaraş İstiklal University, Kahramanmaraş, Türkiye; [2]Technology Transfer Office, Yozgat Bozok University, Yozgat, Türkiye*

## Abstract

Generative artificial intelligence (GenAI) has introduced a novel aspect to educational methodologies and sparked fresh dialogues regarding the creation and evaluation of instructional resources. This project seeks to investigate the impact of GenAI on the development and assessment of online course materials and learners' engagement with these materials in the online learning environment. The study analyzed GenAI-generated multiple-choice questions, fill-in-the-blank exercises, and true-false activities during 3 weeks of a 14-week online course. Subject matter experts assessed these documents in regards to content, relevance, and clarity. Data was collected through an online form with open-ended questions. The interactions of learners with the GenAI-created learning activities were analyzed using log records of the learning management system and compared to the content provided by the course instructor regarding interaction levels. The study's conclusions elucidate the capability of GenAI technologies to produce course-specific content and their efficacy in education. We stress that human specialists' critical evaluations play a crucial part in improving the pedagogical validity of GenAI-powered learning materials. Further research into topics including the ethical dimension, the effect on academic achievement, and student motivation is recommended.

*Keywords*: generative AI, online education, learner engagement, higher education, case study

# Introduction

Education as well as the economic, social, political, and cultural spheres of life are all impacted in different manners by artificial intelligence (AI) technologies. AI tools are becoming increasingly prevalent in education for design, production, distribution, and access to learning, and the potential in these areas continues to grow dramatically. The use of chatbots in education, learning analytics, intelligent tutoring systems, virtual learning assistants, and applications offering personalized learning experiences are just a few examples (Aydemir & Seferoğlu, 2024). Additionally, it is widely acknowledged that AI technologies are frequently used to assist educators, administrators, and learners in various ways, especially in settings that enable face-to-face or online instruction (Kır & Şenocak, 2022). The advancement of AI technologies to generative AI (GenAI), as well as the gradual integration of these tools into people's daily lives, has brought an unanticipated dimension to the applications and projects planned for the very near future. Reflections on GenAI in the field of education have grown to be a hot subject in the past few months, and this prominence has contributed to the widespread implementation of many kinds of AI technologies in educational environments. It is projected that enhanced development and organization using GenAI, in addition to the development of the capacity it provides, will usher in an era of innovation in instructional design (Bozkurt, 2023; Haleem et al., 2022; Kasneci et al., 2023).

As GenAI technologies evolve, they can enhance educational practices by providing personalized learning experiences, facilitating content creation, and promoting innovative teaching strategies. However, the integration of these technologies raises ethical concerns and requires careful consideration of their impact on traditional educational paradigms. GenAI can significantly transform instructional design by enabling the rapid creation of educational materials. For example, AI tools can generate quizzes, lesson plans, and multimedia content based on specific learning objectives, thus allowing educators to focus more on pedagogical strategies rather than content creation (DaCosta & Kinsell, 2024; Wood & Moss, 2024). However, productive AI use in education is not without its challenges. A key concern is the potential for bias towards AI-generated content (Ferrara, 2023). It is crucial for educators to critically evaluate the outputs of AI systems to ensure they align with ethical standards and promote inclusivity.

Many have highlighted the necessity to concentrate on the positive outcomes that the adoption of these technologies would bring for humanity with the arrival of GenAI technology at the end of 2022 (Lee, 2023). It has been suggested that it would be advantageous to adopt a more balanced approach in future studies by taking into account both the possible advantages and disadvantages of GenAI technology in the field of education (Bozkurt, 2023). Furthermore, research into the favorable effects of GenAI and ChatGPT in educational settings has suggested that this new technology will elevate traditional teaching methodologies to a new level and induce a paradigm shift (Tlili et al., 2023). This paradigm shift requires an examination of the influence of GenAI on the instructional design and learning processes inside a course. This study primarily investigated the significance of GenAI technologies in the design of learning materials and their impact on learning efficiency.

## Purpose of the Study

The purpose of this research was to assess the content quality of interactive learning materials created by GenAI for an online course. This series of research questions were put forward:

1. At what level would learners engage with GenAI-generated and human-created interactive content?

2. What are the expert opinions on the questions generated by GenAI in terms of content, clarity, and relevance?

    a. How do experts evaluate multiple-choice questions?

    b. How do experts evaluate true-false questions?

    c. How do experts evaluate the fill-in-the-blank questions?

3. What are the expert opinions on the quality of the questions generated by GenAI?

The efficacy of instructional materials hinges on a triad of core principles: content, clarity, and relevance. Content encompasses the substance of the material—its information, concepts, and skills. To be effective, this content must be accurate, comprehensive, and aligned with learning objectives, while also being engaging enough to sustain student motivation (Sun, 2010). Equally important is clarity, which pertains to the lucidity and logical organization of the presentation. Information must be articulated in a straightforward and unambiguous manner to ensure comprehension and prevent confusion (Sundari et al., 2023). Finally, relevance bridges the material to the learner's world. By connecting to students' lives, interests, and cultural contexts, learning becomes more meaningful (Jamilah et al., 2024), a practice shown to significantly boost learner engagement and motivation (Zabala-Vargas et al., 2021). This synergy of rich content, clear presentation, and contextual relevance is what ultimately supports instructional goals and enhances student learning outcomes.

## Background

The integration of GenAI into open and distance learning (ODL) demands critical examination through established pedagogical frameworks. As a technology that leverages deep learning to produce novel, human-like content (Lim et al., 2023), GenAI has reached a tipping point, sparking a wide spectrum of views on its potential impact on education (Lodge et al., 2023). While the use of AI in education is not new (Alasadi & Baiz, 2023; Chen et al., 2022), GenAI's advanced capabilities necessitate a deeper understanding of its pedagogical implications.

The Community of Inquiry (CoI) framework (Garrison et al., 2000), which emphasizes the interplay of cognitive, social, and teaching presence, provides a valuable lens. The pedagogical quality of GenAI-generated content must be assessed on its capacity to foster meaningful cognitive engagement and replicate core instructional functions. This is particularly relevant in assessment, a long-standing area of AI research (Zawacki-Richter et al., 2019), where GenAI can automate the creation of sophisticated questions using machine learning and natural language processing (Bachiri & Mouncif, 2023; Skanda et al., 2020; Tran et al., 2021). Furthermore, Anderson's (2003) Interaction Equivalency Theorem has suggested that robust learning can occur if one form of interaction, such as learner-content, is highly developed. GenAI can enhance this interaction by personalizing learning materials (Lodge et al., 2023), employing advanced transformer models to create context-aware queries (Kriangchaivech & Wangperawong, 2019). However, this must align with cognitive load theory (Sweller, 1988), ensuring that AI-generated content is designed to support, not overwhelm, learners' cognitive processing.

While GenAI offers powerful tools for developing tailored course materials and even assisting in academic writing (Ali & OpenAI Inc., 2023), its application is fraught with challenges. Balancing the scalability and benefits of AI with profound ethical, privacy, and security considerations is paramount (Lim et al., 2023; Nguyen et al., 2023). Therefore, responsible integration into ODL systems depends on rigorous evaluation and continuous human oversight to ensure both pedagogical effectiveness and ethical use.

# Method

In this research, a case study, one of the qualitative research methods, was adopted. A case study can be a single situation selected from life or it can consist of multiple situations that are limited to a certain period. At this point, the case with well-defined boundaries can be a concrete entity such as a person, institution, or group, as well as an abstract concept such as a process or a relationship (Creswell, 2013; Yin, 2014). Yin (2014) argued that the most important point to focus on in case studies is to seek answers to how and why questions.

## Research Context

The study was conducted in the distance computer programming department of a state university in Turkey. Forty-four students enrolled in the Mobile Programming course participated in the study. The course was about mobile application development. In the course, the theoretical background of the mobile application development process was discussed, the Dart programming language (https://dart.dev/) was explained practically, the creation of a mobile application interface in Flutter (https://flutter.dev/) was explained, and sample mobile applications were programmed in Flutter. The theoretical and practical parts of the course were completely synchronous and remote.

Within the scope of the research, tasks were defined for the students in the learning management system (LMS) for the first 8 weeks after the theoretical course. After the tasks, including the file upload task, assessment activities using the H5P tool (https://h5p.org/) were prepared by the researcher and defined to the students weekly over the course of 3 weeks. In the following 3 weeks, H5P activities prepared by the GenAI were defined. The learning materials included multiple-choice questions, fill-in-the-blank problems, and true-false activities. The process is shown in Figure 1.

**Figure 1**

*Research Process*


*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content. H5P is a free and open-source tool primarily used for creating and sharing interactive content (https://h5p.org).

Finally, the learning materials generated by GenAI were sent to experts and their opinions were collected through an online form. These content experts were asked to evaluate these materials in terms of content, relevance, and clarity through a qualitative data collection form, and the evaluation based on these three criteria was combined and interpreted under the concept of "appropriateness." In addition, to provide in-depth answers to the research questions, students' interactions in the LMS were analyzed and used to interpret the qualitative data.

## Participants

The participants consisted of 10 experts, with some specializing in instructional technology and others in computer engineering. The selection of experts from technical sciences was influenced by the case study's focus on mobile programming. Table 1 presents the demographic information of these experts.

**Table 1**

*Demographic Information of the Ten Experts Participating in the Study*

| Expert | Expertise | Degree | Experience, *n* of years |
|---|---|---|---|
| 1 | Instructional technologies | Master's degree | 5–10 |
| 2 | Instructional technologies | Master's degree | 5–10 |
| 3 | Instructional technologies | Bachelor's degree | 5–10 |
| 4 | Instructional technologies & computer engineering | Master's degree | 10–15 |
| 5 | Computer engineering | Master's degree | 1–5 |
| 6 | Instructional technologies | Master's degree | 1–5 |
| 7 | Computer engineering | Master's degree | 5–10 |
| 8 | Computer engineering | Bachelor's degree | 5–10 |
| 9 | Electrical and electronic engineer | Master's degree | 10–15 |
| 10 | Instructional technologies | Bachelor's degree | 5–10 |

The majority of the experts were in the field of instructional technologies, and the majority of them had obtained a master's degree. Except for two, they were also engaged in pursuing doctoral education. In terms of years of experience, the most common group was those with 5–10 years of experience. The distribution of computer engineering specialists appeared to be balanced, with an additional specialist in electrical and electronic engineering.

## Generation of H5P Activities With Generative Artificial Intelligence

Our study employed Nolej AI (https://www.nolej.io/), a GenAI tool capable of generating H5P activities, albeit with certain limitations in its free usage. This platform, which can accept document or

multimedia content as input, subsequently transforms that content into an H5P activity. In this study, the researcher provided a document containing the theoretical course content to the system during the relevant weeks. The content for each week consisted of an average of 2,000 words. Following the generation of the H5P activities, a review and testing phase was initiated. Once it was determined that there were no technical errors, the content H5P activities was integrated into the course's LMS.

## Research Design

In the research, the interaction of learners with the learning activities created by GenAI and the appropriateness of those activities to the course context were accepted as a case, and the case study method was applied (Yin, 2014). Interaction data were obtained from the records of the course on the LMS, while qualitative data were collected by asking open-ended questions to the participants through an online form. The study group included in the qualitative part of the research consisted of experts with knowledge and experience about the content of the course.

## Ethical Issues

Before the data collection phase of the study, a declaration of ethical compliance was obtained by a state university ethics committee commission numbered 179885 and dated November 20, 2023, and the research process was initiated after this declaration of ethical compliance. In addition, participant selection criteria were created, and those who voluntarily participated were included in the data collection process. In addition, students' interaction data in the LMS were anonymized before being analyzed and reported.

## Data Analysis and Learning Analytics

A multifaceted approach was adopted to address the research questions. In this context, the opinions of experts were collected using an online form. The data obtained were analyzed quantitatively with descriptive statistics and qualitatively with a thematic approach. To understand the result of the analysis in-depth, the interaction data of students in the LMS were also analyzed and visualized using descriptive statistics. RStudio (Version 2024.12.1+563) was employed for the analysis of quantitative data.

During the research period, 25,595 records were logged in the LMS. Before the initiation of data analysis, the records underwent a pre-process of anonymization. Subsequently, H5P interactions included in weeks 9–14 were filtered out. Within the context of this study, two metrics—viewing and completing the relevant activity—were evaluated based on student-based time. This process yielded a total of 925 logs that were subsequently analyzed and reported.

## Limitations of the Study

The findings of this study are constrained to the learning materials produced within the scope of the online course titled Mobile Programming and the data collected from experts in the context of these materials. Furthermore, the situation revealed in the research findings and discussion is limited to the opinions and data obtained from the participants involved in the study. In terms of understanding student interaction and engagement, the study is limited to the user interactions recorded by the Moodle system. It is acknowledged that the LMS data may not be sufficient to make a full assessment of students' learning processes and engagement. Therefore, interaction data were used to understand expert opinions in more depth.

# Results and Discussion

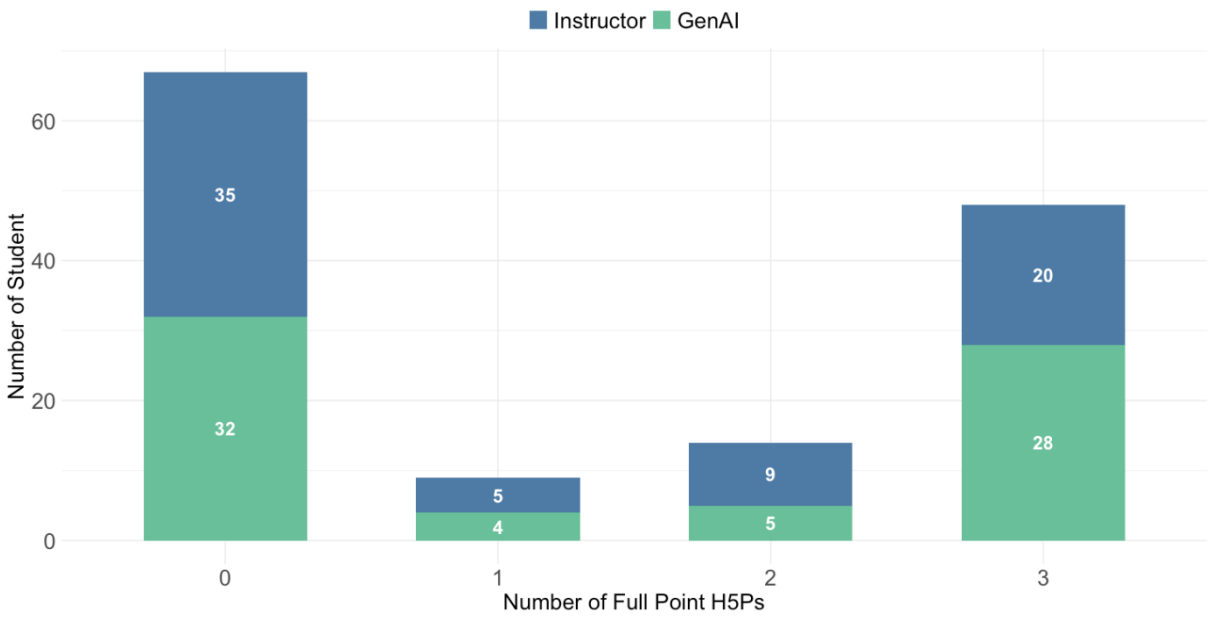
This study explored the engagement levels of students with interactive content generated by GenAI and traditional human-created content. Additionally, it delved into expert assessments of GenAI-generated inquiries, with a particular emphasis on pivotal factors such as content quality, clarity, and relevance. The analysis encompassed a range of question types, including multiple-choice, true-false, and fill-in-the-blank questions. By exploring expert opinions and student engagement, this study has the potential to provide insights into the effectiveness and educational value of interactive content created by GenAI.

## Learner Engagement Levels With GenAI-Generated and Human-Created Interactive Content

Analysis of student engagement with the H5P activities offered insights into performance patterns based on content origin. Out of the 69 students in this distance course, which does not require participation in online activities or completion of assignments, 25 did not participate in these activities. Figure 2 illustrates the performance distribution for the remaining students across three activities created by the instructor and three generated by GenAI. The figure displays the number of students who achieved full points on zero, one, two, or all three activities. The blue bars represent performance on instructor-created content, while the green bars represent performance on GenAI-generated content.

**Figure 2**

*Distribution of Getting Full Points From Instructor-Versus GenAI H5P Activities*




*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content. H5P is a free and open-source tool primarily used for creating and sharing interactive content (https://h5p.org).

As illustrated in Figure 2, there were distinct performance distributions between the two content types. A primary observation was the substantial cohort of students who did not achieve full points in any activity, a figure slightly higher for instructor-created content than for GenAI-created content. In cases

of partial completion (one or two activities), the instructor-created materials appeared to have facilitated incremental success for a slightly larger group of students.

The most significant divergence, however, was evident in the category of complete mastery. A considerably larger group of students successfully obtained full points on all three GenAI-generated activities compared to the instructor-created set. This suggests a polarization of outcomes, particularly with GenAI content; students who engaged successfully were more likely to achieve complete mastery. Conversely, the instructor-created content, while leading to fewer instances of complete mastery, showed a slightly greater capacity to support students in achieving partial success. These findings underscore that GenAI- and instructor-led approaches may produce distinct learning trajectories, with GenAI potentially creating a more uniform "all-or-nothing" performance pattern, while the instructor's activities might possess a more varied difficulty curve. The finding that GenAI-generated content produced a more polarized performance pattern aligns with research on AI-driven personalized learning. As Kasneci et al. (2023) suggested, such tools can create customized learning paths that allow some students to accelerate toward complete mastery, which may explain the higher number of students achieving full points in our study.

Figure 3 provides a granular, student-level visualization of interaction patterns with course materials across 6 weeks, distinguishing between content created by the instructor (weeks 9–11) and by GenAI (weeks 12–14). The central heatmap displays the interaction intensity for each of the 44 students who participated in the activities, where lighter shades of green indicate fewer interactions and darker shades of blue and purple signify a higher number of interactions, as detailed in the color-scale legend. The figure is composed of three interconnected parts:

1. A bar chart at the top summarizes the total number of interactions across all students for each week.

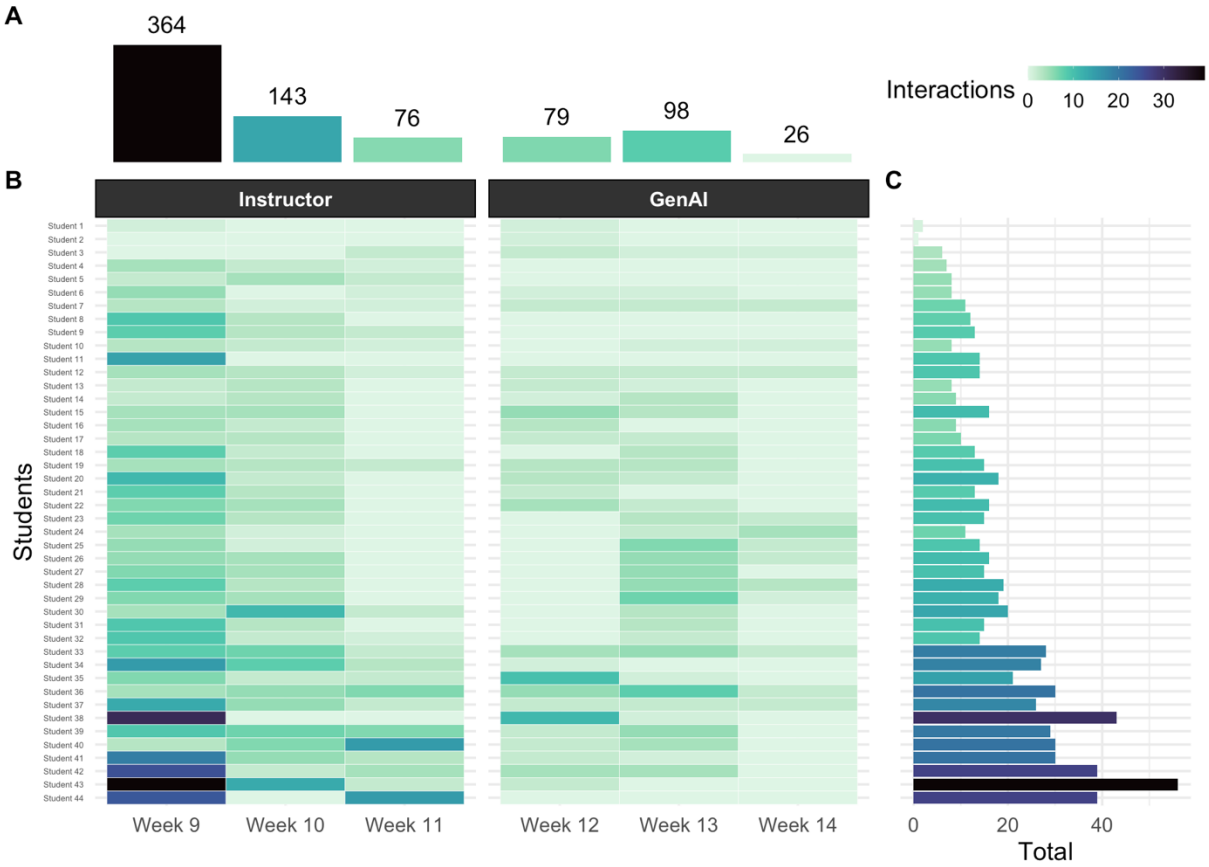
2. The main heatmap illustrates individual student engagement week-by-week.

3. A horizontal bar chart on the right aggregates the total interactions for each student over the entire 6-week period, sorted from least to most engaged.

The interaction data represents the sum of all viewing and completion events recorded in the LMS for each student on the H5P activities.

**Figure 3**

*Distribution of Students' Number of H5P Activity Interactions*




*Note.* Panel A summarizes the total number of interactions across all students for each week. Panel B illustrates individual student engagement week-by-week. Panel C aggregates the total interactions for each student over the entire 6-week period of the activities, sorted from least to most engaged.

Analysis revealed a multi-faceted narrative of student engagement as illustrated in Figure 3. Visually, the most prominent feature was the dramatic peak in interactions seen during week 9, the initial week of the instructor-led activities. This was followed by a sharp and sustained decrease in engagement for all subsequent weeks, showing that the timing within the course or the novelty of the first activities had a significant impact on the overall volume of interactions. Furthermore, stratification of the student body was suggested by the figure: there was a small cohort of highly engaged students (i.e., students 35-44) accounting for a disproportionate number of total interactions, driving the peak in week 9 and remaining the most active throughout the study period.

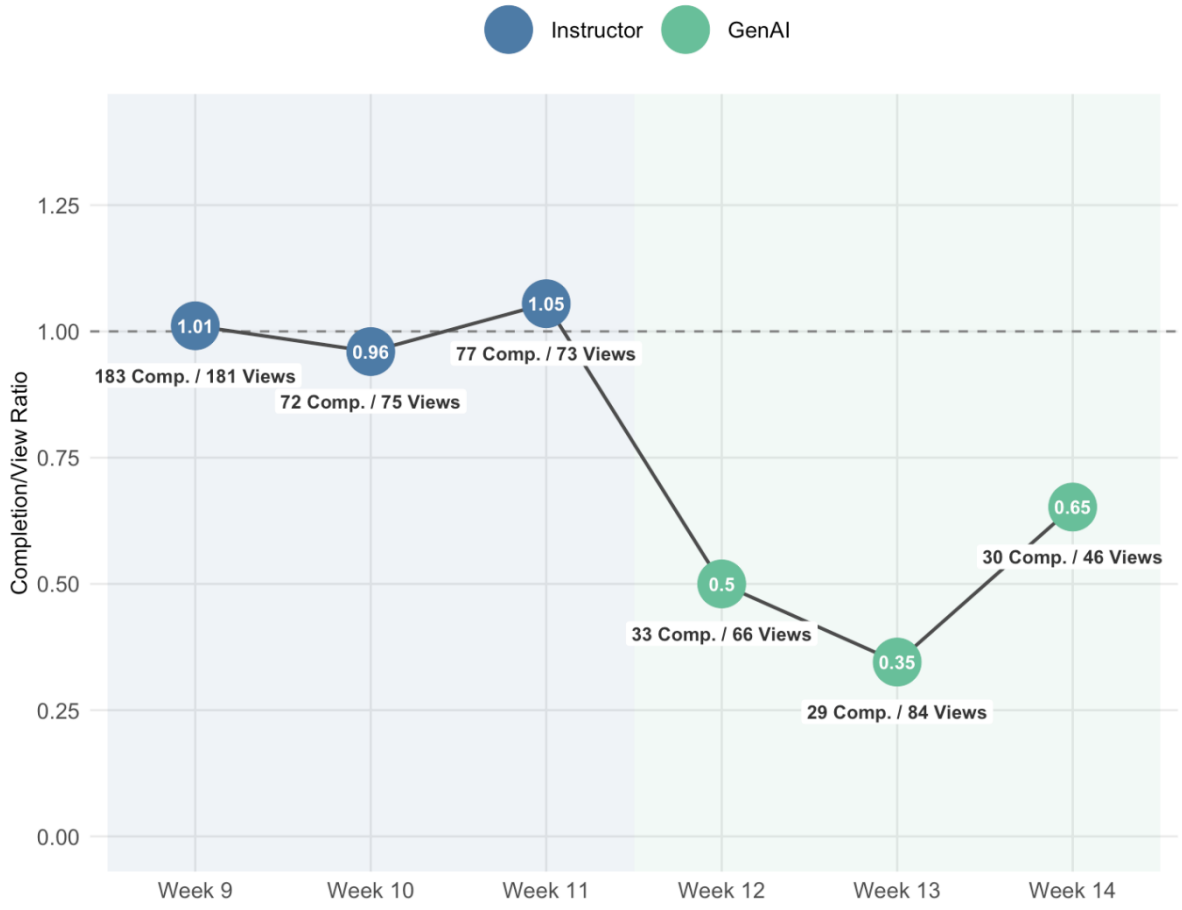
Beyond these visual patterns of timing and distribution, a deeper analysis of the underlying log data exposed critical differences in the nature of these engagements. For the instructor-created content, the ratio of views to completions was 0.99, and activities were often completed more than once. In contrast, for GenAI-generated content, the view-to-completion ratio was 2.13, and activities were typically completed only once.

When viewed together, these two sets of findings provided a comprehensive picture. The high interaction volume in week 9, combined with the efficient completion rate and high repetition, strongly

suggested that the engaged cohort used the instructor's initial activities for mastery-oriented learning. Conversely, the less efficient view-to-completion ratio for GenAI content suggested that students may have struggled with instructional clarity or motivation, leading to more superficial engagement (Figure 4). This dichotomy underscores that while Generative AI is a powerful content-authoring tool, it requires expert human oversight to ensure pedagogical alignment and effectiveness. However, it is crucial to acknowledge that these differing engagement qualities are set against a backdrop of declining overall interaction, reinforcing that student engagement is highly dynamic and influenced by factors beyond just the content's origin.

**Figure 4**

*Weekly Student Interaction Completion/View Ratio of Instructor and GenAI H5P Materials*




*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content. The ratio is calculated by dividing the total number of completion events by the total number of view events for each week. Data labels below each point indicate the raw counts for completions (Comp.) and views used to calculate the ratio for that week. The dashed line at a ratio of 1.0 serves as a baseline, indicating an equal number of completions and views.

Figure 4 reveals a difference in the quality of student engagement between the two phases. The instructor-led H5P activities (Weeks 9-11) are characterized by a high engagement efficiency. The completion-to-view ratio consistently hovers around 1.0, indicating that nearly every view led to a completion. This pattern, particularly the instances where completions exceed views, suggests that students were deeply engaged with the material, likely using it for mastery-oriented learning rather than a cursory review. While the overall interaction volume peaked in Week 9 and then declined, the

efficiency of these interactions remained high, pointing to the pedagogical effectiveness of the instructor-authored content. The initial high volume may indeed be partially attributed to a novelty effect, as students encountered their first H5P activity. Research has indicated that heightened engagement derived from novelty diminishes over time (Liang et al., 2020). Figure 4 is consistent with this finding. However, comparative experimental interventions are needed for a more accurate conclusion.

In contrast, the transition to the GenAI-generated content (Weeks 12-14) is marked by a sharp drop in engagement efficiency. The ratio falls significantly below 1.0, revealing that students viewed the content far more often than they completed it. This suggests a more superficial or exploratory form of engagement, where students may have struggled with instructional clarity or lacked the motivation to complete the activities. While the total number of interactions in this phase appears more stable at a lower level, the quality of these interactions, as measured by the completion ratio, is substantially lower than in the instructor-led phase. However, it is crucial to acknowledge that this apparent difference in performance could be influenced by a confounding variable—namely, that the instructor H5Ps were previously available in the system. Therefore, a dedicated empirical study is needed to isolate the causal impact of content authorship and allow for robust generalizations.

In summary, the findings indicate that a hybrid application of instructor-created and GenAI-generated materials may address different aspects of student engagement. The materials created by the instructor were associated with higher completion-to-view ratios, a pattern consistent with mastery-oriented learning. Conversely, content generated by GenAI corresponded with lower engagement efficiency, suggesting a need for human oversight to enhance its pedagogical alignment and effectiveness. Although AI can be leveraged to reduce unnecessary cognitive load and increase intrinsic motivation (Guo et al., 2023; Xu & Ismail, 2024), the data suggest these outcomes are contingent on careful implementation. The integration of these two methods aligns with tenets of engagement theory, which posits that collaborative interactions can be amplified through technological applications to foster meaningful educational experiences (Bachiri et al., 2023; Huang et al., 2021).

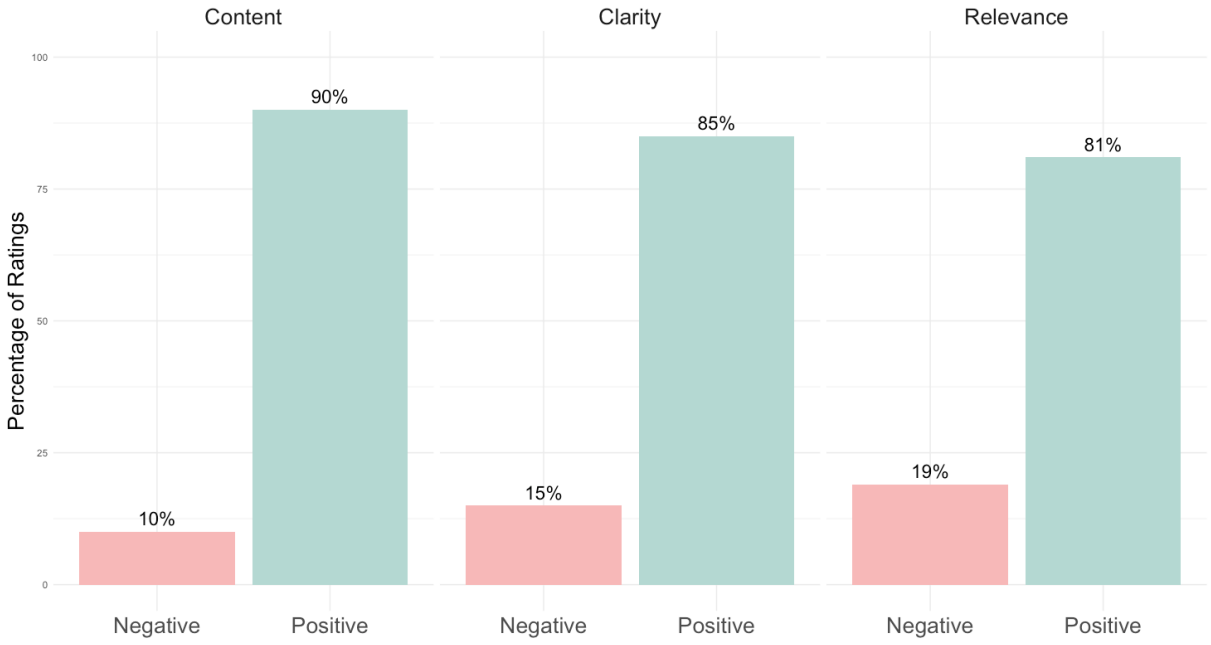
While these quantitative interaction patterns from the LMS provided valuable insight into student engagement, they do not, on their own, explain the underlying reasons for these behaviors. The data revealed what students did, but not intrinsic quality or pedagogical soundness of the materials they interacted with. To evaluate GenAI-generated content, clarity, and relevance, and thus to contextualize the engagement findings, an assessment by subject matter experts was conducted. The following sections present the results of this expert evaluation, which scrutinized GenAI materials from a pedagogical and content-validity perspective.

## Expert Opinions on GenAI-Generated Questions

The expert opinions on the content, clarity, and relevance of the multiple-choice questions generated by Gen AI were analyzed. To evaluate the GenAI-generated questions, subject matter experts rated multiple-choice, true-false, and fill-in-the-blank questions across three criteria: content, clarity, and relevance. For each question and criterion, experts provided a binary rating of 'Negative' (inappropriate) or 'Positive' (appropriate). The percentages shown in the figures represent the aggregate distribution of these ratings. Figure 5 shows the percentage of positive versus negative opinions in each category under consideration.

**Figure 5**

*Expert Evaluation of GenAI-Generated Multiple-Choice Questions*



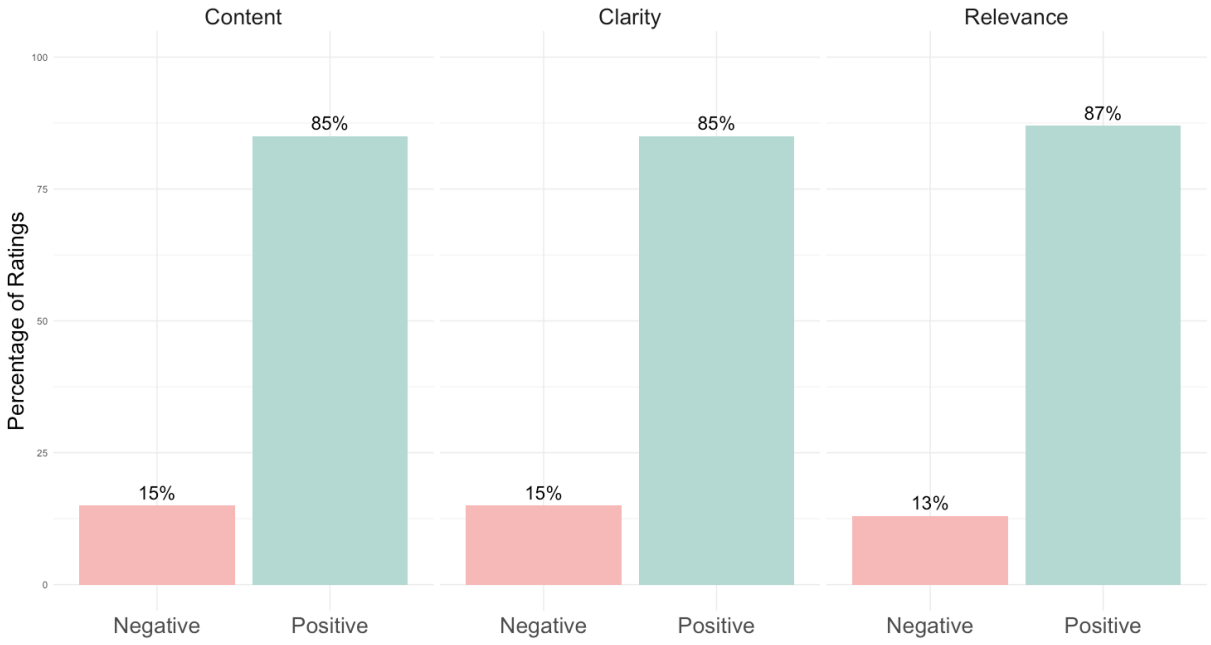

*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content.

As can be seen in Figure 5, the ratings for multiple-choice questions created by GenAI were, on average, 85% positive across the three criteria. Among the criteria for multiple-choice questions, 'Content' received the highest percentage of positive ratings (90%), while 'Relevance' received the lowest (81%).

The ratio of the content, clarity, and relevance levels of the true-false questions that received expert opinion within the scope of the research is presented in Figure 6. It shows consistency and that approximately 85% of the true-false questions prepared by GenAI were found to be appropriate.

**Figure 6**

*Expert Evaluation of GenAI-Generated True-False Questions*

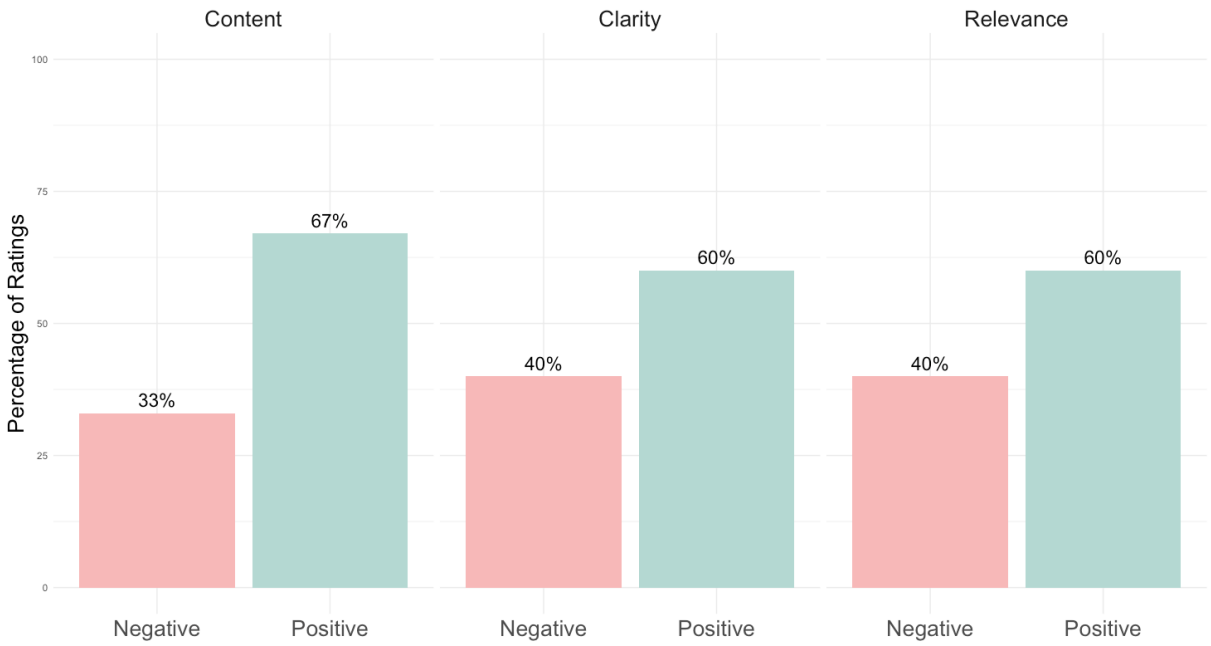



*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content.

The ratio of the content, clarity, and relevance levels of the fill-in-the-blank questions that received expert opinion within the scope of the research is presented in Figure 7.

**Figure 7**

*Expert Evaluation of GenAI-Generated Fill-in-the-Blank Questions*



*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content.

The fill-in-the-blank questions created by GenAI were found appropriate at an average rate of 62%. Experts evaluated the content quality of the questions as the highest and the clarity and relevance of the questions as the lowest. However, in contrast to the other question types, the majority of ratings for fill-in-the-blank questions were not positive, indicating a general inadequacy in all three criteria.

To provide a holistic view of the quantitative data, the expert ratings for all three question types (multiple-choice, true-false, and fill-in-the-blank) were aggregated. Figure 8 presents this general summary, illustrating the overall distribution of 'Positive' and 'Negative' ratings across the core criteria of content, clarity, and relevance.

**Figure 8**

*Aggregated Distribution of Expert Evaluation Across All Question Types*


As shown in Figure 8, which aggregates all ratings, the overall positive rating percentage was highest for 'Content' (80.67%) and lowest for 'Relevance' (76%).

The findings indicated that the content quality of interactive learning materials produced by GenAI was inconsistent and highly dependent on the assessment format. GenAI was proficient at generating structured question types, with both multiple-choice and true-false questions being found appropriate by experts at an average rate of approximately 85%. For these formats, which are often used in online courses to assess foundational knowledge, the AI demonstrated a strong ability to produce factually accurate and clear content, supporting the conclusions of other researchers that GenAI can be a valuable resource for educators. Research by Sihite et al. (2023) showed that GenAI-generated questions can cover a range of cognitive levels, thus providing a valuable resource for educators seeking to improve their assessment methods. Nasution (2023) assessed the validity and reliability of AI-generated biology

questions and showed that GenAI can produce high-quality educational materials. This suggests that for creating a baseline of standard assessment items, GenAI is a capable and efficient tool.

However, our study revealed significant limitations in GenAI's ability to generate content requiring deeper semantic understanding. Fill-in-the-blank questions were found to be largely inadequate, with an average appropriateness rate of only 62%, and experts rated them poorly across all criteria. Furthermore, even for the highly-rated multiple-choice questions, relevance was the least satisfactory criterion. This discrepancy suggests that while GenAI can assemble correct information, it struggles to ensure that the content is contextually aligned with specific instructional goals, a critical component of quality in educational materials.

The most effective educational outcomes are achieved through partnership rather than replacement. The critical role of subject matter experts in this study underscores a central theme in recent literature: while Generative AI is proficient at creating content, its outputs require human evaluation to ensure pedagogical nuance and contextual accuracy (Baidoo-Anu & Ansah, 2023). GenAI should be regarded as a powerful assistant for content creation in online learning, rather than a fully autonomous solution. While students report positive perceptions of AI tools and appreciate their support (Aydemir & Seferoğlu, 2024), our findings emphasized that this enthusiasm must be met with rigorous quality control. The variability in output quality necessitates a "human-in-the-loop" model, where instructional designers and subject matter experts play an indispensable role in validating, refining, and ensuring the pedagogical relevance of all AI-generated materials. While technology can accelerate the development of formative assessments, human oversight remains crucial to guarantee the high quality and instructional integrity required for effective online courses.

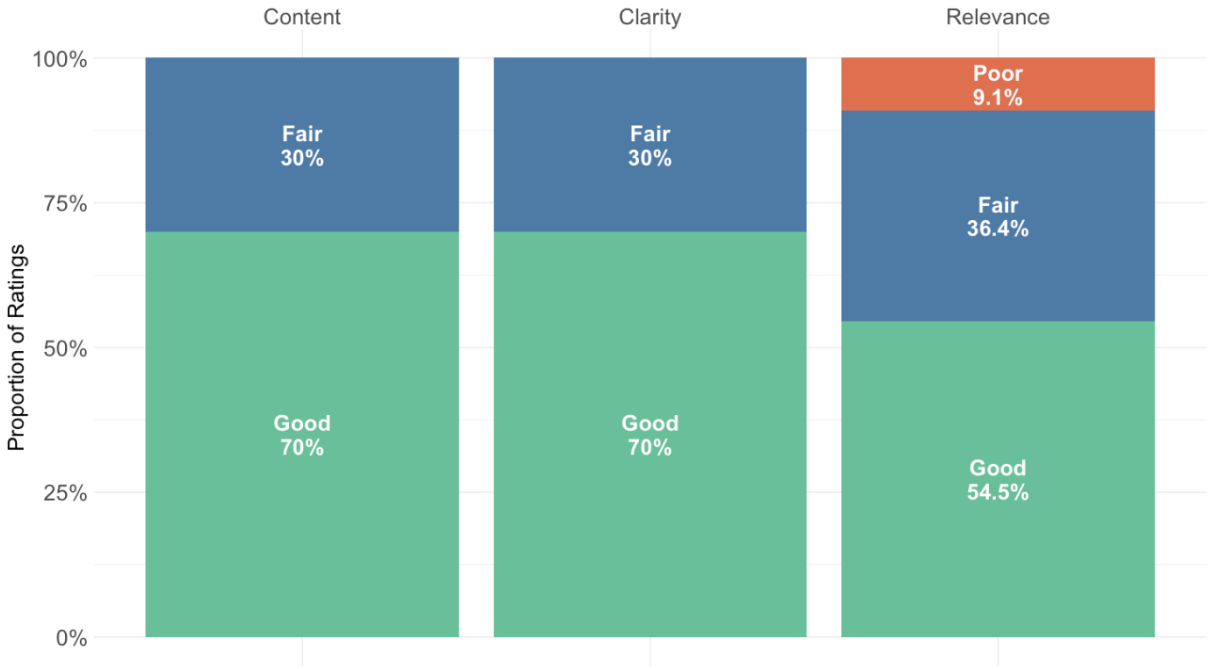
## Qualitative Analysis of Expert Feedback on the Quality of GenAI-Generated Questions

Following the quantitative evaluation where experts provided ratings for the GenAI-generated questions, a qualitative analysis was conducted to explore the reasoning behind their assessments. In addition to the ratings, experts were asked to provide open-ended feedback on the overall quality, advantages, disadvantages, and suggestions for improving the H5P activities. This section presents a thematic analysis of their written responses, providing deeper context for the quantitative results presented previously.

In the qualitative data collection phase of this study, open-ended questions were asked to the participants. In their first three questions, participants were asked to evaluate the GenAI-generated questions on the concepts of content, clarity, and relevance by rating them as good, fair, or poor. Results are shown in Figure 9.

**Figure 9**

*Evaluation of the Content, Clarity, and Relevance of the Questions Generated by GenAI*




*Note.* Generative artificial intelligence (GenAI) is a type of artificial intelligence that uses patterns from existing data to create new and original content.

The evaluation by ten experts indicated that the 'Content' and 'Clarity' of the GenAI-generated questions were rated as "Good" at an identical proportion of 70%, while the 'Relevance' of the questions received a lower proportion of "Good" ratings. Notably, for both 'Content' and 'Clarity', no ratings were in the "Poor" category. However, experts expressed greater concern regarding the 'Relevance' of the questions, where ratings were distributed across "Fair" (36.4%) and "Poor" (9.1%) categories, in addition to "Good". These findings suggested that while GenAI is proficient in generating clear and factually sound content, its ability to align that content with specific pedagogical contexts—its relevance—requires further improvement. Jouault et al. (2016) conducted a comparative study between automatically generated questions and questions generated by human experts. They found that GenAI-generated questions were able to cover a significant portion of knowledge acquisition objectives, thus confirming the effectiveness of GenAI in educational settings (Aydemir & Seferoğlu, 2024).

In the other three questions in the qualitative data collection process, the experts were asked for their opinions about the quality, advantages-disadvantages, and suggestions for improvement of the questions developed by GenAI, respectively.

Figure 10 shows their opinions about quality, categorized by theme.

**Figure 10**

*Experts' Opinions on the Quality of Questions Developed by GenAI*

| Themes | Experts |
|---|---|
| Content appropriateness | E3, E5, E6, E8, E10 |
| Human intervention required | E2, E4, E5, E6 |
| Suitability to learner level | E6, E8, E9 |
| Uncertainties | E4, E1 |

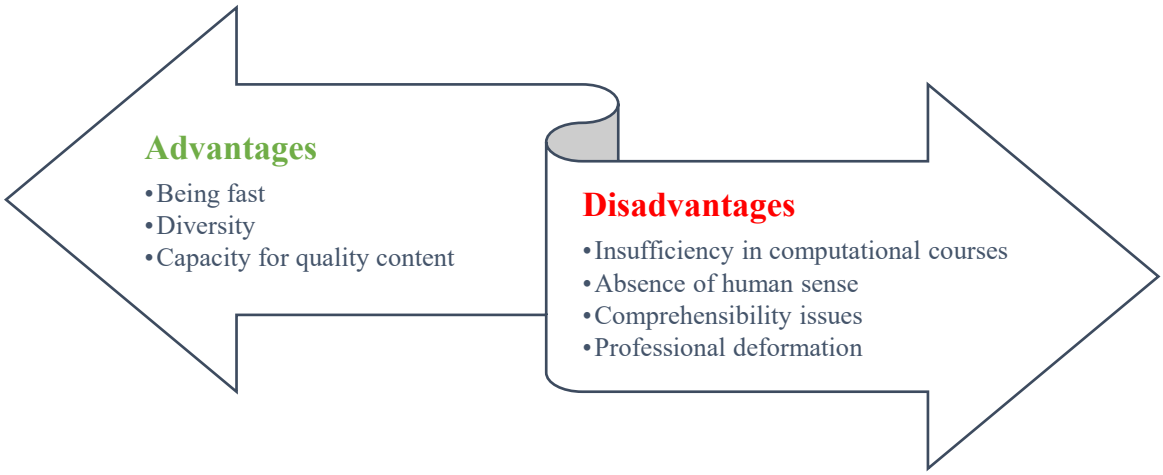
*Note*. E = expert participant.

While *content appropriateness* was the most emphasized theme, identified by five experts, four experts emphasized *human intervention required*. One expert wrote, "I see no harm in using it as long as it is re-evaluated by the instructor. If it is not revised, it can still make simple mistakes." The theme, *suitability to learner level*, was a point that drew the attention of three experts. Finally, two experts drew attention to the fact that some questions contained incomplete information or had poor comprehensibility, and the theme of *uncertainties* emerged.

In the next question, experts were asked about the advantages and disadvantages of the questions developed by GenAI. Figure 11 shows the results of their answers.

**Figure 11**

*Potential Advantages and Disadvantages of AI-Generated Questions*

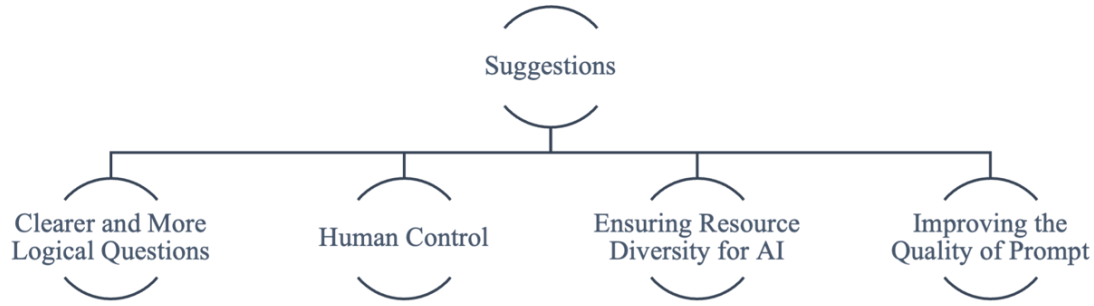


As seen in this table, there were criticisms, such as the issue that GenAI is insufficient in some subjects (such as computational courses) and that some questions contain ambiguity. In terms of advantages, themes of being fast, diversity, and capacity for quality content stood out. Among the disadvantages, comprehensibility issues would have a significant role.

In the last open-ended question, experts were asked for suggestions to improve the questions created by GenAI. The qualitative data obtained from this question were analyzed under four main themes: human control, clearer and more logical questions, ensuring resource diversity for GenAI, and improving the quality of prompts. The connections between these themes is shown in Figure 12.

**Figure 12**

*Suggestions for Enhancing the Questions Formulated by GenAI*




*Note*. AI = Artificial intelligence.

The most frequently mentioned suggestions of the experts were *human control* and *producing clearer and more logical questions*. One participant drawing attention to human control stated: "In my opinion, human control is important to improve the questions produced by GenAI. Human control of the content, appropriateness, and comprehensibility of questions can improve the quality of the questions."

Furthermore, to enhance learning experience and learner success, augmenting the resources provided to GenAI and formulating clearer, more stimulating questions were also highlighted as recommendations for improvement. Enhancing the capacity to formulate questions through the generation of additional inquiries by GenAI was also included among recommendations. A further recommendation was to refine the terminology employed. One expert reported: "To preemptively train GenAI, particularly to uphold and preserve linguistic purity. It may be useful to train and test it in isolation without confusing it about topics to be produced [as] questions." This idea could be considered while giving prompts.

## Conclusion and Future Directions

In this research, course assessment activities for some weeks of a distance course were developed by GenAI, and learners' interactions with these course assessment activities through a LMS were analyzed. In addition, the suitability of these assessment activities to the context of the course was evaluated by 10 experts in course content. In this evaluation process, the questions developed by GenAI were analyzed in terms of content, clarity, and relevance, and the experts were asked for their opinions with open-ended questions and a quantitative evaluation.

This study demonstrates that learners' engagement with content generated by GenAI and produced by instructors may fluctuate based on contextual variables. Results indicate that carefully integrating these two material categories may effectively address learners' diverse demands. More precisely, within the realm of distance education, materials produced by GenAI could serve as adapted instructional materials at the micro-instructional level and enhance educators' pedagogical approaches to design. To fully actualize this potential, it is essential to create professional development programs that empower educators to proficiently integrate AI content and enhance their AI literacy. Additionally, extensive systems such as massive open online course (MOOC) platforms need to establish criteria that define quality requirements for content generated through AI-human collaboration and allocate resources for data-driven enhancements to adaptive learning systems. Forthcoming research will substantially improve the area by elucidating these findings using experimental approaches that investigate learner interaction dynamics and their impact on learning outcomes.

# References

Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education, 100*(8), 2965–2971. https://doi.org/10.1021/acs.jchemed.3c00323

Ali, F., & OpenAI, Inc., C. (2023). Let the devil speak for itself: Should ChatGPT be allowed or banned in hospitality and tourism schools? *Journal of Global Hospitality and Tourism, 2*(1), 1–6. https://doi.org/10.5038/2771-5957.2.1.1016

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distributed Learning, 4*(2). https://doi.org/10.19173/irrodl.v4i2.149

Aydemir, H., & Seferoğlu, S. S. (2024). Investigation of users' opinions, perceptions and attitudes towards the use of artificial intelligence in online learning environments. *European Journal of Open, Distance and E-Learning, 26*(1), 70–92. https://doi.org/10.2478/eurodl-2024-0007

Bachiri, Y.-A., & Mouncif, H. (2023). Artificial intelligence system in aid of pedagogical engineering for knowledge assessment on MOOC platforms: Open EdX and Moodle. *International Journal of Emerging Technologies in Learning (iJET), 18*(05), 144–160. https://doi.org/10.3991/ijet.v18i05.36589

Bachiri, Y.-A., Mouncif, H., & Bouikhalene, B. (2023). Artificial intelligence empowers gamification: Optimizing student engagement and learning outcomes in e-learning and MOOCs. *International Journal of Engineering Pedagogy (iJEP), 13*(8), 4–19. https://doi.org/10.3991/ijep.v13i8.40853

Baidoo-anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Bozkurt, A. (2023). ChatGPT, üretken yapay zekâ ve algoritmik paradigma değişikliği [ChatGPT, generative artificial intelligence and algorithmic paradigm shift]. *Alanyazın 4*(1), 63–72. https://doi.org/10.59320/alanyazin.1283282

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education. *Educational Technology & Society, 25*(1), 28–47. https://www.jstor.org/stable/48647028

Creswell, J. W. (2013). Qualitative inquiry and research design: Choosing among five approaches (3rd ed.). Sage.

DaCosta, B., & Kinsell, C. (2024). Investigating media selection through ChatGPT: An exploratory study on generative artificial intelligence in the aid of instructional design. *Open Journal of Social Sciences, 12*(4), 187–227. https://doi.org/10.4236/jss.2024.124014

Ferrara, E. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci, 6*(1), Article 3. https://doi.org/10.3390/sci6010003

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Guo, Z., Jing, T., Zhu, H., & Xu, X. (2023). Analysis of the impact of artificial intelligence on college students' learning. *The Frontiers of Society, Science and Technology, 5*(13), 110–116. https://doi.org/10.25236/fsst.2023.051318

Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2*(4), Article 100089. https://doi.org/10.1016/j.tbench.2023.100089

Huang, J. Shen, G., & Ren, X. (2021). Connotation analysis and paradigm shift of teaching design under artificial intelligence technology. *International Journal of Emerging Technologies in Learning (iJET), 16*(05), 73–86. https://doi.org/10.3991/ijet.v16i05.20287

Jamilah, Jayadi, K., Yatim, H., Sahnir, N., Djirong, A., & Abduh, A. (2024). The integration of local cultural arts in the context of teaching materials on the implementation of the Merdeka Belajar curriculum. *Journal of Education Research and Evaluation, 8*(2), 404–413. https://doi.org/10.23887/jere.v8i2.78359

Jouault, C., Seta, K., & Hayashi, Y. (2016). Content-dependent question generation using LOD for history learning in open learning space. *New Generation Computing, 34*(4), 367–394. https://doi.org/10.1007/s00354-016-0404-x

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kır, Ş., & Şenocak, D. (2022). Açık ve uzaktan öğrenme sistemlerinde yapay zekânın öğrenen destek hizmeti bağlamında kullanımı [The use of artificial intelligence in the context of learner support services in open and distance learning systems]. *Dijital Teknolojiler Ve Eğitim Dergisi*, *1*(1), 39–56. https://doi.org/10.5281/zenodo.6647642

Kriangchaivech, K., & Wangperawong, A. (2019). *Question generation by transformers*. ArXiv. https://doi.org/10.48550/arxiv.1909.05017

Lee, E. (2023). Is ChatGPT a false promise? *Berkeley Blog*. https://coilink.org/20.500.12592/hzdz6k

Liang, Z., Zhao, Q., Zhou, Z., Yu, Q., Li, S., & Chen, S. (2020). The effect of "novelty input" and "novelty output" on boredom during home quarantine in the COVID-19 pandemic: The moderating effects of trait creativity. *Frontiers in Psychology, 11*, Article 601548. https://doi.org/10.3389/fpsyg.2020.601548

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), Article 100790. https://doi.org/10.1016/j.ijme.2023.100790

Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology, 39*(1), 1–8. https://doi.org/10.14742/ajet.8695

Nasution, N. E. A. (2023). Using artificial intelligence to create biology multiple choice questions for higher education. *Agricultural and Environmental Education, 2*(1), Article em002. https://doi.org/10.29333/agrenvedu/13071

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies, 28*, 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Sihite, M. R., Meisuri, M., & Sibarani, B. (2023). Examining the validity and reliability of ChatGPT 3.5-generated reading comprehension questions for academic texts. *Randwick International of Education and Linguistics Science Journal, 4*(4), 937–944. https://doi.org/10.47175/rielsj.v4i4.835

Skanda, V. C., Jayaram, R., Bukitagar, V. C., & Kumar, N. S. (2020). Automatic questionnaire and interactive session generation from videos. In A. Chandrabose, U. Furbach, A. Ghosh, & M. A. Kumar (Eds.), *IFIP advances in information and communication technology* (Vol. 578; pp. 205–212). Springer. https://doi.org/10.1007/978-3-030-63467-4_16

Sun, Z. (2010). Language teaching materials and learner motivation. *Journal of Language Teaching and Research, 1*(6), 889–892. https://doi.org/10.4304/jltr.1.6.889-892

Sundari, W., Siahaan, S. M., & Ismet, I. (2023). Digital teaching materials based on attention relevance confidence satisfaction substance pressure material. *Journal of Curriculum Indonesia, 6*(1), 69–75. https://doi.org/10.46680/jci.v6i1.77

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), Article 15. https://doi.org/10.1186/s40561-023-00237-x

Tran, S., Krishna, P., Pakuwal, I., Kafle, P., Singh, N., Lynch, J., & Drori, I. (2021). *Solving machine learning problems*. ArXiv. https://doi.org/10.48550/arxiv.2107.01238

Wood, D., & Moss, S. H. (2024). Evaluating the impact of students' generative AI use in educational contexts. *Journal of Research in Innovative Teaching & Learning, 17*(2), 152–167. https://doi.org/10.1108/jrit-06-2024-0151

Xu, B., & Ismail, H. H. (2024). The impact of artificial intelligence-assisted learning applications on oral English ability: A literature review. *International Journal of Academic Research in Progressive Education and Development, 13*(4), 1118–1134. https://doi.org/10.6007/ijarped/v13-i4/23352

Yin, R. K. (2014). *Case study research design and methods*. Sage Publications.

Zabala-Vargas, S., García-Mora, L. H., Arciniegas-Hernández, E., Reina-Medrano, J. I., de Benito-Crosetti, B., & Darder-Mesquida, A. (2021). Strengthening motivation in the mathematical engineering teaching processes—A proposal from gamification and game-based learning. *International Journal of Emerging Technologies in Learning (iJET), 16*(06), 4–19. https://doi.org/10.3991/ijet.v16i06.16163

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16,* Article 39. https://doi.org/10.1186/s41239-019-0171-0


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Psychometric Properties of the Arabic Version of the Unified Theory of Acceptance and Use of Technology (UTAUT-2012) Among Nursing Students

**Latifah Alenazi**
*Nursing Administration and Education Department, College of Nursing, King Saud University, Saudi Arabia*

## Abstract

The integration of artificial intelligence (AI) into nursing education is essential for equipping future nurses with the skills required to navigate an increasingly technology-driven healthcare environment. This study aimed to validate the Arabic version of the Unified Theory of Acceptance and Use of Technology (UTAUT-2012) model in assessing factors influencing nursing students' acceptance and use of AI in healthcare education. A cross-sectional pilot study was conducted with 200 nursing students to evaluate the psychometric properties of the Arabic-translated UTAUT (2012) instrument. Confirmatory factor analysis was performed using covariance-based structural equation modeling (CB-SEM) in SmartPLS (Version 4.1.0). Confirmatory factor analysis supported the construct validity of the nine UTAUT 2012 constructs: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, habit, behavioral intention, and use behavior. All items showed acceptable factor loadings (> .5), composite and construct reliability (> 0.7), and average variance extracted (> 0.5). Discriminant validity was confirmed using the Fornell-Larcker criterion and the heterotrait-monotrait ratio. The findings offer valuable insights into the factors influencing Arabic-speaking nursing students' acceptance and use of AI in healthcare education, supporting the model's validity in this cultural context.

*Keywords*: psychometric properties, UTAUT 2012 model, artificial intelligence, structural equation modeling, nursing education, nursing students

# Introduction

Incorporating artificial intelligence (AI) into nursing education is becoming increasingly important for preparing nurses to meet the evolving demands of modern healthcare. As AI continues to reshape clinical practice, its integration into nursing curricula is essential to ensure that future nurses are well-equipped to navigate these technological changes (O'Connor et al., 2023). Successful implementation of AI in nursing education requires a nuanced understanding of the factors that influence nursing students' intentions and readiness to engage with AI applications (Alenezi, 2023). As emphasized by Alenezi (2023) and O'Connor et al. (2023), identifying these factors is critical for supporting the effective integration of AI within educational environments.

To investigate these factors, researchers must employ robust, psychometrically sound instruments. Accurate measurement of students' acceptance and use of AI depends on tools that demonstrate strong reliability and validity, particularly when adapted for different languages and cultural contexts. However, there remains a notable gap in the literature regarding validated Arabic-translated instruments, especially within the field of nursing education (Taskiran, 2023). Existing technology acceptance models have largely been developed and tested in Western contexts, raising concerns about their generalizability to non-Western populations (McCoy et al., 2007). Although these models have demonstrated high explanatory power in Western settings, their applicability in non-Western cultures has been questioned (Bandyopadhyay & Fraccastoro, 2007).

Recent studies have emphasized the importance of incorporating cultural nuances into technology acceptance research, particularly within educational systems in non-Western countries (Drissi et al., 2022; Gabriel, 2023; Rouibah, 2008). For Arabic-speaking populations, cultural factors significantly influence how technology is perceived and adopted in educational contexts (Drissi et al., 2022; Gabriel, 2023). Consequently, validating these models in Arabic-speaking environments is essential for accurate and meaningful application.

This study adopts the Unified Theory of Acceptance and Use of Technology (UTAUT) 2012 model, which provides a comprehensive framework for analyzing the determinants of technology adoption. The updated UTAUT 2012 model, an evolution of the original framework, has gained substantial recognition as a theoretical advancement due to its continued refinement through empirical research (Venkatesh et al., 2012). It has been extensively applied in educational settings to assess technology acceptance across various populations, particularly in non-Arabic-speaking contexts (Abu-Al-Aish & Love, 2013; Barchielli et al., 2021; Williams et al., 2015; Zhou et al., 2019). Its flexibility and robustness make it particularly valuable for cross-cultural studies, offering insights into students' and nurses' readiness to adopt emerging technologies such as AI in education (Kwak et al., 2022).

The UTAUT 2012 model evaluates user acceptance and technology usage behavior through seven core constructs: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit, which in turn influence behavioral intention and use behavior (Venkatesh et al., 2003, 2012). These constructs are defined as follows: performance expectancy (PE) refers to the perceived usefulness of a system in enhancing job performance; effort expectancy (EE) relates to the ease of use; social influence (SI) concerns the perceived expectations of important others; facilitating conditions (FC) address the perceived support available for system use; hedonic motivation (HM) reflects the enjoyment derived from system use; price value (PV) pertains to the trade-off between benefits and cost; and habit (HT) captures the extent of automatic use behavior (Venkatesh et al., 2003,

2012).

In Arabic-speaking contexts, the UTAUT model has shown promising results. For instance, a study examining Internet banking behaviors across Jordan, Saudi Arabia, and Egypt with 677 participants validated the Arabic version of the model through confirmatory factor analysis (CFA), affirming its applicability in non-Western settings (Al-Qeisi et al., 2015). Additionally, in Saudi Arabia, the UTAUT 2012 model was employed to explore factors affecting the use of virtual classrooms among faculty in various disciplines at the University of Ha'il. This study, which involved 235 teaching staff, validated the measurement model via CFA after modifying select items to improve model fit (SI3, SI4, M4, M5, FC4). The results supported the model's reliability and validity, with satisfactory values for average variance extracted (AVE) and composite reliability (CR; Alshammari, 2021).

Despite these findings, the psychometric properties of the UTAUT 2012 model have not yet been evaluated in the context of nursing education in Saudi Arabia. Given the increasing relevance of AI in healthcare, assessing the model's reliability and validity in this field is essential, particularly for Arabic-speaking nursing students. To date, no studies have directly examined the psychometric validation of the Arabic-translated UTAUT 2012 model within nursing education.

Validating the Arabic version of the UTAUT 2012 model for nursing students holds meaningful implications for both educational practice and policy. Educators can design more targeted learning experiences that promote engagement and readiness for technology-driven healthcare environments by understanding the factors that influence students' acceptance of AI. Institutions may also use the validated tool to identify barriers to AI adoption and provide the necessary support or training. Ultimately, this research contributes not only to measurement science but also to improving digital transformation in nursing education across Arabic-speaking regions.

Evaluating the psychometric properties of research instruments is crucial to ensuring the accuracy, credibility, and generalizability of study findings. Valid and reliable tools are necessary to support the integrity of research outcomes across diverse contexts. Several studies have highlighted the significance of psychometric evaluation in research involving questionnaires, emphasizing its role in strengthening the overall quality of measurement (Dikkema et al., 2021; Speyer et al., 2014; Terwee et al., 2007).

In cross-cultural research, applying models without appropriate linguistic and contextual adaptation can result in measurement error and limit the interpretability of outcomes. Within Arabic-speaking contexts, the lack of validated instruments presents a methodological gap that constrains efforts to assess students' acceptance of educational technologies with precision. This study addresses this limitation by adapting and validating the UTAUT-2012 model for Arabic-speaking nursing students, thereby contributing a culturally and linguistically appropriate measurement tool that can support evidence-based decision-making in educational planning and policy development (Venkatesh et al., 2012).

# Materials and Methods

## Research Design and Participants

The study used a cross-sectional design. To meet the assumptions of confirmatory factor analysis, the sample size was 200 participants (Kline, 2016). The questionnaire form was available on Google Forms from early September to late October 2023. This study included males and females from the first year

to year four who enrolled in the nursing bachelor's program. The study excluded students with academic issues, such as academic suspension.

## Ethical Considerations

The researchers had institutional review board (IRB) approval from the university in the study (KSU-HE; 23-838), and the approval date was 5 September 2023. The researchers used an electronic questionnaire to ensure that no one's identity could possibly be exposed. The electronic questionnaire maintained participant confidentiality to ensure no identifying information was obtained. Each participant was free to discontinue the study at any point, with no impact or consequences, at their will. Each participant signed an electronic consent form contract. Each respondent was asked to indicate their agreement to the terms of an electronic consent form that was part of the survey. The privacy statement was presented before the participant clicked "Yes," indicating they would answer the survey questions. The participant was routed to the disagree page if the respondent clicked "No."

## Research Instrument

### *UTAUT 2012 Model Questionnaire*

The first section of the questionnaire asked the participants about their demographic characteristics (age, gender, and year of study). The study variables were measured using the question items developed by Venkatesh et al. (2012). That is, the questionnaire was a 30-item self-report measure of performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, habit, behavioral intention, and use behavior. Official approval for its use was obtained from the author. For this study, it was adapted to the context of AI usage.

The questionnaire consisted of nine constructs: performance expectancy (4 items), effort expectancy (4 items), social influence (3 items), facilitating conditions (4 items), hedonic motivation (3 items), price value (3 items), habit (3 items), behavioral intention (3 items), and use behavior (3 items). All the items were rated on a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree).

### *Translation of the Study Questionnaire Into Arabic*

After obtaining permission from the questionnaire's original author, the researcher sent the questionnaire to five experts in nursing education to evaluate in order to adapt and translate the instrument. We used a comprehension and cultural relevance index for the questionnaire items, asking each expert to evaluate the items on that basis. The index ranged from 1 (not at all) to 10(very much). Based on the expert's reviews, we made slight modifications to items.

All items were translated from English into Arabic and back-translated by the official translator with the assistance of a linguistic expert to ensure that the content and meaning of the items were the same as the original. Sidani et al. (2010) emphasized the importance of investigating conceptual equivalence while adjusting and translating instruments. Some concepts in one culture may not always have the same meaning in another, depending on how they are perceived, understood, and operationalized (also known as conceptualization and operationalization; Sidani et al., 2010). Moreover, the researcher tested content validity by having experts review the Arabic translation of the questionnaire. Content validity is the degree to which items accurately reflect the content of the characteristic or quality the researcher wants to measure (Polit & Beck, 2016).

## Data Analysis

### *Overview of Statistical Approach*

Data analysis was conducted using SmartPLS (Version 4.1.0). CFA was performed to assess the measurement model and ensure the reliability and validity of the Arabic-translated UTAUT-2012 instrument. Factor loadings were analyzed to determine how well each item reflected its intended construct. Internal consistency and reliability were examined using Cronbach's alpha and rho_C coefficients, which measure how closely related the items within each construct are. Construct validity was evaluated through the average variance extracted (AVE), indicating how much of the item variance is captured by the underlying factor. Discriminant validity, or the degree to which constructs are distinct from one another, was assessed using the Fornell-Larcker criterion and the heterotrait-monotrait ratio (HTMT). In addition, model fit was examined using several indices: chi-square divided by degrees of freedom (CMIN/DF); goodness of fit index (GFI); comparative fit index (CFI); Tucker-Lewis index (TLI); standardized root-mean-square residual (SRMR); and root-mean-square error of approximation (RMSEA) to determine how well the model fit the observed data. These analyses provided the basis for evaluating the psychometric properties of the Arabic version of the UTAUT-2012 model.

### *Confirmatory Factor Analysis*

The type and strength of the interactions between constructs are now the focus instead of the relationships between latent constructs and measured variables. Confirmatory factor analysis (CFA) emphasizes the relationships between measured variables and latent constructs (Kline, 2016). The CFA value was used to evaluate the validity and unidimensionality of the entire assessment model.

This approach does not require distinguishing between dependent and independent variables. Therefore, the oval forms represent latent variables. One-headed connections show a causal relationship between a construct and an indicator, whereas two-headed arrows show covariance between constructs. The factor loadings were assessed for each item representing each construct.

# Results

## Participant Characteristics

The total number of participants in the study was 200. All were nursing students. The majority were female, 98% (*n* = 196), while males made up 2 % of the sample (n = 4). Regarding age, the 18–20-year-old age group represented the majority (72.5%). For the study year, the second year had the highest number of students at 38.5%, followed by the first and third Years at 21% each, and the fourth year at 19.5%, as shown in Table 1.

**Table 1**

*Demographic Characteristics of Participants (*N = *200)*

| Characteristic | *n* | % |
|---|---|---|
| Gender | | |
| Female | 196 | 98.0 |
| Male | 4 | 2.0 |

| Characteristic | *n* | % |
|---|---|---|
| Age (years) | | |
| 18–20 | 145 | 72.5 |
| 21–23 | 55 | 27.5 |
| Study year | | |
| First | 42 | 21.0 |
| Second | 77 | 38.5 |
| Third | 42 | 21.0 |
| Fourth | 39 | 19.5 |

## Confirmatory Factor Analysis

The factor loadings were assessed for each item representing each construct. None of the items was deleted because all factor loadings were more significant than .50, as shown in Figure 1 and Table 2. This indicates that each survey item effectively measured the concept it was intended to assess. The model fit measures used to assess the overall goodness of fit were: CMIN/DF, GFI, adjusted goodness-of-fit index (AGFI), parsimony goodness-of-fit index (PGFI), RMSEA, and SRMR. All the model fit values were within acceptable limits except the RMSEA, which had a value of 0.084, and the GFI and AGFI, which approximated the recommended value, as shown in Table 2. In general, these values suggest that the overall structure of the questionnaire was a good match for the way participants responded. Thus, the model was assumed to fit.

**Table 2**

*Fit Indices for the Arabic Version of the UTAUT-2012 Model*

| Measure | Recommended value | Reference | Estimated model |
|---|---|---|---|
| $\chi^2$ | | | 891.341 |
| Model parameters, *n* | | | 96 |
| Observations, *n* | | | 200 |
| *df* | | | 369 |
| *p*-value | | | 0 |
| CMIN/DF | < 5 | Consiglio et al. (2016) | 2.416 |
| RMSEA | < 0.08 | Hu & Bentler (1998) | 0.084 |
| RMSEA low (90% CI) | | | 0.077 |
| RMSEA high (90% CI) | | | 0.091 |
| GFI | ≥ 0.8 | Hair et al. (2012) | 0.792 |
| AGFI | ≥ 0.8 | Tanaka & Huba (1985) | 0.785 |
| PGFI | ≥ 0.5 | Mulaik et al. (1989) | 0.627 |
| SRMR | < 0.08 | Hu & Bentler (1998) | 0.066 |

*Note*. UTAUT = Unified Theory of Acceptance and Use of Technology; CMIN/DF = chi-square divided by degrees of freedom; RMSEA = root-mean-square error of approximation; CI = confidence interval; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; PGFI = parsimony goodness-of-fit index; SRMR = standardized root-mean-square residual.

## Factor Loadings, Construct Reliability, and Convergent Validity

The consistency of the indicator was assessed and confirmed by all the indicators exhibiting loadings of .6 and above. The construct reliability was assessed using Cronbach's alpha and composite reliability, which is standardized reliability for each construct. Cronbach's alpha for all the constructs was observed to be greater than the required excellent limit of 0.708. The construct validity was measured with average variance extracted (AVE); all the AVE were above the threshold of 0.5. This means that each group of items worked together consistently and accurately reflected the concept they were intended to measure. Therefore, the convergent validity of all the constructs was achieved as shown in Table 3. These results confirm that the Arabic version of the questionnaire is a reliable and valid tool for measuring students' acceptance of AI in nursing education.

**Table 3**

*Results of Factor Analysis for the Relationships Among UTAUT-2012 Questionnaire Constructs*

| Relationship | Outer loading | α | rho_C | AVE |
|---|---|---|---|---|
| Performance expectancy | | 0.881 | 0.882 | 0.653 |
| PE1 | .726 | | | |
| PE2 | .846 | | | |
| PE3 | .829 | | | |
| PE4 | .826 | | | |
| Effort expectancy | | 0.879 | 0.882 | 0.66 |
| EE1 | .703 | | | |
| EE2 | .776 | | | |
| EE3 | .915 | | | |
| EE4 | .839 | | | |
| Social influence | | 0.844 | 0.842 | 0.642 |
| SI1 | .763 | | | |
| SI2 | .807 | | | |
| SI3 | .833 | | | |
| Facilitating condition | | 0.84 | 0.839 | 0.57 |
| FC1 | .697 | | | |
| FC2 | .738 | | | |
| FC3 | .814 | | | |
| FC4 | .768 | | | |
| Hedonic motivation | | 0.888 | 0.888 | 0.727 |
| HM1 | .874 | | | |
| HM2 | .861 | | | |
| HM3 | .822 | | | |
| Price value | | 0.816 | 0.812 | 0.596 |
| PV1 | .732 | | | |
| PV2 | .799 | | | |

| Relationship | Outer loading | α | rho_C | AVE |
|---|---|---|---|---|
| PV3 | .782 | | | |
| Habit | | 0.846 | 0.874 | 0.680 |
| HT1 | .882 | | | |
| HT2 | .911 | | | |
| HT3 | .658 | | | |
| Behavioral intention | | 0.865 | 0.859 | 0.69 |
| BI1 | .701 | | | |
| BI2 | .892 | | | |
| BI3 | .886 | | | |
| Use behavior | | 0.963 | 0.964 | 0.899 |
| UB1 | .927 | | | |
| UB2 | .939 | | | |
| UB3 | .978 | | | |

*Note*. UTAUT = Unified Theory of Acceptance and Use of Technology; AVE = average variance extracted; PE = performance expectancy; EE = effort expectancy; SI = social influence; FC = facilitating condition; HM = hedonic motivation; PV = price value; HT = habit; BI = behavioral intention; UB = use behavior.

## Discriminant Validity

To further affirm the indicators of the measurement model for each construct, discriminant validity was assessed using the heterotrait-monotrait (HTMT) ratio and the Fornell-Larcker criterion. All the values of the HTMT ratio were less than 0.90 as recommended (Ab Hamid et al., 2017) and as shown in Table 4; therefore, all indicators achieved discriminant validity. This indicates that the constructs are distinct from one another, meaning each set of questions measures a different concept. However, the Fornell-Larcker criterion revealed similarities in using behavior and behavioral intention constructs (Ab Hamid et al., 2017). Every construct must have a square root of the AVE that exceeds the correlation value for another construct. Click or tap here to enter text. In comparing other constructs in this study, effort expectancy had a higher association with itself than other constructs, and other constructs showed the same trend except for use of behavior under behavioral intention. This slight overlap suggests a potential connection between students' intentions and their actual technology use, which is common in behavioral research. Table 5 shows the achievement of discriminant validity for other constructs considered in the study using the Fornell-Larcker criterion.

**Table 4**

*Heterotrait-Monotrait (HTMT) Ratio*

| Construct | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Behavioral intention | | | | | | | | | |
| 2. Effort expectancy | .764 | | | | | | | | |
| 3. Facilitating condition | .764 | .82 | | | | | | | |
| 4. Habit | .793 | .618 | .643 | | | | | | |
| 5. Motivation | .71 | .799 | .84 | .531 | | | | | |
| 6. Performance expectancy | .735 | .824 | .802 | .626 | .77 | | | | |
| 7. Price value | .707 | .591 | .637 | .771 | .473 | .587 | | | |

| Construct | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 8. Social influence | .737 | .763 | .767 | .736 | .677 | .732 | .633 | | |
| 9. Use behavior | .854 | .625 | .684 | .698 | .677 | .676 | .614 | .617 | |

**Table 5**

*Fornell-Larcker Criterion*

| Construct | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Behavioral intention | .831 | | | | | | | | |
| 2. Effort expectancy | .717 | .812 | | | | | | | |
| 3. Facilitating condition | .757 | .795 | .755 | | | | | | |
| 4. Habit | .67 | .54 | .557 | .824 | | | | | |
| 5. Motivation | .684 | .779 | .847 | .427 | .853 | | | | |
| 6. Performance expectancy | .75 | .781 | .785 | .559 | .767 | .808 | | | |
| 7. Price value | .707 | .565 | .646 | .723 | .477 | .595 | .772 | | |
| 8. Social influence | .712 | .751 | .766 | .697 | .677 | .727 | .648 | .801 | |
| 9. Use behavior | .855 | .575 | .675 | .599 | .664 | .678 | .618 | .612 | .948 |

# Discussion

The study assessed the psychometric properties of the Arabic version of Venkatesh et al.'s (2012) UTAUT 2012 model for nursing students. This was done to ascertain the reliability and validity of the adapted Arabic-translated version of the instrument. The latent constructs of performance expectation, effort expectation, social influence, facilitating conditions, hedonic motivation, price value, and habit, with behavioral intention and use behavior, optimally loaded, and the factor loadings were more than .5 as recommended (Hair et al., 2012). Furthermore, a good fit was shown using the CFA of the model's nine factors. However, the results showed that the use of behavior and behavioral intention cannot be discriminately validated. This can be attributed to cultural differences.

The study's findings support the significance of the latent constructs in the instrument and align with previous research using the UTAUT 2012 model (Alhamazani, 2020; Alshammari, 2021; Venkatesh et al., 2003, 2012). However, the UTAUT 2012 model's constructs have received significant support from various studies, indicating that they are dynamic and context-dependent rather than static. For example, a Taiwanese study revealed that performance expectancy significantly influenced behavioral intention among nursing students (Lee et al., 2024). Similarly, a study on using the PeduliLindungi app during the COVID-19 pandemic in Indonesia highlighted the favorable impact of social influence, effort expectation, and performance expectation on behavioral intention and usage behavior while facilitating circumstances had no appreciable effect (Akbar et al., 2023). In contrast, facilitating circumstances emerged as a unique predictor of behavioral intention, according to a meta-analytic structural equation modeling analysis aggregating empirical studies using UTAUT in educational contexts. Concurrently, effort expectations and social influence were new predictors of usage behavior (Or, 2023). Research has indicated that several factors, including price value, habit, hedonic motivation, social influence,

performance expectancy, effort expectancy, and facilitating settings, significantly influence students' behavioral intention to use blended learning. Likewise, a study on the adoption of telemedicine by healthcare professionals found that hedonic motivation, habit, and performance expectancy influenced behavioral intention to adopt and use telemedicine systems (Thabet et al., 2023). Finally, a study conducted by Jordanian university faculty members on using digital learning tools discovered a substantial positive association between social effects, performance expectations, effort expectations, and the receptivity level of these tools (Ahmad et al., 2023).

Moreover, the results of this study show that both behavioral intention and use behavior constructs are more correlated with each other than other constructs. Various studies support this result, including an analysis of mobile learning (m-learning; Gonzalez & dos Santos, 2017) and the report of errors in clinical care by nurses in educational and medical organizations (Abry et al., 2022). Lastly, the study of nurses' intention to use blended e-learning systems underscores the need for supportive educational environments to enhance behavioral intentions (Chang et al., 2015).

## Conclusion

This investigation highlights the significance and usefulness of the UTAUT 2012 instrument. The results generally support using the Arabic-translated UTAUT model for the proposed study, which investigates the factors affecting the acceptance of new technologies, views of novel systems, and innovations in healthcare settings. Applying CFA to the Arabic-translated UTAUT 2012 yielded confirmation of a valid and reliable instrument with well-represented latent constructs. This finding supports integrating AI into nursing education, which is crucial for preparing nurses to adapt to the evolving healthcare landscape driven by technology. These findings can help educators and decision-makers in distance learning settings better understand students' readiness for adopting new technologies.

## Acknowledgements

The authors would like to thank the Ongoing Research Funding Program (ORFFT-2025-079-1) at King Saud University, Riyadh, Saudi Arabia, for financial support.

# References

Ab Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (2017). Discriminant validity assessment: Use of Fornell and Larcker criterion versus HTMT criterion. *Journal of Physics: Conference Series, 890*, Article 012163. https://doi.org/10.1088/1742-6596/890/1/012163

Abry, S., Mehrabian, F., Omidi, S., Karimy, M., Kasmaei, P., & Haryalchi, K. (2022). Investigation of factors related to the behavior of reporting clinical errors in nurses working in educational and medical centers in Rasht city, Iran. *BMC Nursing*, *21*, Article 348. https://doi.org/10.1186/s12912-022-01134-3

Abu-Al-Aish, A., & Love, S. (2013). Factors influencing students' acceptance of m-learning: An investigation in higher education. *The International Review of Research in Open and Distributed Learning, 14*(5). https://doi.org/10.19173/irrodl.v14i5.1631

Ahmad, F. B., Al-Nawaiseh, S. J., & Al-Nawaiseh, A. J. (2023). Receptivity level of faculty members in universities using digital learning tools: A UTAUT perspective. *International Journal of Emerging Technologies in Learning*, *18*(13), 209–219. https://doi.org/10.3991/ijet.v18i13.39763

Akbar, G. G., Muchtar, Wahid, A. A., Kania, I., Putri, A. D., & Alamanda, D. T. (2023). Applying UTAUT model to understand use of behavior health applications user in Indonesia during the COVID-19. *Jurnal Ilmu Administrasi: Media Pengembangan Ilmu dan Praktek Administrasi, 20*(1), 29–45. https://jia.stialanbandung.ac.id/index.php/jia/article/view/900

Alenezi, F. Y. (2023). Artificial intelligence versus Arab universities: An enquiry into the Saudi context. *The Scientific Journal of King Faisal University: Humanities and Management Sciences, 24*(1), 82–88. https://doi.org/10.37575/h/edu/220038

Alhamazani, F. H. (2020). Exploratory factor analysis toward adoption of cloud computing by faculty members in Saudi Arabian universities for teaching and learning**.** *International Journal of Computer Science and Network Security, 20*(2), 8–15. https://doi.org/10.22937/IJCSNS.2020.20.2.2

Al-Qeisi, K., Dennis, C., Hegazy, A., & Abbad, M. (2015). How viable is the UTAUT model in a non-Western context? *International Business Research, 8*(2), 204–219. https://doi.org/10.5539/ibr.v8n2p204

Alshammari, S. H. (2021). Determining the factors that affect the use of virtual classrooms: A modification of the UTAUT model. *Journal of Information Technology Education: Research, 20*, 117–135. https://doi.org/10.28945/4709

Bandyopadhyay, K., & Fraccastoro, K. A. (2007). The effect of culture on user acceptance of information technology. *Communications of the Association for Information Systems, 19,* Article 23. https://doi.org/10.17705/1cais.01923

Barchielli, C., Marullo, C., Bonciani, M., & Vainieri, M. (2021). Nurses and the acceptance of innovations in technology-intensive contexts: The need for tailored management strategies. *BMC Health Services Research, 21*, Article 639. https://doi.org/10.1186/s12913-021-06628-5

Chang, T.-F., Chao, C.-M., & Cheng, B.-W. (2015). Framework and verification of a blended e-learning system behavioral intention model among clinical nurses. *Journal of Baltic Science Education*, *14*(6), 733–743. https://files.eric.ed.gov/fulltext/EJ1427489.pdf

Consiglio, C., Borgogni, L., Di Tecco, C., & Schaufeli, W. B. (2016). What makes employees engaged with their work? The role of self-efficacy and employee's perceptions of social context over time. *The Career Development International*, *21*(2), 125–143. https://psycnet.apa.org/record/2016-19688-002

Dikkema, Y., Nieuwenhuis, M. K., van der Schans, C. P., & Mouton, L. J. (2021). Questionnaires to assess facilitators and barriers of early mobilization in critically ill patients: Which one to choose? A systematic review. *Clinical Nursing Research, 30(4),* 442–454. https://doi.org/10.1177/1054773820948268

Drissi, N., Ouhbi, S., Amiri, L., Al Mugaddam, F., Jan, R. K., & Isomursu, M. (2022). A conceptual framework to design connected mental health solutions in the United Arab Emirates: Questionnaire study. *JMIR Formative Research*, *6*(2), Article e27675. https://doi.org/10.2196/27675

Gabriel, R. (2023). Correlations between learning style preferences and Arab-speaking Gulf region first-year college students' EFL performance: A literature review. *Journal of Language Teaching and Research, 14*(3), 709–714. https://doi.org/10.17507/jltr.1403.18

Gonzalez, I. P., Jr., & dos Santos, E. M. (2017). A proposal for UTAUT model extension in the virtual learning environments use as presential learning support context. *International Journal of Technology and Human Interaction, 13*(3), 33–46. https://doi.org/10.4018/IJTHI.2017070103

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*, 414–433. https://doi.org/10.1007/s11747-011-0261-6

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kwak, Y., Seo, Y. H., & Ahn, J. W. (2022). Nursing students' intent to use AI-based healthcare technology: Path analysis using the unified theory of acceptance and use of technology. *Nurse Education Today, 119*, Article 105541. https://doi.org/10.1016/j.nedt.2022.105541

Lee, Y., Lim, W., & Eng, H. S. (2024). A systematic review of UTAUT2 constructs' analysis among MSMEs in non-OECD countries. *Journal of Science and Technology Policy Management*, *15*(4), 765–793. https://doi.org/10.1108/JSTPM-08-2022-0140

McCoy, S., Galletta, D. F., & King, W. R. (2007). Applying TAM across cultures: The need for caution. *European Journal of Information Systems, 16*(1), 81–90.

https://doi.org/10.1057/palgrave.ejis.3000659

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430–445. https://doi.org/10.1037/0033-2909.105.3.430

O'Connor, S., Peltonen, L.-M., Ronquillo, C., Chu, C. H., Topaz, M., Chen, L.-Y. A., & Lee, J. J. (2023, February 20). *A framework for embedding artificial intelligence in nursing: Education, innovation, collaboration, and implementation—The ALIGN Model*. OSF Preprints. https://doi.org/10.31219/osf.io/3b8pc

Or, C. (2023). Revisiting Unified Theory of Technology and Use of Technology using meta-analytic structural equation modelling. *International Journal of Technology in Education and Science*, *7*(1), 83–103. https://doi.org/10.46328/ijtes.420

**Polit, D. F., & Beck, C. T. (2016).** *Nursing research: Generating and assessing evidence for nursing practice* (10th ed.). Lippincott Williams & Wilkins.

Rouibah, K. (2008). Social usage of instant messaging by individuals outside the workplace in Kuwait: A structural equation model. *Information Technology & People, 21*(1), 34–68. https://doi.org/10.1108/09593840810860324

Sidani, S., Guruge, S., Miranda, J., Ford-Gilboe, M., & Varcoe, C. (2010). Cultural adaptation and translation of measures: An integrated method. *Research in Nursing & Health, 33*(2), 133–143. https://doi.org/10.1002/nur.20364

Speyer, R., Cordier, R., Kertscher, B., & Heijnen, B. J. (2014). Psychometric properties of questionnaires on functional health status in oropharyngeal dysphagia: A systematic literature review. *BioMed Research International*, *2014*(1), Article 458678. https://doi.org/10.1155/2014/458678

Tanaka, J. S., & Huba, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, *38*(2), 197–201. https://psycnet.apa.org/record/1986-10882-001

Taskiran, N. (2023). Effect of an artificial intelligence course in nursing on students' medical artificial intelligence readiness: A comparative quasi-experimental study. *Nurse Educator, 48*(5), e147–e152. https://doi.org/10.1097/NNE.0000000000001446

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for the measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Thabet, Z., Albashtawi, S., Ansari, H., Al-Emran, M., Al-Sharafi, M. A., & AlQudah, A. A. (2023). Exploring the factors affecting telemedicine adoption by integrating UTAUT2 and IS success model: A hybrid SEM–ANN approach. *IEEE Transactions on Engineering Management*, *71*, 8938–8950. https://ieeexplore.ieee.org/abstract/document/10195252

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Williams, M. D., Rana, N. P., & Dwivedi, Y. K. (2015). The Unified Theory of Acceptance and Use of Technology (UTAUT): A literature review. *Journal of Enterprise Information Management, 28*(3), 443–488. https://doi.org/10.1108/JEIM-09-2014-0088

Zhou, L. L., Owusu-Marfo, J., Asante Antwi, H., Antwi, M. O., Kachie, A. D. T., & Ampon-Wireko, S. (2019). Assessment of the social influence and facilitating conditions that support nurses' adoption of hospital electronic information management systems (HEIMS) in Ghana using the unified theory of acceptance and use of technology (UTAUT) model. *BMC Medical Informatics and Decision Making, 19*, Article 230. https://doi.org/10.1186/s12911-019-0956-z


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Blended Learning Effectiveness and College Students' Deep Learning Perceptions: The Community of Inquiry Perspective

**Dandan Shen[1], Chiung-Sui Chang[2]*, and Junjie Yang[1]**
*[1]School of Computer Science and Intelligence Education, Lingnan Normal University, Guangdong, China; [2]Department of Educational Technology, Tamkang University, New Taipei, Taiwan; *Corresponding author*

## Abstract

Emerging technologies and innovative instructional methods have revolutionized education, making blended learning the new standard in the artificial intelligence era. However, poor integration of online and face-to-face learning has led to challenges such as superficial student engagement. This study developed a Community of Inquiry-based blended learning model and evaluated its effectiveness with 92 college students using a quasi-experimental approach. Over 16 weeks, the experimental group (n = 48) adopted the blended learning model, while the control group (n = 44) used traditional learning conditions. Learning effectiveness and deep learning perceptions were evaluated, revealing the blended learning group demonstrated superior learning effectiveness (d = 0.83) and reported higher deep learning perceptions ($\eta 2$ = .05–.072) compared to the traditional learning group. These results provide valuable insights for educators aiming to design blended learning models that foster deep learning and improve overall learning effectiveness.

*Keywords*: blended learning, learning effectiveness, deep learning perception, Community of Inquiry, CoI, higher education

# Introduction

The widespread adoption of digital technologies has transformed education, leading to the rise of diverse instructional methods such as online learning and personalized instruction, which have spurred the growth of blended learning (Perera et al., 2020). Blended learning, which combines the strengths of traditional and online education, addresses the limitations of both methods and plays a crucial role in enhancing teaching quality and talent development (Han, 2023). In such environments, educators guide students in cultivating creativity, problem-solving, and deep learning skills, all essential for higher education reform and the 21st-century workforce (Qi et al., 2020).

The Community of Inquiry (CoI) framework is an effective model for designing blended environments, promoting learning engagement through active participation, meaningful discussion, and timely instructor feedback (Armellini et al., 2021; Liu & Deris, 2022). The three key elements of the CoI framework—teaching presence, social presence, and cognitive presence—are vital for enhancing students' knowledge acquisition, skill development, and motivation (Zhang, 2020).

Empirical research has suggested that students in blended learning environments tend to achieve superior learning outcomes and report more positive learning experiences compared to those in traditional classroom settings (Vallée et al., 2020). The CoI framework further supports blended course design by fostering collaboration, critical thinking, and meaningful knowledge construction (Zhang, 2020). Despite these promising findings, further research is required to determine whether CoI-based blended learning is more effective than traditional pedagogical methods in enhancing both learning outcomes and deep learning. This study addressed this gap by investigating the effectiveness of a CoI-based blended learning model, facilitated through the Superstar Learning platform (https://www.chaoxing.com/), in promoting high-quality learning experiences. By examining the impact of this model, the study contributes to the growing body of research on evidence-based frameworks for blended learning, offering valuable insights for educators and instructional designers seeking to optimize student engagement and academic achievement.

# The Blended Learning Effect

Blended learning is widely recognized for improving learning outcomes, especially when using diverse information and communications technology tools and resources that create immersive, student-centered environments (Bizami et al., 2023; Shamir-Inbal & Blau, 2021). These environments effectively enhance learning by integrating traditional and digital methods. For instance, Li and Cao's (2020) blended learning model for undergraduate English students combined virtual reality, online content, and interactive classroom sessions, significantly improving test scores. Similarly, Antonelli et al. (2023) found that virtual reality labs reduced student errors and enhanced practical skills. Through a systematic review, Almusaed et al. (2023) confirmed that blended learning environments enabled with artificial intelligence (AI) have the potential to enhance educational quality. Technologies such as chatbots, intelligent tutoring systems, and personalized learning platforms can increase student engagement and support motivation.

Blended learning also fosters deep learning competencies. Innovative models, such as micro-learning and computer-based collaborative learning, have improved students' self-learning abilities and goal achievement (Astiwardhani Sobandi, 2024). Courses incorporating video lectures and online materials have boosted English learners' motivation, autonomy, and satisfaction (Wang et al., 2021). Zhao (2022)

leveraged the Rain Classroom platform's network and mobile technologies (https://www.yuketang.cn/en) to create a blended model that increased learning quality and interest. Similarly, AI-powered blended courses have been more effective than traditional teaching in promoting student engagement, deep learning, critical thinking, and independent and cooperative learning (He et al., 2023).

## Deep Learning Dimensions

Unlike surface learning, which focuses on memorization driven by external motivation, deep learning fosters meaningful knowledge construction and intrinsic motivation (Darling-Hammond & Oakes, 2021). Deep learning encourages learners to engage deeply with content, develop critical thinking and problem-solving skills, and enhance knowledge transfer (Shen & Chang, 2023). It also promotes skills such as learning how to learn, complex problem-solving, and effective communication and collaboration (Mthethwa-Kunene et al., 2022).

At the core of deep learning is the development of higher-order cognitive abilities, including problem-solving, decision-making, and critical thinking (Kurniawan, 2021). This process empowers students to creatively solve real-world problems (Pan et al., 2023). Mobile technology has been shown to support higher-order thinking by providing timely access to information and facilitating active cooperation and engagement (Hye et al., 2020; Yaniawati et al., 2022). Hu and Hwang (2024) proposed a problem-posing approach that was mobile, self-adapted, and based on concept mapping within a virtual museum context and found that it significantly enhanced learners' critical thinking and problem-solving abilities. In this study, learning behaviors involving application, analysis, evaluation, and creation were categorized as higher-order cognition.

Deep learning also emphasizes communication and collaboration, essential 21st-century skills (Mthethwa-Kunene et al., 2022). Effective communication involves organizing learning content through presentations, group work, and interactive projects, while collaboration focuses on learner-centered activities such as sharing, interaction, and discussion (Islam et al., 2022; Mthethwa-Kunene et al., 2022). Interactive learning is supported by flipped classroom models that promote active communication through online and in-person discussions and timely feedback (Shen & Chang, 2023). Blended learning, incorporating diverse computer-mediated collaboration strategies, enhances participation and student satisfaction (Belda-Medina, 2021; Vlachopoulos & Makri, 2019). Intelligent computers can simulate interactions between learners and the external environment by performing tasks and providing timely feedback based on individual mistakes. Additionally, collaboration between AI and learners has the potential to optimize learning outcomes (Weber et al., 2025).

Reflective thinking is another key aspect of deep learning, focusing on self-reflection in academic research and practice (Rogers et al., 2019). It plays a crucial role in shaping students' professional identity and ensuring sustainable learning (Annansingh, 2019). Reflective learning involves both self-reflection and reflection on others' learning behavior (Priddis & Rogers, 2018). Intelligent learning tools—such as ChatGPT, Apple's Shortcuts, and LINE—have been shown to facilitate this process by enabling students to monitor their progress and develop reflective thinking skills (Wu et al., 2023). In this study, reflective learning primarily involved learning under supervision and guidance.

Lastly, emotional experience is a hallmark of deep learning (Pellegrino & Hilton, 2012). Self-efficacy, motivation, and positive attitudes toward learning significantly impact engagement and effectiveness

(Meyer et al., 2018). Well-designed online and face-to-face activities in blended courses can foster positive attitudes, while clear goals and challenging tasks can enhance self-efficacy and motivation (Zhao & Song, 2022; Zhao et al., 2021). In this study, emotional experiences included learning attitudes, self-efficacy, and motivation.

## Research Hypotheses

This study operationally defined four core features of deep learning: higher-order cognition, interactive learning, reflective learning, and emotional experience. It measured students' deep learning perceptions in these four dimensions to verify the effectiveness of the CoI-based blended learning model. Furthermore, this study, using the Digital Film and Television Directing and Production course as an example, applied a CoI-blended learning model and compared the blended (BL) and traditional (TL) classroom learning outcomes represented by three scores: online quizzes, a final exam, and film and television assignments.

While prior research has generally supported the efficacy of blended learning, some studies have reported no significant differences in learning outcomes between BL and TL students (e.g., Müller & Mildenberger, 2021). These discrepancies may stem from inadequate instructional design, insufficient scaffolding, superficial integration of online and offline learning, and underuse of the potential benefits of blended learning. The CoI framework, which focuses on enhancing students' learning presence, can effectively guide the design of blended learning activities, fostering a more seamless integration of online and face-to-face learning modes (Armellini et al., 2021). CoI-based blended courses are expected to promote deep learning and improve overall learning effectiveness by balancing online and offline instruction (Tabassum & Mohd Saad, 2024).

Thus, the following hypotheses were proposed:

- H1: BL students will exhibit significantly higher learning effectiveness than TL students.

    H1-1a: BL students will achieve higher scores on lower-order online quiz questions.

    H1-1b: BL students will achieve higher scores on higher-order online quiz questions.

    H1-2a: BL students will achieve higher scores on lower-order final exam questions.

    H1-2b: BL students will achieve higher scores on higher-order final exam questions.

    H1-3: BL students will score significantly higher on film and television assignments.

- H2: BL students will demonstrate stronger deep learning perceptions than TL students across the following dimensions:

    H2-1: Higher-order cognition

    H2-2: Interactive learning

    H2-3: Reflective learning

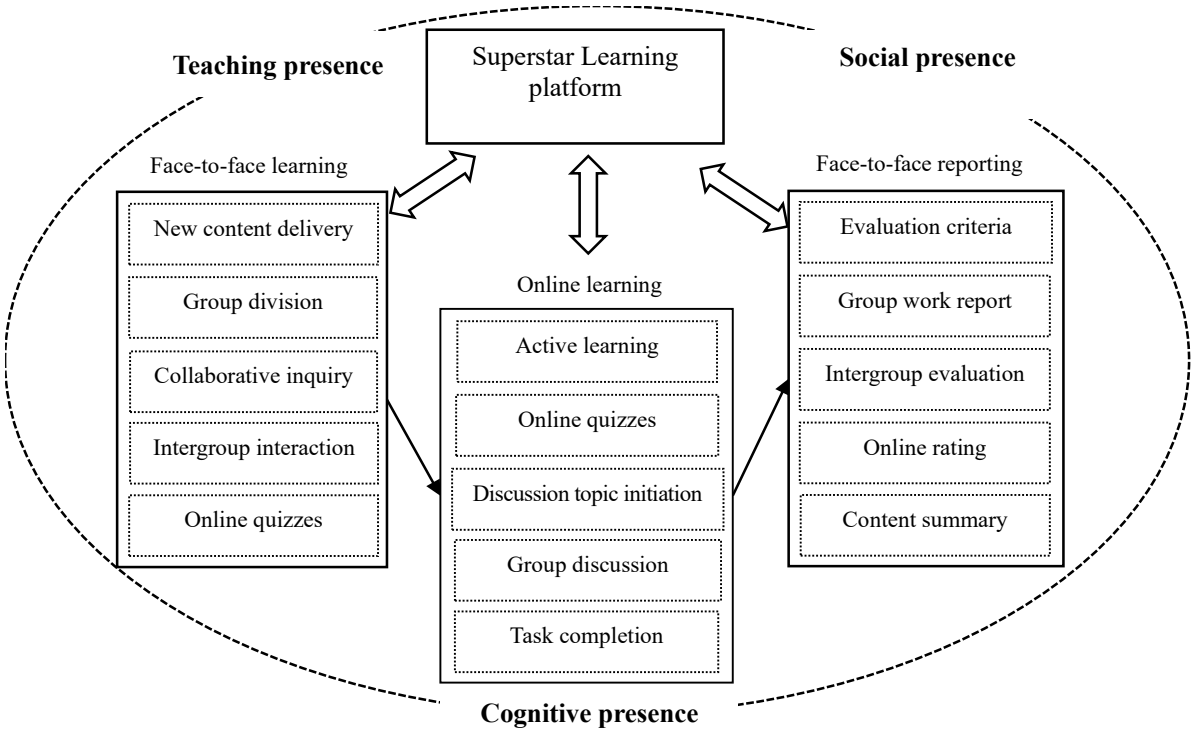
H2-4: Emotional experience

## Community of Inquiry-Based Blended Learning Model Design

This study examined the implementation of a blended learning model in the Digital Film and Television Directing and Production course, using the Superstar Learning platform developed by Beijing Century Superstar Information. The model integrated three key elements of the community of inquiry (CoI) framework: teaching, social, and cognitive presence, within three learning stages—face-to-face learning, online learning, and face-to-face reporting (Figure 1). The 16-week course was divided into three phases: early creation, mid-stage shooting, and post-editing, with the initial week dedicated to orientation and the remaining time equally distributed across the three teaching modules.

**Figure 1**

*The CoI-Based Blended Learning Model*




*Note*. CoI = Community of Inquiry.

The course design and implementation effectively addressed the three elements of the Community of Inquiry (CoI) framework. In terms of teaching presence, the course featured clearly defined learning objectives, detailed syllabi, and well-organized materials, while instructors actively provided timely feedback and facilitated problem-solving discussions. The intelligent learning platform further enabled tailored instructional strategies by monitoring student progress and adjusting content delivery accordingly.

Regarding social presence, the curriculum incorporated collaborative activities, such as online discussion forums, group projects, and peer reviews, which fostered a sense of community among

students. Real-time communication tools and virtual breakout sessions were employed to enhance interaction and support, while integrated social networking features promoted informal exchanges of ideas beyond formal class sessions.

In addressing cognitive presence, the course used inquiry-based tasks, case studies, and problem-solving assignments to encourage critical analysis and reflective thinking. Interactive modules and scenario-based exercises were implemented to deepen understanding and facilitate the application of theoretical concepts, while continuous reflection activities, including reflective journals and self-assessment quizzes, supported the development of critical thinking and knowledge construction.

## Face-to-Face Learning Stage

In this stage, in-class learning was paired with the Superstar Learning platform. The first two weeks of each module followed a structured five-step process:

1. The teacher introduced new concepts, encouraging active participation.
2. Students formed self-selected groups and learning behavior norms for participation.
3. Inquiry tasks were assigned, prompting group discussions and collaborative exploration. Results were uploaded to the platform.
4. Groups conducted peer evaluations, refining their work based on feedback, with guidance from the teacher.
5. Online quizzes assessed learning outcomes, followed by teacher feedback and content summaries.

## Online Learning Stage

This stage featured three online discussion questions and one practice task, each spanning 2 weeks and following five steps:

1. The teacher provided learning materials (video, PowerPoints, quizzes) on the platform for students to engage with independently.
2. Students completed online quizzes with immediate feedback, and the teacher provided support when needed.
3. The teacher initiated discussions, with students tackling two discussion questions in the first week and one in the second.
4. Group leaders supervised discussions, fostering deeper engagement and collaboration, while the teacher facilitated reflection and knowledge co-construction.
5. Groups completed tasks based on discussion outcomes and uploaded their work to the platform.

### Face-to-Face Reporting Stage

This stage focused on the online learning tasks and summarizing content. The last week of each learning module followed a five-step process.

1. The teachers shared evaluation criteria, and students familiarized themselves with the guidelines.

2. Group representatives presented their work, showcasing innovative ideas and application of new concepts. Other groups and the teacher provided feedback.

3. Intergroup evaluations were conducted, with students providing feedback based on evaluation criteria and engaging in on-site discussions.

4. Groups posted online ratings and comments on other groups' work, and the teachers stored the evaluation results for post-class review.

5. The teachers guided students through content summarization, encouraging critical thinking and integration of new material

## Research Methodology

### Participants

The participants were 92 sophomore educational technology students enrolled in the two compulsory Digital Film and Television Production classes at a university in China in the spring semester of 2023. Both classes lasted for 16 weeks and were taught by the same teacher. One class was randomly assigned as the experimental group ($n$ = 48), while the other class was the control group ($n$ = 44). The male-female ratio for the experimental and control groups were 1:2 (16 males and 32 females) and 1:2.4 (13 males and 31 females), respectively. All students were between 19 and 21 years old, owned smartphones, and were proficient Superstar Learning platform users.

The final exam scores from the prerequisite course Fundamentals of Photography Skills served as the prior knowledge assessment criteria. Independent sample $t$-test results indicated that there were no statistically significant prior knowledge differences ($t$ = 0.86, $p$ > .05) between the experimental group, $M$($SD$) = 81.38(7.47), and the control group, $M$($SD$) = 80.11(6.59).

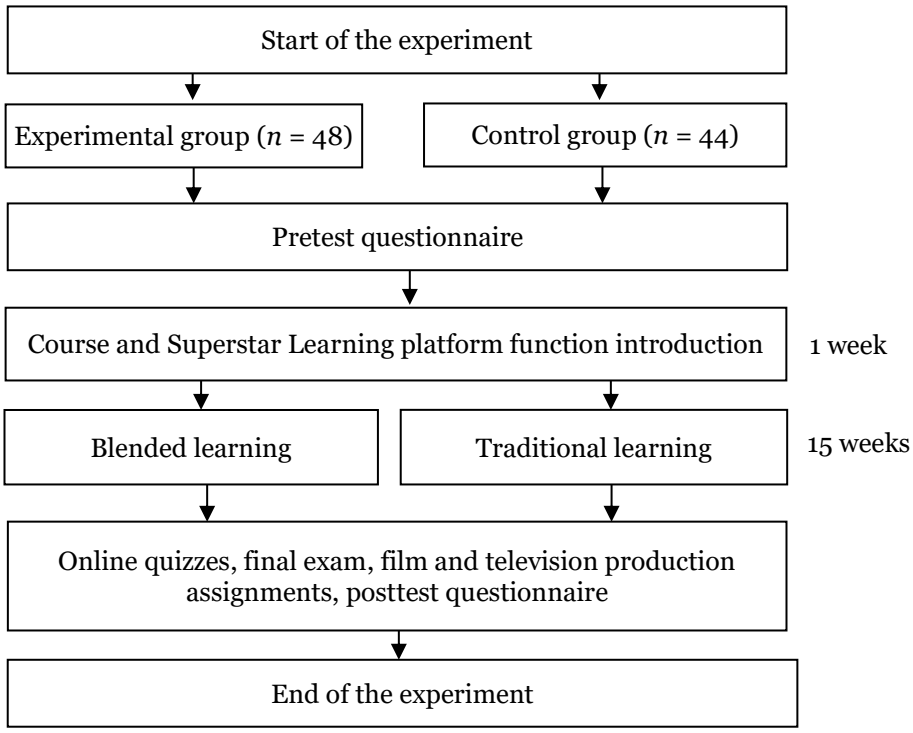
### Experimental Procedure

This study focused on learning effectiveness and students' deep learning perceptions in the CoI-based blended learning course (Figure 2). In the pre-intervention stage, the participants took the pretest and were introduced to the basic functions of the Superstar Learning platform. In the intervention stage, while both groups participated in face-to-face learning with some Superstar Learning activities done in class, only the experimental group took part in the online learning stage. The control group students listened to the teacher's lectures, participated in class discussions, collaborated in groups, completed online quizzes, and completed the same post-class learning tasks. As the Digital Film and Television Directing and Production course focused on both theory and practice, students in both groups were required to take the final exam and submit comprehensive film and television production assignments. At the end of the

course, both groups completed the perception posttest.

**Figure 2**

*Experimental Procedure*




## Instruments

This study used various assessments to compare learning effectiveness between students in blended and traditional learning environments: online quizzes, a final exam, and group assignments. Additionally, pre- and posttests were conducted to measure deep learning perceptions in both groups.

### *Online Quizzes*

On the Superstar Learning platform, both groups completed identical online quizzes, worth 280 points in total, consisting of 64 single-choice questions (128 points), 26 multiple-choice questions (104 points), and 6 essay questions (48 points). Based on Bloom's revised taxonomy (Krathwohl, 2002), the quizzes were divided into lower-order (knowledge and comprehension) and higher-order (application, analysis, and evaluation) cognitive levels. Lower-order questions totaled 186 points, while higher-order questions totaled 94 points.

### *Final Exam*

Both groups took the same final exam, categorized into lower-order and higher-order cognitive questions. For instance, a lower-order question required students to briefly describe the fundamental principles of shot assembly, whereas a higher-order question asked them to provide an example of how the close-up method is applied in effect sound processing. The exam was worth 100 points, with 10 fill-in-the-blank questions (20%), 15 multiple-choice questions (30%), 5 brief questions (35%), and 1 design

question (15%). Each cognitive level accounted for 50 points. Content validity was reviewed by three educational technology experts with over 10 years of experience.

## *Film and Television Production Assignments*

Students worked in groups to produce 10-minute films on self-chosen themes. Each class was divided into six groups. To assess individual contributions, students provided task descriptions in their final reports. The total score (100 points) consisted of intragroup scores (30%), intergroup score (30 points), and the teacher's evaluation (40%).

## *Deep Learning Perception Scale*

The questionnaire comprised two sections. The first section collected students' basic information, including student number, gender, and experience with the Superstar Learning platform. The second section contained the 20-item Deep Learning Perception Questionnaire (DLPQ), designed to measure students' perceptions of deep learning. Items related to higher-order cognition, interactive learning, and reflective learning were adapted from Shen and Chang (2022), while emotional experience items were synthesized from Alqurashi (2019) and Zimmerman and Kulikowich (2016). Responses were scored on a 5-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). To ensure content validity, three experts with over 10 years of experience in educational technology reviewed the scale, refining item structure, clarifying meanings, and offering revision suggestions. The finalized questionnaire was pretested with 156 students informed of the study's purpose, yielding 122 valid responses and a return rate of 78.2%.

Exploratory factor analysis (EFA) was conducted using IBM SPSS Statistics (Version 25.0) to evaluate the DLPQ's factor structure. Results are shown in Table 1. The Kaiser-Meyer-Olkin value was 0.83, and Bartlett's chi-square value was 1573.67 ($df = 190$, $p < .001$), indicating the scale's suitability for factor analysis. Principal component analysis with the maximum variance method was employed to analyze the pretest data. Factors were selected based on eigenvalues greater than 1, with item retention requiring factor loadings exceeding 0.5. The analysis revealed four dimensions, explaining a cumulative variance of 66.78% (see Table 1). The interactive learning dimension (5 items, e.g., "I can offer useful suggestions for my peers") accounted for close to half the variance, with factor loadings ranging from 0.64 to 0.86. The reflective learning dimension (5 items, e.g., "I can analyze the reasons for failure to solve problems") explained around another one tenth of the variance, with loadings from 0.51 to 0.83. The higher-order cognition dimension (5 items, e.g., "I can find logical relationships between knowledge points") accounted for an additional small percentage of the variance, with loadings between 0.64 and 0.75. Finally, the emotional experience dimension (5 items, e.g., "I am willing to adjust my learning style to meet course requirements") explained a further small fraction of the variance, with loadings from 0.60 to 0.83. These findings indicate strong structural validity for the scale. Additionally, Cronbach's alpha values for the four dimensions demonstrated good reliability.

**Table 1**

*The Reliability and Validity of DLP Scale*

| Item | Dimension | | | |
|---|---|---|---|---|
| | Interactive learning | Reflective learning | Higher-order cognition | Emotional experience |
| HC1 | | | 0.69 | |
| HC2 | | | 0.75 | |
| HC3 | | | 0.75 | |
| HC4 | | | 0.73 | |
| HC5 | | | 0.64 | |
| IL6 | 0.64 | | | |
| IL7 | 0.74 | | | |
| IL8 | 0.86 | | | |
| IL9 | 0.69 | | | |
| IL10 | 0.77 | | | |
| RL11 | | 0.62 | | |
| RL12 | | 0.83 | | |
| RL13 | | 0.72 | | |
| RL14 | | 0.72 | | |
| RL15 | | 0.51 | | |
| AE16 | | | | 0.67 |
| AE17 | | | | 0.67 |
| AE18 | | | | 0.68 |
| AE19 | | | | 0.60 |
| AE20 | | | | 0.83 |
| Cronbach's $\alpha$ | 0.88 | 0.86 | 0.83 | 0.87 |
| Eigenvalues | 8.84 | 1.83 | 1.52 | 1.17 |
| Variance (%) | 44.19 | 9.13 | 7.62 | 5.84 |
| Cumulative variance (%) | 44.19 | 53.32 | 60.94 | 66.78 |

*Note*. DLP = deep learning perception; HC = higher-order cognition; IL = interactive learning; RL = reflective learning; AE = emotional experience

# Data Collection and Analysis

## Learning Effectiveness

The learning effectiveness data came from online quizzes (30%), film and television assignments (30%), and a final exam (40%). For online quizzes, the objective questions (single-choice, multiple-choice) were automatically scored by the Superstar Learning platform, while the teacher scored the subjective questions (essay questions). The assignment scores comprised the weighted average of the students' and teacher's scores. The final exam papers were graded solely by the teacher. An independent sample *t*-test was used to explore the differences between the two groups in the lower-order cognitive and higher-order cognitive questions on the online quiz and final exam scores, as well as in the assignment

scores.

A significance level of $p < .05$ was adopted, and Cohen's $d$ was used to measure the effect size to evaluate further differences between the experimental and the control groups. Cohen's $d$ values of 0.2~0.5 represented a small effect size, 0.5~0.8 denoted a medium effect size, and values greater than 0.8 indicated a large effect size (Cohen, 1988).

### Deep Learning Perception

Formal questionnaires were administered to investigate the pretest and posttest deep learning perceptions in both the experimental and the control groups. A total of 92 students completed the questionnaire. Out of 48 students in the experimental group, 46 submitted valid questionnaires, resulting in a high return rate of 95.8%. Out of 44 students in the control group, 41 submitted valid questionnaires, with a robust return rate of 93.2%.

Descriptive statistics were used to analyze the pre- and posttest data probing students' higher-order cognition, interactive learning, reflective learning, and emotional experience. Mean and standard deviation were employed to assess students' deeper learning perceptions in four dimensions. Subsequently, covariate analysis was conducted to examine the deep learning perception differences between the experimental and the control groups. η2 was reported to calculate the effect size, with values .01~.07 indicating a small effect size, values .07~.14 denoting a moderate effect size, and values greater than .14 signifying a large effect size (Cohen, 1992).

## Research Results

### Learning Effectiveness

There was a significant difference in the learning effectiveness between the experimental and the control groups (Table 2). The experimental group achieved significantly better learning outcomes than the control group, with a large effect size. These findings support hypothesis H1.

**Table 2**

*Learning Effectiveness Differences*

| Score source | Experimental ($n$ = 48) | | Control ($n$ = 44) | | $t$ | Cohen's $d$ |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | | |
| Lower-order questions (Online quizzes) | 179.63 | 5.74 | 177.41 | 7.04 | 1.66 | 0.35 |
| Higher-order questions (Online quizzes) | 86.67 | 6.21 | 82.41 | 8.22 | 2.78** | 0.58 |
| Lower-order questions (Final exam) | 36.88 | 6.98 | 34.14 | 6.36 | 1.96 | 0.41 |
| Higher-order questions | 36.54 | 6.82 | 32.39 | 5.09 | 3.29*** | 0.69 |

| Score source | Experimental ($n$ = 48) | | Control ($n$ = 44) | | $t$ | Cohen's $d$ |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | | |
| (Final exam) | | | | | | |
| Film and television assignments | 88.81 | 3.94 | 86.86 | 4.14 | 2.31* | 0.48 |
| Learning effectiveness | 84.54 | 5.35 | 80.51 | 4.34 | 3.95*** | 0.83 |

*Note.* $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Also shown in Table 2, there were no significant differences between the experimental and control groups in lower-order question scores for online quizzes and final exam scores, indicating that hypotheses H1-1a and H1-2a were not supported. However, the experimental group scored significantly higher than the control group on higher-order questions in the online quizzes, with a moderate effect size. Similarly, for the final exam, the experimental group outperformed the control group on higher-order question scores, thereby supporting hypotheses H1-1b and H1-2b.

In terms of film and television assignment scores, Table 2 shows a significant difference. The mean assignment score for the experimental group was significantly higher than that of the control group. Although the effect size was small, these findings suggest that students in the experimental group achieved significantly better performance in practical tasks compared to those in the control group, thus supporting hypothesis H1-3.

## Deep Learning Perceptions

This study assessed deep learning perceptions across the dimensions of higher-order cognition, interactive learning, reflective learning, and emotional experience both before and after the completion of the blended course. The findings are shown in Table 3. Results indicated that prior to the course, all students' deep learning competencies in these dimensions were above average, with scores ranging from 3 to 4. After engaging in the blended course, these competencies approached a score of 4, reflecting a significant enhancement in the deep learning abilities of the experimental group. In contrast, the control group, in post-test results, showed less of an improvement, suggesting that the blended learning model played a role in enhancing deep learning competencies.

**Table 3**

*Deep Learning Perceptions in the Experimental and Control Groups*

| Dimension | | Experimental | | Control | |
|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* |
| Higher-order cognition | Pretest | 3.59 | 0.57 | 3.60 | 0.46 |
| | Posttest | 3.96 | 0.55 | 3.76 | 0.46 |
| Interactive learning | Pretest | 3.52 | 0.50 | 3.59 | 0.56 |
| | Posttest | 3.90 | 0.56 | 3.73 | 0.44 |
| Reflective learning | Pretest | 3.48 | 0.49 | 3.50 | 0.49 |

| Dimension | | Experimental | | Control | |
|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* |
| Emotional experience | Posttest | 3.89 | 0.49 | 3.73 | 0.41 |
| | Pretest | 3.49 | 0.49 | 3.54 | 0.49 |
| | Posttest | 3.95 | 0.53 | 3.77 | 0.38 |

Covariate analysis was performed to assess the model's effectiveness further and to elucidate the disparities in higher-order cognition, interactive learning, reflective learning, and emotional experience between the two groups. The results of the intragroup regression coefficient homogeneity test revealed that none of the dimensions reached the significance level: for high-order cognition, $F = 3.23$, $p > .05$; for interactive learning, $F = 0.04$, $p > .05$; for reflective learning, $F = 0.89$, $p > .05$; and for emotional experience, $F = 0.26$, $p > .05$. Consequently, the null hypotheses were accepted, aligning with the assumption of homogeneity of regression coefficients in covariate analysis, justifying the continuation of covariate analysis.

Covariate analyses results revealed significant perception differences between the two groups. Specifically, significant differences were found in higher-order cognition, interactive learning, reflective learning, and emotional experience perceptions, revealing small effect sizes in the first three dimensions (tables 4, 5, and 6) and moderate effect size in the last dimension (Table 7). These findings suggest that students enrolled in the CoI-based blended course had significantly more positive higher-order cognition, interactive learning, reflective learning, and emotional experience perceptions than their traditional classroom counterparts, thereby supporting hypotheses H2-1, H2-2, H2-3, and H2-4.

**Table 4**

*Higher-Order Cognition Perception Differences*

| Source | *SS* | *df* | *MS* | *F* | $\eta^2$ |
|---|---|---|---|---|---|
| Intervention | 8.99 | 1 | 8.99 | 61.32*** | .42 |
| Experimental effect | 0.62 | 1 | 0.62 | 4.25* | .05 |
| Residuals | 12.32 | 84 | 0.15 | | |
| Interpreted sum | 21.88 | 86 | | | |

*Note.* $R^2 = .44$ (Adjusted $R^2 = .42$).
*$p < .05$. ***$p < .001$.

**Table 5**

*Interactive Learning Perception Differences*

| Source | *SS* | *df* | *MS* | *F* | $\eta^2$ |
|---|---|---|---|---|---|
| Intervention | 8.76 | 1 | 8.76 | 59.26*** | .41 |
| Experimental effect | 0.69 | 1 | 0.69 | 4.64* | .05 |

| Source | *SS* | *df* | *MS* | *F* | $\eta^2$ |
|---|---|---|---|---|---|
| Residuals | 12.41 | 84 | 0.15 | | |
| Interpreted sum | 21.56 | 86 | | | |

*Note*. $R^2$ = .42 (Adjusted $R^2$ = .41).
$^{*}p < .05$. $^{***}p < .001$.

**Table 6**

*Reflective Learning Perception Differences*

| Source | *SS* | *df* | *MS* | *F* | $\eta^2$ |
|---|---|---|---|---|---|
| Intervention | 7.91 | 1 | 7.91 | $77.57^{***}$ | .48 |
| Experimental effect | 0.44 | 1 | 0.44 | $4.31^{*}$ | .05 |
| Residuals | 8.57 | 84 | 0.10 | | |
| Interpreted sum | 16.84 | 86 | | | |

*Note*. $R^2$ = .49 (Adjusted $R^2$ = .48).
$^{*}p < .05$. $^{***}p < .001$.

**Table 7**

*Emotional Experience Perception Differences*

| Source | Sum of squares | df | Mean of squares | F | $\eta^2$ |
|---|---|---|---|---|---|
| Intervention | 6.47 | 1 | 6.47 | $45.43^{***}$ | .35 |
| Experimental effect | 0.93 | 1 | 0.93 | $6.54^{*}$ | .072 |
| Residuals | 11.96 | 84 | 0.14 | | |
| Interpreted sum | 19.15 | 86 | | | |

*Note*. $R^2$ = .38 (Adjusted $R^2$ = .36).
$^{*}p < .05$. $^{***}p < .001$.

# Discussion

## Students' Learning Effectiveness

The findings support hypothesis H1, indicating that students in the blended learning (BL) group demonstrated significantly better learning effectiveness than those in the traditional learning (TL) group. This aligns with previous studies, such as Vallée et al. (2020), which found that students in blended learning environments exhibited superior learning outcomes compared to their traditional counterparts. Similarly, Yin and Yuan (2021, 2022) concluded that CoI-based blended learning models not only increased learning interest but also positively impacted academic performance, suggesting that the structure of blended learning fosters deeper engagement and critical thinking, which can enhance learning quality beyond traditional methods (Vo et al., 2017).

For hypotheses H1-1a and H1-2a, no significant difference was observed between the BL and TL groups on lower-order cognitive questions. This finding may be attributed to the independent learning ability required for mastering lower-order cognitive content, where learning methods appear to have less impact. This result is consistent with Shen and Chang (2023), who found that the instructional approach had limited influence on students' mastery of lower-order cognitive knowledge. Similarly, Lozano-Lozano et al. (2020) reported no significant differences between BL and TL students in their comprehension of theoretical knowledge.

In contrast, hypotheses H1-1b and H1-2b were supported, with BL students scoring significantly higher on higher-order cognitive questions compared to their TL peers. This confirms that blended learning has a substantial impact on students' understanding of higher-order cognitive knowledge, consistent with prior research on CoI-based models. Blended learning promotes deeper cognitive engagement, allowing students to move beyond superficial learning and engage in higher-order thinking, leading to improved academic performance and greater understanding of complex concepts (Chen, 2022; Guo et al., 2021). The integration of inquiry-based learning and open communication within the CoI framework enhances social presence, which, in turn, positively affects students' performance on higher-order cognitive tasks (Tan, 2021). These findings emphasize the importance of designing blended learning activities that foster teaching, cognitive, and social presence to enhance higher-order cognitive skills (Kurniawan, 2021). Given the distinct advantage of BL in fostering higher-order thinking, educators should carefully consider the cognitive complexity of the content when designing such courses.

Hypothesis H1-3 was also supported, as BL students demonstrated significantly better practical skills compared to TL students. The substantial impact of blended learning on students' application abilities can be attributed to the opportunities for hands-on practice and experimentation that BL environments provide. Virtual simulations, online laboratories, and other digital tools in blended settings offer students interactive experiences that aid in mastering practical skills (Antonelli et al., 2023). For instance, combining offline and virtual experiments have been shown to improve student performance by enabling faster knowledge acquisition and exploration (Kerimbayev et al., 2023). Virtual laboratories offer safe, interactive platforms for students to observe phenomena that may not be possible in physical settings (Seifan et al., 2020). These immersive environments, along with enhanced communication and teamwork facilitated by blended learning, promote collaboration, knowledge sharing, and innovative problem-solving (Li & Cao, 2020). Furthermore, Liu et al. (2021) found that a blended model incorporating computer-assisted self-study, group discussion, and simulation training significantly improved students' practical application skills.

Overall, the findings suggest that blended learning positively impacts students' ability to apply knowledge in practical contexts, fostering the development of essential operational skills through enriched learning environments and collaborative opportunities.

## Students' Deep Learning Perception

The research findings robustly support hypotheses H2-1, H2-2, H2-3, and H2-4, indicating that students in the BL group had significantly higher perceptions across the four dimensions of higher-order cognition, interactive learning, reflective learning, and emotional experience compared to their TL counterparts. This is consistent with previous studies showing that CoI-based blended courses facilitate deeper learning (Tabassum & Mohd Saad, 2024; Zhang, 2020).

For hypothesis H2-1, the results indicated that BL students exhibited significantly better perceptions in the higher-order cognitive dimension. This finding highlights the model's effectiveness in promoting engagement in advanced learning activities, such as application, analysis, evaluation, and creation. Yaniawati et al. (2022) similarly observed significant improvements in students' creative thinking and problem-solving skills within mobile-based blended learning environments. The intelligent pedagogy framework, which integrates various learning approaches, also showed enhanced higher-order thinking capabilities among students (Meng & Zhang, 2020). Notably, during the COVID-19 pandemic, blended learning became a normative approach that effectively supported higher-order thinking skills. Well-designed blended learning models leverage both online and offline modalities, enhancing cognitive abilities and fostering deeper engagement in learning activities.

In terms of hypothesis H2-2, BL students reported significantly higher perceptions of interactive learning. This can be attributed to the flexible nature of online components within blended courses, which provided opportunities for interaction and allowed students to engage with in-depth discussions at their own pace (Wut et al., 2022). Castro (2019) emphasized the role of innovative technologies in facilitating effective student interaction in higher education blended learning. Enhanced interactive opportunities in blended learning environments improve communication, collaboration, and knowledge sharing, thereby enriching students' cognitive and interactive capabilities (Islam et al., 2022).

Regarding hypothesis H2-3, the BL group exhibited enhanced reflective learning perceptions, suggesting improved recognition of learning challenges and engagement in reflective practices. This improvement may stem from the facilitated problem presentation and discussion typical of blended environments, which foster self-regulation and reflection (Bizami et al., 2023; Wu et al., 2023). Zhu and Bonk (2019) noted that feedback mechanisms and reflective questions significantly promote self-monitoring and reflection. Additionally, real-time technology-supported interactions likely motivate learners and enhance engagement, thereby improving reflective practices and overall learning quality (Zhu et al., 2021).

For hypothesis H2-4, the BL group demonstrated superior perceptions of emotional experiences, with a moderate effect size, indicating the positive impact of blended learning on the affective domain. This enhancement can be linked to personalized learning interventions that improve students' attitudes, motivation, and self-efficacy (Zhang et al., 2020). Ballouk et al. (2022) highlighted the role of teaching guidance in fostering active learning and motivation in blended environments. A supportive, user-friendly blended learning context—augmented by accessible digital tools—transforms students' perceptions of learning, leading to positive emotional experiences (Masadeh, 2021). The availability of diverse learning resources, such as quizzes and case studies, further enhances self-efficacy (Prifti, 2022). Consequently, blended learning environments shift away from traditional teaching paradigms, providing enriched emotional experiences that contribute to improved learning outcomes.

# Conclusions

This study examined the effectiveness of a Community of Inquiry (CoI)-based blended learning model implemented through an intelligent learning platform, revealing significant improvements in both students' learning outcomes and their deep learning perceptions. These findings affirm the model's efficacy. Educators designing such models should capitalize on the strengths of both online and offline learning to foster robust teaching, social, and cognitive presences. Using digital resources enhances

students' access to information and supports creative learning. Additionally, providing targeted guidance through online platforms can facilitate student reflection on their learning strategies and outcomes. Incorporating both online and offline discussions fosters a more engaging and satisfying interactive learning experience, enhancing emotional engagement.

These insights are valuable for future AI-driven blended course design, underscoring the importance of strategically integrating online and offline components to optimize learning outcomes. The effectiveness of CoI-based blended learning models in promoting deep learning among college students is clearly demonstrated. However, several limitations warrant consideration. First, the sample comprised sophomores from specific educational technology classes at a university in southern China, which limits the generalizability of the findings. Second, the instructional design requirements may differ for courses outside Digital Film and Television Directing and Production, meaning the proposed model's effectiveness could vary across courses and disciplines. Third, the CoI-based blended learning model was only compared to traditional learning; its efficacy against other blended models remains to be explored.

Future research could address these limitations by expanding the sample to include learners from various majors and courses. Comparative analyses could assess the differential impacts of blended learning across different disciplines. Additionally, investigating the relationship between students' sense of presence, learning effectiveness, and deep learning perceptions in blended environments would provide deeper insights into the mechanisms that contribute to successful blended learning experiences. Furthermore, future research should examine the implementation differences between the CoI framework in traditional blended learning and AI-driven distance education. Specifically, exploring how AI technologies can collaborate with the CoI framework to optimize instructional design, enhance student engagement, and improve personalized learning experiences will be a key area of future investigation. In conclusion, while this study highlights the benefits of CoI-based blended learning, future research should aim to address its limitations and explore the broader applicability and underlying mechanisms that enhance blended learning effectiveness.

## Acknowledgments

This research was financially supported by The Philosophy and Social Science Planning Project of Guangdong Province (GD24XJY39), The Characteristic Innovation Project of Colleges and Universities in Guangdong Province (2024WTSCX041), The Education Science Planning Project of Guangdong Province (2018JKZ022, 2024GXJK331), and The Education and Teaching Achievement Award Cultivation Project of Lingnan Normal University (LSGJ2314).

# References

Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing student engagement: Harnessing "AIED"'s power in hybrid education—A review analysis. *Education Sciences, 13*(7), Article 632. https://doi.org/10.3390/educsci13070632

Alqurashi, E. (2019). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education, 40*(1), 133–148. https://doi.org/10.1080/01587919.2018.1553562

Annansingh, F. (2019). Mind the gap: Cognitive active learning in virtual learning environment perception of instructors and students. *Education and Information Technologies, 24*(6), 3669–3688. https://doi.org/10.1007/s10639-019-09949-5

Antonelli, D., Christopoulos, A., Laakso, M.-J., Dagienė, V., Juškevičienė, A., Masiulionytė-Dagienė, V., Mądziel, M., Stadnicka, D., & Stylios, C. (2023). A virtual reality laboratory for blended learning education: Design, implementation and evaluation. *Education Sciences, 13*(5), Article 528. https://doi.org/10.3390/educsci13050528

Armellini, A., Teixeira Antunes, V., & Howe, R. (2021). Student perspectives on learning experiences in a higher education active blended learning context. *TechTrends, 65*(4), 433–443. https://doi.org/10.1007/s11528-021-00593-w

Astiwardhani, W., & Sobandi, A. (2024). Transforming Educational Paradigms: How Micro Learning Shapes Student Understanding, Retention, and Motivation?. Journal of Education Action Research, 8(2), 300-309. https://doi.org/10.23887/jear.v8i2.77711

Ballouk, R., Mansour, V., Dalziel, B., McDonald, J., & Hegazi, I. (2022). Medical students' self-regulation of learning in a blended learning environment: A systematic scoping review. *Medical Education Online, 27*(1), Article 2029336. https://doi.org/10.1080/10872981.2022.2029336

Belda-Medina, J. (2021). Enhancing multimodal interaction and communicative competence through task-based language teaching (TBLT) in synchronous computer-mediated communication (SCMC). *Education Sciences, 11*(11), Article 723. https://doi.org/10.3390/educsci11110723

Bizami, N. A., Tasir, Z., & Kew, S. N. (2023). Innovative pedagogical principles and technological tools capabilities for immersive blended learning: A systematic literature review. *Education and Information Technologies, 28*(2), 1373–1425. https://doi.org/10.1007/s10639-022-11243-w

Castro, R. (2019). Blended learning in higher education: Trends and capabilities. *Education and Information Technologies*, *24*(4), 2523–2546. https://doi.org/10.1007/s10639-019-09886-3

Chen, R. H. (2022). Effects of deliberate practice on blended learning sustainability: A community of inquiry perspective. *Sustainability, 14*(3), Article 1785. https://doi.org/10.3390/su14031785

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*(1),

155–159. https://doi.org/10.1037/0033-2909.112.1.155

Darling-Hammond, L., & Oakes, J. (2021). *Preparing teachers for deeper learning*. Harvard Education Press.

Guo, P., Saab, N., Wu, L., & Admiraal, W. (2021). The Community of Inquiry perspective on students' social presence, cognitive presence, and academic performance in online project-based learning. *Journal of Computer Assisted Learning, 37*(5), 1479–1493. https://doi.org/10.1111/jcal.12586

Han, F. (2023). Relations between students' study approaches perceptions of the learning environment, and academic achievement in flipped classroom learning: Evidence from self-reported and process data. *Journal of Educational Computing Research, 61*(6), 1252–1274. https://doi.org/10.1177/07356331231162823

He, J., Ma, T., & Zhang, Y. (2023). Design of blended learning mode and practice community using intelligent cloud teaching. *Education and Information Technologies*, *28*(8), 10593–10615. https://doi.org/10.1007/s10639-023-11606-x

Hu, Y., & Hwang, G.-J. (2024). Promoting students' higher order thinking in virtual museum contexts: A self-adapted mobile concept mapping-based problem posing approach. *Education and Information Technologies, 29*(3), 2741–2765. https://doi.org/10.1007/s10639-023-11930-2

Hye, J. K., Yi, P., & Hong, J. I. (2020). Students' academic use of mobile technology and higher-order thinking skills: The role of active engagement. *Education Sciences, 10*(3), Article 47. https://doi.org/10.3390/educsci10030047

Islam, M. K., Sarker, M. F. H., & Islam, M.S. (2022). Promoting student-centered blended learning in higher education: A model. *E-Learning and Digital Media, 19*(1), 36–54. https://doi.org/10.1177/20427530211027721

Kerimbayev, N., Umirzakova, Z., Shadiev, R., & Jotsov, V. (2023). A student-centered approach using modern technologies in distance learning: A systematic review of the literature. *Smart Learning Environments, 10*(1), Article 61. https://doi.org/10.1186/s40561-023-00280-8

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212–218. https://www.jstor.org/stable/1477405

Kurniawan, S. (2021). The effect of blended learning of student teams achievement division (STAD) and jigsaw towards higher-order thinking skills. *EDUTEC: Journal of Education and Technology, 5*(1), 121–133. https://doi.org/10.29062/edu.v5i1.229

Li, X.-D., & Cao, H.-H. (2020). Research on VR-supported flipped classroom based on blended learning. A case study in "Learning English through news." *International Journal of Information and Education Technology, 10*(2), 104–109. https://doi.org/10.18178/ijiet.2020.10.2.1347

Liu, W., Wang, J., Zhang, H., Yu, C., Liu, S., Zhang, C., Yu, J., Liu, Q., & Yang, B. (2021). Determining the effects of blended learning using the community of inquiry on nursing students' learning gains in sudden patient deterioration module. *Nursing Open, 8*(6), 3635–3644. https://doi.org/10.1002/nop2.914

Liu, X., & Deris, F. D. (2022). CoI-based teaching practices to promote EFL learners' online discussion in China's blended learning context. *Asian Journal of University Education*, *18*(2), 477–488. https://files.eric.ed.gov/fulltext/EJ1347554.pdf

Lozano-Lozano, M., Fernández-Lao, C., Cantarero-Villanueva, I., Noguerol, I., Álvarez-Salvago, F., Cruz-Fernández, M., Arroyo-Morales, M., & Galiano-Castillo, N. (2020). A blended learning system to improve motivation, mood state, and satisfaction in undergraduate students: Randomized controlled trial. *Journal of Medical Internet Research, 22*(5), Article e17101. https://doi.org/10.2196/17101

Masadeh, T. S. Y. (2021). Blended learning: Issues related to successful implementation. *International Journal of Scientific Research and Management, 9*(10), 1897–1907. https://doi.org/10.18535/ijsrm/v9i10.el02

Meng, Q., Jia, J., & Zhang, Z. (2020). A framework of smart pedagogy based on the facilitating of high order thinking skills. *Interactive Technology and Smart Education*, *17*(3), 251–266. https://doi.org/10.1108/ITSE-11-2019-0076

Meyer, O., Coyle, D., Imhof, M., & Connolly, T. (2018). Beyond CLIL: Fostering student and teacher engagement for personal growth and deeper learning. In Martínez Agudo, J. D. D. (Ed.), *Emotions in second language teaching: Theory, research and teacher education* (pp. 277–297). Springer. https://doi.org/10.1007/978-3-319-75438-3_16

Mthethwa-Kunene, K., Rugube, T., & Maphosa, C. (2022). Rethinking pedagogy: Interrogating ways of promoting deeper learning in higher education. *European Journal of Interactive Multimedia and Education, 3*(1), Article e02204. https://doi.org/10.30935/ejimed/11439

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review, 34,* Article 100394. https://doi.org/10.1016/j.edurev.2021.100394

Pan, Q., Zhou, J., Yang, D., Shi, D., Wang, D., Chen, X., & Liu, J. (2023). Mapping knowledge domain analysis in deep learning research of global education. *Sustainability, 15*(4), Article 3097. https://doi.org/10.3390/su15043097

Pellegrino, J. W., & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press. https://www.fsba.org/wp-content/uploads/2014/04/Education-for-life-and-work.pdf

Perera, C. J., Zainuddin, Z., Piaw, C. Y., Cheah, K. S. L., & Asirvatham, D. (2020). The pedagogical frontiers of urban higher education: Blended learning and co-lecturing. *Education and Urban Society*, *52*(9), 1305–1329. https://doi.org/10.1177/0013124519894966

Priddis, L., & Rogers, S. L. (2018). Development of the reflective practice questionnaire: Preliminary findings. *Reflective Practice, 19*(1), 89–104. https://doi.org/10.1080/14623943.2017.1379384

Prifti, R. (2022). Self-efficacy and student satisfaction in the context of blended learning courses. *Open Learning: The Journal of Open, Distance and e-Learning, 37*(2), 111–125. https://doi.org/10.1080/02680513.2020.1755642

Qi, D., Zhang, M., & Zhang, Y. (2020). Resource integration, value co-creation and continuance intention in MOOCs learning process. *Interactive Learning Environments*, *31*(2), 701–713. https://doi.org/10.1080/10494820.2020.1802299

Rogers, S. L., Priddis, L. E., Michels, M., Tieman, M., & Van Winkle, L. J. (2019). Applications of the reflective practice questionnaire in medical education. *BMC Medical Education, 19*, Article 47. https://doi.org/10.1186/s12909-019-1481-6

Seifan, M., Robertson, N., & Berenjian, A. (2020). Use of virtual learning to increase key laboratory skills and essential non-cognitive characteristics. *Education for Chemical Engineers, 33*, 66–75. https://doi.org/10.1016/j.ece.2020.07.006

Shamir-Inbal, T., & Blau, I. (2021). Facilitating emergency remote K–12 teaching in computing-enhanced virtual learning environments during COVID-19 pandemic—Blessing or curse? *Journal of Educational Computing Research, 59*(7), 1243–1271. https://doi.org/10.1177/0735633121992781

Shen, D.-D., & Chang, C.-S. (2022). Exploring college students' deeper learning perceptions in the blended learning environment: Scale development, validation, and experimental comparison. *International Journal of Technology and Human Interaction, 18*(1), 1–21. http://doi.org/10.4018/IJTHI.313184

Shen, D., & Chang, C.-S. (2023). Implementation of the flipped classroom approach for promoting college students' deeper learning. *Educational Technology Research and Development*, *71*, 1323–1347. http://doi.org/10.1007/s11423-023-10186-4

Tabassum, Z., & Mohd Saad, M. R. B. (2024). A decadal examination of community of inquiry and blended learning in EFL/ESL development: A systematic review. *Arab World English Journals, 15*(1), 401–422. https://dx.doi.org/10.24093/awej/vol15no1.25

Tan, C. (2021). The impact of COVID-19 pandemic on student learning performance from the perspectives of Community of Inquiry. *Corporate Governance: The International Journal of Business in Society*, *21*(6), 1215–1228. https://doi.org/10.1108/CG-09-2020-0419

Vallée, A., Blacher, J., Cariou, A., & Sorbets, E. (2020). Blended learning compared to traditional learning in medical education: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *22*(8), Article e16504. https://doi.org/10.2196/16504

Vlachopoulos, D., & Makri, A. (2019). Online communication and interaction in distance higher education: A framework study of good practice. *International Review of Education, 65*(4), 605–632. https://doi.org/10.1007/s11159-019-09792-3

Vo, H. M., Zhu, C., & Diep, N. A. (2017). The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies in Educational Evaluation, 53*, 17–28. https://doi.org/10.1016/j.stueduc.2017.01.002

Wang, N., Chen, J., Tai, M., & Zhang, J. (2021). Blended learning for Chinese university EFL learners: Learning environment and learner perceptions. *Computer Assisted Language Learning, 34*(3), 297–323. https://doi.org/10.1080/09588221.2019.1607881

Weber, F., Wambsganss, T., & Söllner, M. (2025). Enhancing legal writing skills: The impact of formative feedback in a hybrid intelligence learning environment. *British Journal of Educational Technology*, *56*(2), 650–677. https://doi.org/10.1111/bjet.13529

Wu, T.-T., Lee, H.-Y., Li, P.-H., Huang, C.-N., & Huang, Y.-M. (2023). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *Journal of Educational Computing Research, 61*(8), 1539–1567. https://doi.org/10.1177/07356331231191125

Wut, T. M., Xu, J. (B.), Lee, S. W., & Lee, D. (2022). University student readiness and its effect on intention to participate in the flipped classroom setting of hybrid learning. *Education Sciences*, *12*(7), Article 442. https://doi.org/10.3390/educsci12070442

Yaniawati, P., Maat, S. M., Supianti, I. I., & Fisher, D. (2022). Mathematics mobile blended learning development: Student-oriented high order thinking skill learning. *European Journal of Educational Research, 11*(1), 69–81. https://doi.org/10.12973/eu-jer.11.1.69

Yin, B., & Yuan, C.-H. (2021). Precision teaching and learning performance in a blended learning environment. *Frontiers in Psychology, 12*, Article 631125. https://doi.org/10.3389/fpsyg.2021.631125

Yin, B., & Yuan, C.-H. (2022). Blended learning performance influence mechanism based on community of inquiry. *Asia Pacific Journal of Education, 44*(4), 977–992. https://doi.org/10.1080/02188791.2022.2061912

Zhang, J.-H., Zou, L.-c., Miao, J.-j., Zhang, Y.-X., Hwang, G.-J., & Zhu, Y. (2020). An individualized intervention approach to improving university students' learning performance and interactive behaviors in a blended learning environment. *Interactive Learning Environments*, *28*(2), 231–245. https://doi.org/10.1080/10494820.2019.1636078

Zhang, R. (2020). Exploring blended learning experiences through the community of inquiry framework. *Language Learning & Technology*, *24*(1), 38–53. https://doi.org/10125/44707

Zhao, J., Hwang, G.-J., Chang, S.-C., Yang, Q.-f., & Nokkaew, A. (2021). Effects of gamified interactive e-books on students' flipped learning performance, motivation, and meta-cognition tendency in a mathematics course. *Educational Technology Research and Development, 69*, 3255–3280. https://doi.org/10.1007/s11423-021-10053-0

Zhao, S., & Song, J. (2022). Unpacking the emotional experiences of learners in a blended learning context. *Frontiers in Psychology, 13*, Article 879696.

https://doi.org/10.3389/fpsyg.2022.879696

Zhao, W. (2022). An empirical study on blended learning in higher education in "Internet+" era. *Education and Information Technologies, 27*(6), 8705–8722. https://doi.org/10.1007/s10639-022-10944-6

Zhu, M., Berri, S., & Zhang, K. (2021). Effective instructional strategies and technology use in blended learning: A case study. *Education and Information Technologies, 26*(5), 6143–6161. https://doi.org/10.1007/s10639-021-10544-w

Zhu, M., & Bonk, C. J. (2019). Designing MOOCs to facilitate participant self-monitoring for self-directed learning. *Online Learning Journal, 23*(4), 106–134. https://doi.org/10.24059/olj.v23i4.2037

Zimmerman, W. A., & Kulikowich, J. M. (2016). Online learning self-efficacy in students with and without online learning experience. *The American Journal of Distance Education, 30*(3), 180–191. https://doi.org/10.1080/08923647.2016.1193801


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Enhancing Team Performance in Hybrid-Flexible Courses: The Role of Team Communication and Atmosphere

**Miaohui Wang[1]*, Wei Zhan[1], Yan Li[2], and Wu Song[2]**
*[1]College of Mechanical Engineering and Automation, Huaqiao University, Xiamen, China; [2]School of Fine Arts, Nanjing Normal University, Nanjing, China; *Corresponding Author*

## Abstract

Hybrid-flexible (hyflex) courses have become a prominent open and distributed learning (ODL) approach post-COVID-19, offering students the flexibility to attend classes online, in-person, or both. While this model promotes instructional continuity and student-centered learning, it presents unique ODL challenges in managing effective team collaboration. This quantitative study investigated the underexplored relationships among effective communication, team atmosphere, and team performance in hyflex learning environments; the study sought to provide novel insights into optimizing team dynamics and enhancing learning outcomes in ODL-integrated educational settings. The study comprised data from 310 college students enrolled in the authors' hyflex course offered on China's Super Star learning platform at an ODL institution with over 700 million registered users. An online survey was conducted using validated scales; confirmatory factor analysis (CFA) and multiple linear regression analyses were used to test the hypotheses. The study found effective communication significantly enhanced team performance ($\beta = 0.389$, $p < .001$). Team atmosphere moderated this relationship ($\beta = -0.118$, $p < .05$); while a positive atmosphere generally supported performance, an overly positive one may have reduced the critical engagement necessary for optimal outcomes. The study concluded that both effective communication and a balanced team atmosphere were critical for maximizing team performance in hyflex courses. We suggest ODL educators foster communication strategies that encourage open dialogue and critical thinking while maintaining a supportive team environment. Overemphasis on maintaining positivity may hinder constructive critique and diminish performance in distance and hybrid settings.

*Keywords*: input-process-output theory, team atmosphere, team communication, team performance, hybrid-flexible course

# Introduction

According to Thurnau and Byun (2022), hybrid-flexible (hyflex) courses have become one of the 10 most prominent pedagogical innovations since the COVID-19 pandemic, yet their implications for collaborative learning dynamics remain underexplored. While hyflex was first conceptualized and implemented by Beatty (2005) as an extension of established ODL principles, it gained unprecedented visibility in post-pandemic education. As an established but increasingly adopted approach in education, hyflex courses have provided unique flexibility, allowing students to attend classes online, in person, or through a combination of both. This model has played a critical role in maintaining instructional continuity while promoting student-centered learning, particularly during the global transition to remote education (Cumming et al., 2024; Detyna et al., 2023). Such flexibility, a crucial and fundamental characteristic of ODL, has been researched broadly across its various approaches for decades (Bozkurt, 2019). It has led to significant shifts in team-based learning dynamics and pedagogical approaches, requiring educators to adapt to new ways of facilitating collaboration and interaction in mixed-modality environments (Kohnke & Moorhouse, 2021; Wong et al., 2023).

However, this flexibility has also introduced new challenges in managing effective team collaboration and ensuring equitable participation among students, particularly when team members engage in different modes of learning simultaneously. Prior research has indicated that achieving equity in hybrid learning environments was often hindered by issues such as proximity bias, where in-person participants tended to dominate interactions with varying levels of engagement among online and in-person students as a result (Lohmann et al., 2021; Qureshi et al., 2023; Zhang & Li, 2024). Additionally, the use of diverse digital communication tools can further complicate team dynamics, potentially leading to misunderstandings and reduced cohesion if not managed effectively (Wilson & Alexander, 2021).

While numerous studies have investigated team communication (Johnson et al., 2014; Razali et al., 2015), and team atmosphere in traditional learning environments (Kanaris & Mujtaba, 2023; Lyons, 2024; Sargeant et al., 2008), there is a paucity of studies that integrated these dimensions within hyflex learning environments. Moreover, existing research on distributed learning frameworks has often focused on fully online or asynchronous settings, leaving a gap in understanding how hybrid modalities, such as hyflex courses, uniquely shape team dynamics (Brown, 2024). This study addressed this gap by examining the interplay between communication and team atmosphere in hyflex courses, as exemplified in Appreciation of Traditional Fujian Arts and Crafts, a course offered at Huaqiao University. Our intent was not only to advance our theoretical understanding of team dynamics in hyflex learning settings but also to provide a fresh perspective on team collaboration in ODL environments. Furthermore, this study was designed to provide practical guidelines for educators in the design and implementation of hyflex courses.

# Team Communication in Hyflex Course Learning

## Defining Hyflex Course Learning and Course Learning Team Communication

Hyflex courses represent a pedagogical model that offers students multiple participation options, accommodating diverse learning preferences and circumstances. Students can choose to attend classes in person, participate synchronously online, or engage asynchronously through recorded materials. While not new to the field of ODL, the hyflex model originated in 2005 when Brian J. Beatty developed

it to address the needs of working graduate students by providing multiple participation pathways (Beatty, 2019). Drawing from established ODL principles of flexibility, student choice, and equivalence of learning experiences, hyflex represented an evolution of hybrid learning approaches that have long been fundamental to distance education (Miller et al., 2013). This model, increasingly significant in higher education after the COVID-19 pandemic, enhanced accessibility and flexibility by allowing students to tailor their learning experience and switch between modes as needed (Beatty, 2019; Detyna et al., 2023; Kohnke & Moorhouse, 2021). Unlike other ODL formats, which have often focused on asynchronous delivery or fully online formats, hyflex combines multiple modalities to meet varied student needs (Vetrivel & Mohanasundaram, 2024). Research has shown that this flexibility improved student engagement by helping them balance academic and personal responsibilities while enabling institutions to address diverse learning requirements (Graham et al., 2013; Wong et al., 2023).

Course learning team communication refers to the structured exchange of knowledge, ideas, and feedback among students working on academic tasks (Johnson & Johnson, 1991). It supports collaborative learning by enabling students to discuss concepts, address misunderstandings, and solve problems together (Penrod, 2023). This communication has often involved formal mechanisms like group projects, peer reviews, and structured discussions aligned with course objectives (Johnson et al., 2014).

In hyflex courses, team communication must bridge the gaps among in-person, synchronous online, and asynchronous participants, ensuring effective collaboration despite differences in time, space, and technological access (Wong et al., 2023). Effective communication is essential for fostering engagement and connection across participation modes, supporting the diverse ways in which students learn, and creating an inclusive learning environment.

## The Challenge of Team Communication in Hybrid-Flexible Course Learning

The effectiveness of hyflex courses has been widely studied, highlighting benefits such as increased student engagement and satisfaction (Bockorny et al., 2024). Research has indicated that hyflex courses improve learning outcomes by fostering a more inclusive environment that accommodates diverse learning ways (Graham et al., 2013). Additionally, their flexibility reduces barriers for non-traditional students, such as those balancing work and study, thereby supporting a more diverse student population (Detyna et al., 2023; Kohnke & Moorhouse, 2021).

However, hyflex courses also present unique challenges due to their integration of multiple participation modes. This hybrid nature has created a complex communication landscape that is distinct from other ODL formats, which often focus on either synchronous or asynchronous online learning (Wilson & Alexander, 2021). Research has indicated that remote participants in hyflex courses have often experienced proximity bias (Raes et al., 2020), feeling marginalized compared to in-person peers. For example, online participants may struggle with reduced visibility in discussions, leading to less active participation and a sense of isolation (Qureshi et al., 2023; Wong et al., 2023). Additionally, reliance on digital tools like Zoom or Microsoft Teams can introduce technological barriers, exacerbate varying levels of digital literacy, and present communication challenges, further hindering team cohesion (Detyna et al., 2023; Zhang & Li, 2024).

Despite extensive research on team communication in traditional and fully online settings, there is limited exploration of the unique challenges posed by hyflex environments. Studies on ODL have

frequently focused on fully online or asynchronous learning, overlooking the hybrid nature of hyflex courses, where in-person and online participation coexist (Penrod, 2023). While effective communication strategies are critical for fostering collaboration and trust (Sargeant et al., 2008), little research has examined how these strategies can be adapted to hyflex contexts. This is particularly important as hyflex courses require communication approaches that ensure equitable participation and engagement across all modes (Wong et al., 2023).

Addressing this research gap is essential for understanding how communication dynamics influence student engagement and learning outcomes, specifically in team environments within hyflex courses. Developing tailored communication strategies that address the hybrid nature of hyflex learning can help educators support interactions across different modes, fostering a more cohesive and equitable team experience (Brown, 2024; Lohmann et al., 2021; Wong et al., 2023). By tackling these challenges, institutions can enhance educational experiences and ensure equitable participation for all students.

## Team Atmosphere in Hyflex Course Learning

Team atmosphere in hyflex course learning refers to the collective emotional and social environment among students. It is characterized by the quality of interactions, mutual support, and the sense of belonging, which can significantly influence engagement and collaboration (Beatty, 2019). A positive team atmosphere fosters trust and open communication, essential for effective teamwork and successful learning outcomes (Bower et al., 2015). In contrast to traditional face-to-face or fully online learning environments, the hyflex model introduces unique complexities in cultivating a positive team atmosphere due to its simultaneous integration of multiple participation modes (Vetrivel & Mohanasundaram, 2024).

Research has shown that a supportive team atmosphere enhances engagement and collaboration across participation modes, helping students feel connected regardless of how they attend (Garrison et al., 1999; Graham et al., 2013). According to Detyna et al. (2023) and Lyons (2024) a strong sense of community encouraged active participation and bridged the gap between online and in-person learners. However, challenges such as feelings of isolation among remote participants, proximity bias favoring in-person learners, and varying levels of technological proficiency have been shown to hinder team cohesion and effective communication (Hadjipieris, 2024; Qureshi et al., 2023; Zhang & Li, 2024). The ability to use digital tools effectively has emerged as particularly critical in hyflex settings (Wilson & Alexander, 2021).

Despite these challenges, it is crucial to cultivate a positive team atmosphere in hyflex courses as it not only influences individual student performance but also affects overall group dynamics. It has been shown to mitigate feelings of isolation, enhance motivation, and promote peer support, leading to better learning outcomes (Kohnke & Moorhouse, 2021; Lohmann et al., 2021; Penrod, 2023). Research on hyflex environments, characterized by the coexistence of synchronous and asynchronous participation modes, has been limited compared to studies of fully online settings (Wong et al., 2023). Further studies are needed to explore how specific communication practices, technology use, and instructor interventions can influence student interactions and foster a sense of community among diverse participants in hyflex settings (Sargeant et al., 2008). Addressing these challenges will help educators to optimize hyflex course design and improve student experiences.

### Team Performance in Hyflex Courses

Team performance in hyflex course learning refers to the effectiveness with which student teams collaborate to achieve educational objectives within a flexible learning framework (Beatty, 2019). This performance encompasses various dimensions, including communication, collaboration, and the ability to adapt to different modes of participation whether in-person, synchronous online, or asynchronous (Beatty, 2019). Effective team performance is essential for maximizing learning outcomes and ensuring that all members contribute meaningfully to group tasks. Recent studies have emphasized that in hyflex learning, adaptability and responsiveness to different participation modes are critical for maintaining team cohesion and achieving learning objectives (Detyna et al., 2023; Kohnke & Moorhouse, 2021).

Research has demonstrated that structured collaboration and effective use of technology are critical for success (Akcaoglu & Lee, 2016; Al-Rahmi et al., 2015; Jarwati, 2024) Using collaborative tools such as Google Docs and Zoom have facilitated communication and ensured equitable participation among team members (Graham et al., 2013; Kohnke & Moorhouse, 2021). However, compared to traditional ODL courses, where asynchronous communication often dominates, hyflex courses require teams to balance synchronous and asynchronous interactions, which can create disparities in engagement and participation (Heilporn & Lakhal, 2021). Challenges such as proximity bias (Raes et al., 2020), where in-person participants dominate discussions at the expense of remote members, have hindered effective collaboration (Zhang & Li, 2024). Moreover, disparities in technological proficiency may also exacerbate engagement issues, impacting the overall effectiveness of team interactions (Brown, 2024; Hadjipieris, 2024).

Despite these insights, there remains a notable gap in comprehensive studies examining how specific factors such as team communication and atmosphere influence performance outcomes in undergraduate hyflex educational context. Hyflex courses have introduced a hybrid dynamic that requires further investigation beyond traditional ODL research (Keshavarz, 2023). There is limited understanding of how these elements interact with team atmosphere to shape outcomes in hyflex courses (Penrod, 2023). This gap is particularly significant as hyflex courses demand integrative approaches that address the interplay among synchronous and asynchronous communication, team atmosphere, and performance outcomes (Heilporn & Lakhal, 2021). Understanding this interplay will provide insights for educators aiming to enhance collaborative learning experiences in ODL environments.

## Theoretical Framework: Input-Process-Output Theory

The proposed framework for this study is grounded in the input-process-output (IPO) theory, a well-established model in organizational behavior research (Hackman & Morris, 1975). The IPO model delineates how various inputs—such as team composition, communication styles, and environmental context—impact team processes, including collaboration and conflict resolution, ultimately influencing outputs like performance outcomes (Forsyth, 2014; Hackman & Morris, 1975). In the context of higher education, the IPO model has increasingly been employed to understand team dynamics within flexible learning environments, such as hyflex courses.

Compared to traditional face-to-face educational settings, where inputs such as physical presence, synchronous communication, and standardized learning materials are the norm, the hyflex context has introduced more complex and dynamic inputs. These include asynchronous and synchronous

participation modes, a variety of digital communication platforms, and the need for technological proficiency among team members (Beatty, 2019; Kohnke & Moorhouse, 2021). Due to the physical and temporal separation of team members, these unique inputs not only shape but also challenge team processes such as effective communication, task coordination, and conflict resolution (Bond et al., 2021). Thus, the IPO model in hyflex settings must account for the fluid and multimodal nature of interactions, which is less emphasized in traditional IPO applications.

Moreover, processes in hyflex learning environments are mediated by technological interfaces that facilitate or hinder communication and collaboration. For instance, while traditional IPO applications have often assumed consistent face-to-face interactions, hyflex processes must consider asynchronous discussions, time lags, and technological barriers as integral components of the team process (Raes et al., 2020). This complexity adds new dimensions to the IPO framework, where the quality of digital tools, Internet accessibility, and digital literacy become critical process variables.

Outputs in the hyflex context are also distinct, as performance outcomes are not limited to academic achievements but extend to students' digital collaboration skills, adaptability, and technological proficiency, reflecting a broader spectrum of success metrics compared to traditional learning environments (Kohnke & Moorhouse, 2022). Applying the IPO model in hyflex settings thus requires a nuanced understanding of how digital communication tools and flexible participation modes shape collaborative processes and learning outcomes.

To investigate the effects of communication and atmosphere within hyflex course learning teams on team performance in undergraduate hyflex education setting, we formed the following core research questions:

1. Does effective hyflex course learning team communication significantly enhance team performance?

2. What is the impact of a positive team atmosphere on team performance?

3. How does team atmosphere moderate the relationship between communication and team performance in a hyflex learning environment?

To address these research questions, we proposed three hypotheses.

Hypothesis 1: Effective course learning team communication significantly enhances team performance in an undergraduate hyflex setting.

Effective communication is particularly crucial in hyflex learning environments. Recent research in higher education has emphasized that clear and structured communication channels are essential for maintaining team cohesion and ensuring active participation across different learning modalities (Brown, 2024; Detyna et al., 2023). Studies have indicated that open communication channels enhance collaboration and information sharing, which are essential for achieving high team performance (Mathieu et al., 2006). Furthermore, effective communication has been shown to mitigate misunderstandings and foster a sense of belonging among team members, ultimately leading to improved outcomes (Höddinghaus et al., 2023).

Hypothesis 2: A positive team atmosphere positively influences team performance in an undergraduate

hyflex setting.

A supportive atmosphere encourages active participation, idea sharing, and effective collaboration among members (Baker et al., 2006). Research in educational psychology has shown that a positive emotional climate can significantly increase students' willingness to engage in collaborative activities, thereby enhancing overall team effectiveness in hybrid learning settings (Keshavarz, 2023; Salas et al., 2005).

Hypothesis 3: Team atmosphere moderates the relationship between communication and team performance in an undergraduate hyflex setting.

This hypothesis highlighted the interplay between communication and the emotional climate of the team. A positive atmosphere may amplify the benefits of effective communication by creating an environment where members feel valued and understood (González - Romá et al., 2009). Conversely, in the absence of a supportive atmosphere, students may be less likely to engage actively in discussions, limiting the effectiveness of communication strategies (Brown, 2024; Wong et al., 2023). Both factors must be considered together to fully understand their impact on team outcomes. It is crucial for educators designing hyflex courses to foster environments where communication and atmosphere complement each other to support optimal team performance (Detyna et al., 2023).

# Methodology

## Context

This study adopted a quantitative survey design using a seven-point Likert scale to investigate relationships among team communication, atmosphere, and performance in a hyflex setting. The research was conducted at a Chinese undergraduate institution. Students participated in hyflex learning teams in the *Appreciation of Traditional Fujian Arts and Crafts* course on the Super Star learning platform, a major ODL platform in China with over 700 million registered users. This hyflex learning environment enabled students to choose among in-person attendance, synchronous online participation, or asynchronous engagement with course materials, all of which were well-established open and distributed learning practices

## Participants and Sampling

The target population consisted of undergraduate students enrolled in *Appreciation of Traditional Fujian Arts and Crafts* over two semesters from September 1, 2023 to July 20, 2024. A total of 332 responses were collected using in-class surveys administered via Wenjuanxing, an online crowdsourcing platform used in China, and similar to Amazon Mechanical Turk. Screening criteria excluded responses with insufficient answering time (less than 45 seconds, $n = 17$) or patterned answers ($n = 5$). This process resulted in 310 valid responses included in the analysis.

## Survey Instrument Design

The survey instrument was designed to measure three key constructs: team atmosphere, team communication, and team performance. All survey items were presented in Chinese, and responses were collected using a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

The final instrument was developed using recommended translation and back-translation procedures performed by five experienced hyflex educators. A pilot study involving 75 students was conducted to test the validity and clarity of the questionnaire and led to refinement before its use in the main study.

Team atmosphere was measured using nine items adapted from Fan et al.'s (2017) team atmosphere scale, with minor modifications to fit the hyflex learning context. The scale evaluated three dimensions: personal relationships, psychological safety, and team trust, with reliability coefficients (α) of 0.77, 0.76, and 0.79, respectively. The complete items have been provided below.

1. Colleagues are reliable in important matters.
2. We invest emotional effort in working relationships.
3. I take the initiative to share my thoughts, problems, or feelings with my course team leader.
4. The course team leader seeks input from team members when making decisions.
5. Team members can maintain their own opinions even after decisions are made.
6. Team members can express disagreement with leaders.
7. There is mutual trust among team members.
8. We believe everyone acts with good intentions.
9. We trust that the information shared is truthful.

Team communication was measured using the following four items adapted from He et al. (2014) and Wang (2008), with a reliability coefficient (α) of 0.83.

1. Team members express their opinions clearly and accurately.
2. Communication among team members is characterized by mutual respect and a positive atmosphere.
3. Team members frequently share views and opinions.
4. There are multiple ways for team communication and information exchange.

Finally, team performance was assessed across two dimensions: goal achievement and member satisfaction. Each dimension consisted of three items, adapted from Huang (2011) and Zheng (2006), with Cronbach's alpha reliability coefficients of 0.79 and 0.78, respectively. The complete list has been provided below.

1. I have a clear understanding of the team's objectives.
2. Team members complete their tasks accurately.
3. The team consistently meets its goals.

4. I am satisfied with how the team works.
5. I am satisfied with my own performance.
6. I am satisfied with the performance of other team members.

### Validity and Reliability

The reliability of the scales was confirmed using Cronbach's alpha values for each construct. The final Chinese version of the questionnaire was developed using the recommended translations and back-translation procedures (Brislin, 1970). Five experienced teachers from the research site who had extensive experience in hyflex courses conducted the back-translation of the questions. The accuracy of the translated items was tested through a pilot study involving 75 students who had previously participated in a hyflex team-learning course. Based on the pilot results, the questionnaire was refined and finalized. Random post-pilot interviews were also conducted with 30 students to further validate the instrument's content.

### Data Collection

As mentioned above, data was collected over two semesters from 332 learners enrolled in the authors' hyflex course delivered through Super Star learning platform. Unlike typical Wenjuanxing surveys, this study was administered during scheduled learning sessions, ensuring that only students who had experience with the hyflex course could participate, thereby minimizing the risk of duplicated or ineligible responses.

To mitigate the possible risk of common method variance, the sequence of questionnaire items was carefully structured. Respondents' anonymity was fully protected.

### Data Analysis

After data cleaning and screening, statistical analyses were performed on the final dataset of 310 valid responses. Analytical techniques included descriptive statistics and regression analyses to examine relationships among team atmosphere, communication, and performance. All analyses were conducted using IBM SPSS Amos 28, a structural equation modeling software.

## Results

We began by conducting a confirmatory factor analysis (CFA) to validate the four measurement models, followed by multiple linear regression to test the hypotheses. The CFA assessed the reliability and validity of the constructs, while the regression analysis tested the relationships between variables. Centered values from the CFA were used in the regression analysis. To address potential multicollinearity, which could bias the regression estimates, we used variance inflation factor (VIF) analysis. Additionally, the condition number was calculated to evaluate the model's stability, confirming reduced multicollinearity and ensuring the robustness of the model.

**Table 1**

*Model Fit Indices*

| Fit index | Model value | Reference value | Overall model fit |
|---|---|---|---|
| $\chi^2/df$ | 1.695 | < 3 | Yes |
| CFI | 0.969 | > 0.9 | Yes |
| IFI | 0.969 | > 0.9 | Yes |
| GFI | 0.916 | > 0.9 | Yes |
| SRMR | 0.030 | < 0.05 | Yes |
| RMSEA | 0.047 | < 0.08 | Yes |

Several indices were used to assess the model fit, including degrees of freedom (*df*), chi-square value ($\chi^2$), $\chi^2$ /*df* ratio, comparative fix index (CFI), goodness-of-fit index (GFI), Bollen's incremental fit index (IFI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA). Table 1 shows that all indices met the recommended reference values, indicating an overall good fit for the measurement models.

**Table 2**

*Summary Statistics of CFA (*N = *310)*

| Construct | Index number | Convergent validity | | | Model-fit indices | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $SMC_{min}$ | CR | AVE | $\chi2$ | *df* | $\chi2/df$ | GFI | AGFI | RMSEA |
| TC (Team communication) | 4 | 0.51 | 0.84 | 0.57 | 6.58 | 2 | 3.29 | 0.99 | 0.95 | 0.086 |
| TA (Team atmosphere) | 9 | 0.63 | 0.90 | 0.51 | 64.68 | 27 | 2.40 | 0.96 | 0.93 | 0.067 |
| TP (Team performance) | 6 | 0.49 | 0.89 | 0.56 | 25.08 | 9 | 2.79 | 0.97 | 0.94 | 0.076 |

The constructs—team communication (TC), team atmosphere (TA) and team performance (TP) —were evaluated for internal consistency reliability, convergent validity, and discriminant validity (see Table 2). The results showed that the composite reliability (CR) for each construct ranged from 0.81 to 0.90, exceeding the acceptable threshold of 0.60, thus confirming internal consistency reliability (Bagozzi & Yi, 1989; Fornell & Larcker, 1981). Additionally, the factor loadings for all items in the four-factor model were all significant ($p < .001$), providing evidence of convergent validity. The average variance extracted (AVE) for all constructs ranged from 0.51 to 0.58, exceeding the 0.50 threshold, further supporting acceptable convergent validity.

Overall, the CFA results confirmed the reliability and validity of the measurement models. These validated constructs were subsequently used in the regression analysis to test the study's hypotheses.

**Table 3**

*Mean, Standard Deviation, Reliability, and Correlations of Constructs*

| Construct | Mean | SD | TC | TA | TP |
|---|---|---|---|---|---|
| TC | 5.19 | 1.03 | 0.755 | | |
| TA | 5.22 | 0.98 | 0.519** | 0.714 | |
| TP | 5.35 | 1.05 | 0.553** | 0.542** | 0.748 |

*Note:*** $p < .01$ (two-tailed); $N = 310$.

The square roots of AVE for discriminant validity are in parentheses along the diagonal.

Table 3 shows the mean, standard deviation, reliability, and correlations of the variables. TC was positively corelated with both TA ($r = 0.519$, $p < .01$) and TP ($r = 0.553$, $p < .01$). Similarly, TP was positively correlated to TA ($r = 0.542$, $p < .01$). Additionally, the square roots of the AVE of each variable, shown along the diagonal in Table 3, were higher than the correlations among the constructs, indicating strong discriminant validity.

To address potential multicollinearity and improve the interpretation of interaction effects, all continuous independent variables were centered by subtracting each variable's mean from its raw value. This centering reduced correlations between main effects and interaction terms, ensuring clear interpretation of the regression results.

**Table 4**

*Regression Coefficients*

| Items | Unstandardized coefficient | | Standardized coefficient | *t* | *p* | Collinearity statistics | |
|---|---|---|---|---|---|---|---|
| | B | *SE* | Beta | | | Tolerance | VIF |
| (Constant) | 5.470 | 0.032 | | 172.126 | 0.000 | | |
| TC (centered) | 0.389 | 0.046 | 0.382 | 8.457 | 0.000 | 0.326 | 3.071 |
| TA (centered) | 0.267 | 0.037 | 0.296 | 7.196 | 0.000 | 0.394 | 2.538 |
| TC*TA | -0.139 | 0.018 | -0.306 | -7.592 | 0.000 | 0.411 | 2.434 |

*Note*: Dependent factor was team performance.

Table 4 presents the regression coefficients for the model. The maximum variance inflation factor (VIF) values ranged from 3.071 to 2.434, well below the recommended threshold of 10, indicating minimal multicollinearity (O'Brien, 2007). The condition number after the regression was calculated as 3.418, confirming the absence of multicollinearity and ensuring the stability of the regression model. The model produced an $R^2$ of 0.796, with a significant F change ($p < 0.001$), and a Durbin-Walson statistic of 1.877, indicating a good fit and no issues with autocorrelation.

The following are the results of hypothesis testing. H1 was supported; TC had a significant positive impact on TP, with an unstandardized coefficient of 0.389 ($t$ = 8.457, $p$ <0.001). This result aligned with previous research, underscoring the crucial role of effective communication in enhancing team performance.

H2 was also supported; TA had a significantly positive effect on TP, with an unstandardized coefficient of 0.267 ($t$ = 7.196, $p$ <0.001). This finding was consistent with existing literature, which emphasized the importance of supportive team atmosphere in driving performance outcomes (Yamagata-Lynch, 2014).

H3 was confirmed; TA negatively moderated the relationship between TC and TP. The unstandardized coefficient was 0.389 ($t$ = -0.139, $p$ <0.001), implying that as team atmosphere improved, the positive effect of team communication on performance was diminished. This nuanced result highlights a complex interaction between communication and atmosphere, in line with theories suggesting diminishing returns in highly cohesive teams (Mathieu et al., 2008).

## Discussion

Our findings suggested that effective communication is a critical driver of team performance in hyflex courses. Equally, a positive team atmosphere significantly enhances team performance, it also moderates the effect of communication. Specifically, our results revealed that while robust communication enhanced performance, its incremental benefits may diminish in a highly positive team atmosphere.

Effective communication is essential for collaboration in ODL environments, especially in hyflex courses where students predominantly engage synchronously or asynchronously online, with minimal or no face-to-face interaction. Clear communication strategies are vital to overcome the spatial and temporal barriers inherent in ODL, ensuring equitable engagement despite physical separation (Detyna et al., 2023; Kohnke & Moorhouse, 2021). Providing multiple interaction channels (e.g., discussion boards, video conferencing, collaborative documents) further enhance meaningfully engagement with peers and course content (Wong et al., 2023).

The positive impact of communication on team performance aligned with findings from studies indicating that collaborative learning activities can foster deeper understanding and retention of material (Johnson et al., 2014). In a hyflex ODL environment, where students are rarely or never physically co-located, leveraging technology to facilitate communication is crucial. Tools such as Slack, Microsoft Teams, and other online collaboration platforms have been shown to reduce communication barriers in hyflex settings, helping to create a more inclusive learning environment (Brown, 2024; Detyna et al., 2023; Penrod, 2023).

Our findings suggested that although effective communication was vital, its benefits may be less pronounced in an already positive environment. Literature has supported this notion by highlighting that in some cases, an overly harmonious atmosphere has led to complacency among team members, potentially stifling critical discussions and constructive feedback (Garrison & Vaughan, 2008; Yamagata-Lynch, 2014). In such cases, teams may have experienced diminishing returns on communication efforts as they prioritized maintaining harmony over engaging in challenging conversations necessary for growth and improvement. This phenomenon can be explained through the

lens of group dynamics theory, which posited that highly cohesive teams often develop strong social bonds that may inadvertently discourage dissent or critical engagement (Forsyth, 2014). Such dynamics are particularly relevant in hyflex courses, where the diverse modalities of participation may further reduce the likelihood of dissenting opinions being expressed, especially in asynchronous settings where immediacy and direct feedback are limited.

Cultural factors may also partially explain the diminishing returns of communication benefits in highly positive team atmospheres observed in this study, as participants were Chinese university students from a cultural background that emphasizes group harmony, interdependence, and respect for authority (Hofstede, 1984; Triandis, 2018). This collectivist orientation may have amplified tendencies to avoid conflict in highly cohesive teams. In ODL contexts, these cultural dynamics are further complicated by the reduced social presence and limited non-verbal cues inherent in online communication (Garrison et al., 2010). Research has shown that the absence of face-to-face interaction in virtual teams can exacerbate cultural tendencies toward conflict avoidance (Ke & Kwak, 2013). Maintaining harmony in online collectivist contexts often supersedes critical engagement, even when constructive disagreement could enhance team performance.

Moreover, research has indicated that while collaboration is beneficial, there comes a point where too much emphasis on consensus can hinder decision-making processes (Murray, 1983; Priem et al., 1995). In hyflex courses, fostering an environment that encourages critical engagement and constructive dissent is especially important; it can balance the need for harmony with the need for rigorous academic discourse (Brown, 2024; Kohnke & Moorhouse, 2021; Wong et al., 2023). This balance is particularly critical in hyflex settings, where the heterogeneity of participation modes necessitates deliberate efforts to ensure that all voices are heard. For example, asynchronous participants may require structured opportunities to contribute their perspectives, while synchronous or in-person participants may need to be encouraged to actively seek input from their peers in other modes (Lyons, 2024).

## Contributions and Implications

Our study has contributed to the existing body of literature on ODL and hyflex learning by providing empirical evidence that highlights the importance of effective communication and team atmosphere in enhancing team performance. Previous studies have emphasized the role of communication in collaborative learning settings (Beatty, 2019; Hadjipieris, 2024; Penrod, 2023), but often overlooked the unique challenges of building team cohesion without physical co-presence. Our research addressed this gap by demonstrating the moderating role of team atmosphere in ODL contexts where spontaneous interactions are limited. This adds complexity to existing theories surrounding team dynamics, suggesting that while effective communication is essential, its impact on performance varies based on the level of supportiveness within the distributed team environment (Detyna et al., 2023; Kohnke & Moorhouse, 2021).

The observation regarding diminishing returns on communication benefits within a highly positive team atmosphere introduced a critical nuance to existing theories. While positive atmospheres are generally associated with enhanced performance, our findings indicated that an overly harmonious environment may lead to complacency or reduced critical engagement among team members. This aligns with recent discussions in educational psychology, which have emphasized the need for a balance between support and critical dialogue to maintain team effectiveness (Brown, 2024; Salas et al., 2005). In the context of ODL, this balance is particularly important. Moore's transactional distance theory

established that spatial separation increases psychological distance between educational participants, necessitating structured dialogue mechanisms (Moore, 2013). Garrison's community of inquiry framework further emphasized the essential equilibrium between social and cognitive presence in effective online learning environments (Garrison, 2016). Empirical investigations have confirmed that successful ODL implementation requires balanced interaction structures that support both interpersonal cohesion and critical discourse (Miao & Ma, 2022). By highlighting this potential pitfall, our research encourages educators to maintain this balance in environments where face-to-face regulatory cues are absent.

The practical implications of the moderation effect are particularly noteworthy. For instance, in highly cohesive teams where trust and mutual understanding are already well-established, educators may consider reducing the frequency of mandatory team meetings or communication check-ins to avoid redundancy and cognitive overload (Marks et al., 2001). Instead, they could focus on designing activities that challenge students to critically evaluate ideas or solve complex problems, ensuring that team discussions remain meaningful and productive. Educators should prioritize both effective communication strategies and a supportive team atmosphere across ODL contexts. In synchronous settings, for instance, supportive atmospheres could be developed through structured turn-taking in video conferences and collaborative digital whiteboarding (Martin et al., 2022). Asynchronously, support could emerge from consistent instructor forum presence and peer review protocols with constructive feedback guidelines (Borup et al., 2020), amongst other strategies. These structured communication approaches would ensure all students remained engaged, regardless of participation mode (Detyna et al., 2023; Keshavarz, 2023).

Educators must also be mindful of cultivating a balanced team atmosphere that encourages openness while avoiding excessive conformity. To achieve this balance, strategies such as role rotation, peer feedback sessions, and structured debates would encourage critical discussions while maintaining a positive atmosphere (Penrod, 2023). By actively promoting dissenting opinions in a respectful manner, educators can harness the benefits of both effective communication and a positive atmosphere.

Educators should implement online feedback mechanisms that allow students to reflect on their experiences regarding communication and team atmosphere. This could involve methods such as anonymous surveys, reflective journals, or real-time feedback tools that help instructors adjust strategies and maintain an optimal balance between support and challenge within teams (Brown, 2024; Detyna et al., 2023; Wong et al., 2023).

In conclusion, our findings have advanced hyflex pedagogy in two key dimensions. First, the discovery of a negative moderation effect of team atmosphere challenges conventional assumptions that more positivity is always better, a novel insight with direct implications for avoiding complacency in highly cohesive teams (Yamagata-Lynch, 2014). Second, we formed actionable strategies that addressed the core hyflex challenge of balancing flexibility with equitable participations, which is critical for institutions scaling hybrid online-physical models. By extending these insights to fully online and other ODL contexts, this study has provided a foundation for future research and practice aimed at improving team performance in increasingly diverse and flexible learning environments.

The findings have also offered important implications for institutional policies. Institutions should consider implementing policies that promote structured online communication and balanced team dynamics to maximize collaborative learning outcomes. For example, institutions could provide

training for instructors on how to foster a supportive yet critically engaging team atmosphere in both synchronous video sessions and asynchronous discussion boards. This training could include strategies for encouraging constructive dissent, such as using anonymous online feedback tools or assigning rotating leadership roles within teams (Forsyth, 2014).

While this study has contributed valuable insights into hyflex course learning, several limitations must be acknowledged. The findings were based on a sample from within a collectivist cultural context, which may limit their generalizability. Additionally, students in our study had access to face-to-face sessions, which likely influenced collaboration dynamics differently than in fully online environments. As hyflex learning is increasingly adopted worldwide, future research should examine how various combinations of modalities (physical/virtual) affect collaboration across different cultural backgrounds, thereby improving the external validity of the findings.

The cross-sectional nature of this study captured data at a single point in time, limiting the ability to infer causal relationships among communication, team atmosphere, and performance. This approach provided only a snapshot of team dynamics, making it difficult to understand how these interactions change over time. Longitudinal studies would allow researchers to track changes over time and assess how these dynamics evolve throughout the course.

Future research should also consider conducting longitudinal studies to examine the evolution of communication effectiveness and team atmosphere over multiple semesters or academic terms. Such studies might shed light on how teams adapt their communication strategies as they progress through hyflex courses, enabling programs and educators to develop more dynamic and responsive instructional strategies that support sustained team performance over time.

## Acknowledgements

The study was funded by Project of Fujian Provincial Social Science Fund (Project No. FJ2025B238), Xiamen Natural Science Foundation Project (No. 3502Z202573039), Key Project of Fujian Provincial Social Science Fund (Project No.FJ2024A025), Humanities and Social Sciences Research Project of the Ministry of Education (No.24YJA760065) and Xiamen Natural Science Foundation Project (No. 3502Z202473049).

## Declaration of Use of Generative AI and AI-Assisted Technologies

While preparing this work, the authors used ChatGPT 4.0 to help generate ideas, enhance clarity, and improve the overall structure of the writing. After using this tool/service, the authors reviewed and edited the content as needed; they take full responsibility for the article's content and references.

# References

Akcaoglu, M., & Lee, E. (2016). Increasing social presence in online learning through small group discussions. *The International Review of Research in Open and Distributed Learning*, *17*(3). https://doi.org/10.19173/irrodl.v17i3.2293

Al-Rahmi, W., Othman, M. S., & Yusuf, L. M. (2015). The role of social media for collaborative learning to improve academic performance of students and researchers in Malaysian higher education. *The International Review of Research in Open and Distributed Learning*, *16*(4). https://doi.org/10.19173/irrodl.v16i4.2326

Bagozzi, R. P., & Yi, Y. J. (1989). On the use of structural equation models in experimental designs. *Journal of Marketing Research*, *26*, 271–284. https://doi.org/10.1177/002224378902600302

Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Services Research*, *41*(4p2), 1576–598. https://doi.org/10.1111/j.1475-6773.2006.00566.x

Beatty, B. J. (2019). *Hybrid-Flexible course design: Implementing student-directed hybrid classes*. EdTech Books. https://edtechbooks.org/hyflex

Bockorny, K. M., Giannavola, T. M., Mathew, S., & Walters, H. D. (2024). Effective engagement strategies in hyflex modality based on intrinsic motivation in students. *Active Learning in Higher Education*, *25*(3), 455–472. https://doi.org/10.1177/14697874231161364

Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, *18*, 1–24. https://doi.org/10.1186/s41239-021-00282-x

Borup, J., Jensen, M., Archambault, L., Short, C. R., & Graham, C. R. (2020). Supporting students during COVID-19: Developing and leveraging academic communities of engagement in a time of crisis. *Journal of Technology and Teacher Education*, *28*(2), 161–169. https://doi.org/10.70725/150472phreji

Bower, M., Dalgarno, B., Kennedy, G. E., Lee, M. J. W., & Kenney, J. (2015). Design and implementation factors in blended synchronous learning environments: Outcomes from a cross-case analysis. *Computers & Education*, *86*, 1–17. https://doi.org/10.1016/j.compedu.2015.03.006

Bozkurt, A. (2019). From distance education to open and distance learning: A holistic evaluation of history, definitions, and theories. In *Handbook of research on learning in the age of transhumanism* (pp. 252–273). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-8431-5.ch016

Brislin, R. W. (1970). Back-Translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*, 185–216. https://doi.org/10.1177/135910457000100301

Brown, M. (2024). Enhancing “hyflex” instruction: Best practices, criticisms, and results from a case

study in graduate education. *International Journal of Studies in Education and Science*, *5*(1), 43–54. https://doi.org/10.46328/ijses.81

Cumming, T. M., Gilanyi, L., & Han, C. (2024). Hyflex delivery mode in a postgraduate course: Instructor and student perspectives. *Discover Education*, *3*(1), 24. https://doi.org/10.1007/s44217-024-00110-0

Detyna, M., Sanchez-Pizani, R., Giampietro, V., Dommett, E. J., & Dyer, K. (2023). Hybrid flexible (hyflex) teaching and learning: Climbing the mountain of implementation challenges for synchronous online and face-to-face seminars during a pandemic. *Learning Environments Research*, *26*(1), 145–159. https://doi.org/10.1007/s10984-022-09408-y

Fan, L., Yang, J., & Shi, J. (2017). The influence of female leadership characteristics on employee performance: Based on the mediating effect of team climate. *Journal of Nanjing Audit University*, 14(4), 34–43. https://link.cnki.net/urlid/32.1867.F.20170629.1654.016

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*, 39–50. https://doi.org/10.1177/002224378101800104

Forsyth, D. R. (2014). *Group dynamics* (6th ed.). Wadsworth Cengage Learning.

Galperin, B. L. (2003). Can workplace deviance be constructive? In A. Sagie, S. Stashevsky, & M. Koslowsky (Eds.), *Misbehavior and dysfunctional attitudes in organizations* (pp. 154–170). Palgrave Macmillan. https://doi.org/10.1057/9780230288829

Garrison, D. R. (2016). *E-learning in the 21st century: A community of inquiry framework for research and practice*. Routledge. https://doi.org/10.4324/9781315667263

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education*, *13*(1–2), 31–36. https://doi.org/10.1016/j.iheduc.2009.10.002

Garrison, D. R., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines*. John Wiley & Sons. https://id.erudit.org/iderudit/1065292ar

González-Romá, V., Fortes-Ferreira, L., & Peiró, J. M. (2009). Team climate, climate strength and team performance: A longitudinal study. *Journal of Occupational and Organizational Psychology*, *82*(3), 511–536. https://doi.org/10.1348/096317908X370025

Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). *A framework for institutional adoption and implementation of blended learning in higher education*. *The Internet and Higher Education, 18*(3), 4–14. https://doi.org/10.1016/j.iheduc.2012.09.003

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. *Advances in Experimental Social Psychology*, *8*, 45–99. https://doi.org/10.1016/S0065-2601(08)60248-8

Hadjipieris, P. A. (2024). *Hyflex course design: Exploring student experiences of an equity-centered technology enhanced curriculum. A qualitative case study* [Doctoral dissertation, University of California, San Diego]. https://scholarworks.calstate.edu/downloads/js956p80v

He, J., Jiang, X., & Yu, G. (2014). Coupling relations of collective efficacy and team performance: The moderating role of team communication. *Science and Technology Management Research*, 34(04), 169–173.

Heilporn, G., & Lakhal, S. (2021). Converting a graduate-level course into a hyflex modality: What are effective engagement strategies? *The International Journal of Management Education*, *19*(1), 100454. https://doi.org/10.1016/j.ijme.2021.100454

Höddinghaus, M., Nohe, C., & Hertel, G. (2023). Leadership in virtual work settings: What we know, what we do not know, and what we need to do. *European Journal of Work and Organizational Psychology*, *33*(2), 188–212. https://doi.org/10.1080/1359432X.2023.2250079

Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). Sage. https://doi.org/10.1177/017084068300400409

Huang, Q. (2011). *Research on the influencing mechanism of transformational leadership on team performance*. [Doctoral dissertation, Wuhan University of Technology]. China National Knowledge Infrastructure.

Jarwati. (2024). Book Review: *IDEAL Distance Education and Blended Learning Handbook* (8th ed.) authored by Vanek, J., Simpson, D., Harris, J., & Goumas, J. International Review of Research in Open and Distributed Learning, 25(1), 157-159, https://doi.org/10.19173/irrodl.v25i1.7253

Johnson, D., & Johnson, R. (1991). *Learning together and alone: Cooperative, competitive, and individualistic learning* (4th ed.). Allyn & Bacon, Inc.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). *Cooperative learning: Improving university instruction by basing practice on validated theory*. https://celt.miamioh.edu/index.php/JECT/article/view/454

Kanaris, M. E., & Mujtaba, B. G. (2023). Trust shaping the social relationship of diverse learners in the online education environment. *Environment and Social Psychology*, *9*(2). https://doi.org/10.54517/esp.v9i2.2197

Ke, F., & Kwak, D. (2013). Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction. *Computers & Education*, *61*, 43–51. https://doi.org/10.1016/j.compedu.2012.09.003

Keshavarz, M. (2023). Hybrid-Flexible course design: Implementing student-directed hybrid classes. *Quarterly Review of Distance Education*, *24*(1), 83-86.

Kohnke, L., & Moorhouse, B. L. (2021). Adopting hyflex in higher education in response to COVID-19: Students' perspectives. *Open Learning: The Journal of Open, Distance and e-Learning*, *36*(3), 231–244. https://doi.org/10.1080/02680513.2021.1906641

Kohnke, L., & Moorhouse, B. L. (2022). Facilitating synchronous online language learning through Zoom. *RELC Journal*, *53*(1), 296–301. https://doi.org/10.1177/0033688220937235

Lohmann, M. J., Randolph, K. M., & Oh, J. H. (2021). Classroom management strategies for hyflex instruction: Setting students up for success in the hybrid environment. *Early Childhood Education Journal*, *49*(5), 807–814. https://doi.org/10.1007/s10643-021-01201-5

Lyons, E. (2024). *Teaching and social presence in hyflex classrooms: Setting the climate for effective group cohesion* [Unpublished master's thesis]. University of Windsor.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*(3), 356–376. https://doi.org/10.5465/amr.2001.4845785

Martin, F., Wu, T., Wan, L., & Xie, K. (2022). A meta-analysis on the community of inquiry presences and learning outcomes in online and blended learning environments. *Online Learning*, *26*(1), 325–359. https://doi.org/10.24059/olj.v26i1.2604

Mathieu, J. E., Gilson, L. L., & Ruddy, T. M. (2006). Empowerment and team effectiveness: An empirical test of an integrated model. *Journal of Applied Psychology*, *91*(1), 97. https://doi.org/10.1037/0021-9010.91.1.97

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410–476. https://doi.org/10.1177/0149206308316061

Miao, J., & Ma, L. (2022). Students' online interaction, self-regulation, and learning engagement in higher education: The importance of social presence to online learning. *Frontiers in Psychology*, *13*, 815220. https://doi.org/10.3389/fpsyg.2022.815220

Miller, J., Risser, M., & Griffiths, R. (2013). Student choice, instructor flexibility: Moving beyond the blended instructional model. *Issues and Trends in Educational Technology*, *1*(1), 8–24. https://doi.org/10.2458/AZU_ITET_V1I1_16464

Moore, M. G. (2013). The theory of transactional distance. In *Handbook of distance education* (pp. 66–85). Routledge. https://doi.org/10.4324/9780203803738

Murray, S. C. (1983). [Review of the book *Groupthink: Psychological studies of policy decisions and fiascoes*. By I. L. Janis]. *Presidential Studies Quarterly, 13(4)*. 654-656. https://www.jstor.org/stable/27547987

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality &*

*Quantity*, *41*, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Penrod, J. M. (2023). *Considerations for implementing and supporting hybrid and hyflex learning in a higher education institution*. Ohio University.

Priem, R. L., Harrison, D. A., & Muir, N. K. (1995). Structured conflict and consensus outcomes in group decision making. *Journal of Management*, *21*(4), 691–710. https://doi.org/10.1016/0149-2063(95)90006-3

Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2023). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments*, *31*(4), 2371–2391. https://doi.org/10.1080/10494820.2021.1884886

Raes, A., Detienne, L., Windey, I., & Depaepe, F. (2020). A systematic literature review on synchronous hybrid learning: Gaps identified. *Learning Environments Research*, *23*, 269–290. https://doi.org/10.1007/s10984-019-09303-z

Razali, S. N., Shahbodin, F., Hussin, H., & Bakar, N. (2015). Factors affecting the effective online collaborative learning environment. In *Pattern analysis, intelligent security and the Internet of Things* (pp. 215–224). Springer International Publishing. https://doi.org/10.1007/978-3-319-17398-6_20

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork? *Small Group Research*, *36*(5), 555–599. https://doi.org/10.1177/1046496405277134

Sargeant, J., Loney, E., & Murphy, G. (2008). Effective interprofessional teams: "Contact is not enough" to build a team. *Journal of Continuing Education in the Health Professions*, *28*(4), 228–234. https://doi.org10.1002/chp.189

Thurnau, L., & Byun, J. (2022). HyFlex: A hybrid model providing flexibility for learners. *C2C Digital Magazine*, *1*(17), 7. https://scholarsspace.jccc.edu/c2c_online/vol1/iss17/7

Triandis, H. C. (2018). *Individualism and collectivism*. Routledge. https://doi.org/10.4324/9780429499845

Vetrivel, S. C., & Mohanasundaram, T. (2024). Beyond the blackboard: Embracing hybrid learning spaces. In *Global perspectives on micro-learning and micro-credentials in higher education* (pp. 10–28). IGI Global. https://doi.org10.4018/979-8-3693-0343-6.ch002

Wang, H. (2008). *The study on influence between the team process and the team effectiveness*. [Doctoral dissertation, Tianjin University of Finance and Economics].

Wilson, T. J., & Alexander, M. (2021). HyFlex course delivery: Addressing the change in course modality brought on by the pandemic. *Journal of the International Society for Teacher Education*, *25*(2), 41–58. https://doi.org/10.26522/jiste.v25i2.3668

Wong, B. T., Li, K. C., Chan, H. T., & Cheung, S. K. (2023). HyFlex learning research and practice: A longitudinal analysis. *Sustainability*, *15*(12), 9699. https://doi.org/10.3390/su15129699

Yamagata-Lynch, L. C. (2014). Blending online asynchronous and synchronous learning. *International Review of Research in Open and Distributed Learning*, *15*(2), 189–212. https://doi.org/10.19173/irrodl.v15i2.1778

Zhang, Y., & Li, D. (2024). Enhancing project performance through effective team communication: A comprehensive study integrating project management quotient, trust, and management information systems. *Journal of Information Systems Engineering and Management*, *9*(1), 25574. https://doi.org/10.55267/iadt.07.14317

Zheng, B. (2006). *A study of the mechanism within leadership style, organizational citizenship behavior and team performance*. [Master's thesis, Zhejiang University].


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Tutors' Perspectives of Advancing Distance Learning Programs: A Comprehensive Understanding

**Mohammad Rezaul Islam[1] and Md Nazim Mahmud[2]**
*[1]School of Education, Bangladesh Open University*
*[2]Department of Educational Administration, Foundations, and Psychology, Faculty of Education, University of Manitoba*

## Abstract

This study critically examined tutors' perspectives on advancing the academic development of teacher education programs delivered via open and distance learning (ODL) at Bangladesh Open University (BOU). Tutors play a pivotal role as frontline facilitators of instruction, yet their experiential insights are often underrepresented in institutional decision-making. Drawing on a constructivist paradigm and grounded theory methodology, this qualitative inquiry engaged 82 tutors across eight tutorial centres using open-ended survey questions. Through classical content analysis, eleven major themes emerged, including attendance in tutorial sessions, curriculum and module design, tutorial session frequency, physical resources, tutor professional development, and supervision of practice teaching. The findings reveal that tutors emphasise the need for structured learner engagement, participatory curriculum revision, robust infrastructural support, and institutional investment in tutor capacity-building. The study also highlights disparities between current program structures at BOU and international norms, suggesting the need for extended program duration and more integrated practicum experiences. Implications are drawn for institutional policy, academic design, and participatory governance in ODL. By foregrounding tutors' voices, this study contributes to a more inclusive model of academic development and underscores what tutors perceive as the need to bridge the gap between policy directives and pedagogical realities in distance education.

*Keywords*: open and distance learning, ODL, teacher education programs, tutors' perspectives, academic development, Bangladesh Open University

# Introduction

In contemporary higher education, the continual enhancement and restructuring of academic programs are central to aligning institutional offerings with evolving pedagogical standards and accreditation benchmarks (Klement et al., 2017). Nowhere is this more salient than in open and distance learning (ODL) systems, where innovation must be matched by responsiveness to the unique demands of distance learners and the mediators who facilitate their academic journeys. Among those mediators, tutors play an indispensable role, often functioning as the academic face of the university and embodying the bridge between institutionally designed curriculum and learner engagement (Ntuli & Gumbo, 2019; Sutherland, 2018).

While educational reforms are typically driven from the top down, research consistently suggests that meaningful curriculum development must be informed by those responsible for enacting policy at the classroom level—namely, tutors (Bascia et al., 2014; Johnson, 2001). In the ODL context, tutors serve multiple functions: they interpret course content, diagnose learner needs, deliver instruction, offer emotional support, and liaise between learners and institutions (Tait, 2000). Their role is therefore uniquely positioned to provide crucial feedback that could inform academic program improvement. However, tutors at the Bangladesh Open University (BOU)—the country's sole national provider of ODL—perceive that they have had limited involvement in curriculum decision-making processes, even though they consider their role central to program delivery.

Given that BOU's teacher education programs are designed for working professionals, particularly school teachers, there is a need to continually evaluate and strengthen the academic scaffolding of these offerings. The absence of a formal platform for tutor voices in academic development represents a missed opportunity for grounded, context-sensitive reform. This study attempted to bridge that gap by systematically investigating tutors' perspectives on how teacher education programs could be academically enhanced, asking how BOU tutors perceive and recommend academic development strategies for the teacher education programs in the context of ODL. In addressing this question, the study positioned tutors not merely as implementers of policy but as vital stakeholders in academic innovation. The aim was to provide policymakers, curriculum designers, and institutional leaders with empirically grounded insights for the ongoing enhancement of teacher education programs within ODL environments.

# Tutors of Bangladesh Open University

Tutors at the School of Education (SoE) of BOU are appointed through referrals from the study centres, which serve as the university's examination and tutorial centres. Most of these tutors work as academics or regular teachers at the study centres. The university often selects the tutor if the recommendation complies with its appointment rules. Following program enrollment, learners are directed to attend tutorial sessions by the SoE. Usually, that is the first time tutors and students meet; before then, the tutors have no responsibility for those learners.

The tutorials, which generally consist of 16 sessions for each program, are scheduled by SoE. Students must turn in their assignments to their tutors. The SoE teachers determine assignment themes. Following the conclusion of the tutorial sessions, the SoE releases the exam schedule, and study centres carry out the tasks

associated with the test. This suggests that the primary function of administering examinations is carried out by the study centre tutors. Tutors and the teachers of the SoE evaluate the answer scripts, while assignments are evaluated only by the tutors. The university's exam division produces the final results for learners. Based on these tasks and responsibilities, it can be concluded that tutors' primary duties are to lead tutorial sessions and conduct examination-related activities. In the ODL system, tutors are appointed as part-time and contractual staff to deliver supporting instruction, with minimal opportunity to participate in the academic development of programs. Even though they are well-known and respected academicians elsewhere, the second-class position (Lentell, 2001) in the job can demotivate and have a long-term effect on their professional activities in the ODL system (Shelley et al., 2006).

# Literature Review

## Tutors as Facilitators of Learning in Open and Distance Education

ODL programs such as bachelor's and master's degrees in education depend heavily on tutors to support student learning. As primary links between learners and institutions, tutors provide academic and emotional support in place of regular in-person instruction (Ntuli & Gumbo, 2019). Tutors mediate between course content and student needs, acting as mentors and facilitators (Sutherland, 2018). Particularly in large ODL institutions, tutors significantly shape students' academic experiences.

Tait (2000) outlined three types of learner support in ODL: cognitive, affective, and systemic. Tutors deliver cognitive support through instruction and feedback, affective support through encouragement, and systemic support by helping students navigate institutional requirements. Tutors assume multiple roles, including explaining content, moderating discussions, and motivating learners. The University of South Africa (UNISA) tutor model exemplifies this multifaceted role (Barrow & Grant, 2012). Tutors are crucial to bridging gaps between students and institutional structures (Leibowitz et al., 2015), especially in teacher education programs, where they model both academic content and pedagogical practices.

## Academic Support and Development Through Tutoring

Tutors play a vital role in promoting academic development by helping students build knowledge, skills, and critical thinking. Student support in ODL comprises academic and non-academic dimensions. Tutors contribute to academic success by clarifying content, scaffolding learning, and providing timely feedback (Sánchez-Elvira Paniagua & Simpson, 2018).

In distance-based teacher education programs, tutors combine subject-matter expertise with pedagogical guidance. They support constructivist learning approaches by encouraging dialogue and knowledge construction (Li et al., 2017). Tutors also indirectly influence curriculum through feedback loops, identifying student difficulties that may guide content revision. Although they seldom influence curricular policy directly, they personalize the learning experience in ways static course materials cannot. Their support is linked to increased persistence and success (Mori et al., 2022; Walker, 2015).

## Pedagogical Practices and Tutor Professional Development

The quality of tutoring depends on tutors' pedagogical skills, including digital facilitation and student engagement. Researchers de Metz and Bezuidenhout (2018) found that tutors often face mismatches

between institutional expectations and their actual training. While many tutors see themselves as central to student learning, they report feeling underprepared in key areas, including online teaching and technology use (Hofer et al., 2021).

Effective tutoring requires continuous professional development. Tutors trained in student-centred techniques and adult learning principles are better equipped to foster engagement. Yet some studies have revealed persistent gaps in tutors' instructional alignment with ODL pedagogical best practices (Govender, 2018). Institutions are urged to provide support that enables tutors to meet these challenges.

Collaborative learning opportunities, such as peer mentoring and communities of practice, have been found to enhance tutor effectiveness. Tutors engaged in structured peer-exchange programs have reported adopting improved instructional strategies (Muazam et al., 2021; Walker, 2015). These exchanges reinforce tutors' professional growth and align practices across programs. As ODL expands, especially post-pandemic, the need for pedagogically skilled tutors is greater than ever (Rapanta et al., 2020; Mori et al., 2022).

## Institutional Roles and Policy Engagement of Tutors

Despite their instructional importance, tutors often have limited roles in institutional decision-making. Typically hired as part-time faculty, they may have minimal involvement in curriculum design or governance (Barrow & Grant, 2012; Sutherland, 2018). Felber (2020) highlighted concerns over tutors teaching courses they did not design, which limits their ability to improve content.

In many cases, tutor feedback is not meaningfully incorporated into program development. Even structured tutor models like UNISA's provide limited pathways for tutors' voices to influence policy (Ntuli & Gumbo, 2019). This lack of agency is problematic, as tutors hold valuable insights about student needs and course delivery. Yet, institutional systems often relegate tutors to delivery roles without avenues for curriculum review or policy dialogue (Leibowitz et al., 2015).

Researchers and organizations have been advocating for tutor inclusion in institutional development. The Commonwealth of Learning and UNESCO stressed the need to value tutor input as part of a renewed educational social contract (UNESCO & International Task Force on Teachers for Education 2030, 2024). Felber (2020) proposed practical strategies, such as involving adjunct tutors in course revision workshops. Institutionalizing tutor engagement could improve curriculum relevance, strengthen tutor commitment, and elevate student experience. Still, barriers such as limited time and institutional resistance persist. Tutors have expressed frustration with top-down decision-making that fails to reflect their frontline realities (Tait, 2018).

## Tutors' Contributions to Quality and Student Success

Though often marginal in institutional structures, tutors substantially impact academic quality and student success. Studies have confirmed that strong tutoring correlates with better pass rates, retention, and satisfaction. At UNISA, e-tutors reported helping students succeed by clarifying content and maintaining motivation (Joubert & Snyman, 2020).

Tutors also contribute to the quality of the learning process, not just outcomes. Their feedback helps correct misunderstandings and ensures academic integrity. Through regular interaction, tutors foster student accountability and deeper engagement. Researchers de Metz and Bezuidenhout (2018) described e-tutors as "critical success factors" in shaping students' perceptions of online learning (para. 1).

Institutional support is essential to maximize tutors' contributions. High tutor-student ratios or limited resources can constrain personalized support. Nonetheless, many tutors adapt creatively by offering additional help or materials. Their commitment to student learning reflects a broader professional ethos. However, their limited input into policy remains a challenge. Tutors carry much responsibility for academic standards but often lack influence over institutional frameworks. This gap hinders sustainable improvements. Still, tutors take initiative to meet learning goals through supplemental practices.

In sum, tutors in ODL programs are integral to academic success, pedagogical quality, and institutional credibility. They work at the intersection of learner support and content delivery, particularly in professional programs in education. Yet their expertise remains underused in governance and curriculum design. Expanding professional development and engaging tutors in institutional processes will be crucial to realizing the full potential of ODL education.

# Methodology

This study adopted a qualitative research approach, framed within the constructivist paradigm which posits that reality is socially constructed and contextually situated (Charmaz, 2014; Lincoln & Guba, 1985). This methodological framework enabled a nuanced exploration of tutors' perspectives within the contextual boundaries of BOU's ODL system. While the study was not designed to generate formal grounded theory in the traditional Glaserian or Straussian sense, it adopted grounded theory principles—particularly from Charmaz's (2014) constructivist tradition—as an analytic orientation. These principles included iterative engagement with the data, constant comparison across participants' responses, and the emergence of inductively generated categories. However, classical content analysis (Leech & Onwuegbuzie, 2007) served as the primary analytic strategy to identify thematic frequencies and institutional patterns. Thus, the study does not claim theory development but rather uses grounded theory concepts to deepen the interpretive process and ensure analytical transparency. This approach aligns with constructivist paradigms that value co-constructed meaning and recognise the researcher's interpretive role

Constructivism guided the formulation of open-ended, context-sensitive questions aimed at eliciting participants' lived realities rather than imposing pre-defined categories. The survey prompts were deliberately open to multiple interpretations to reflect the socially constructed nature of knowledge. This orientation shaped the design and language of data collection tools, inviting tutors to co-construct meaning rather than act as passive data sources. Grounded theory principles further structured the analytic process, allowing emergent themes to surface inductively from tutors' narratives. The iterative coding, categorisation, and constant comparison methods helped reveal latent patterns embedded within tutor perspectives, thereby grounding theoretical insights in practice-based realities (Birks & Mills, 2015; Charmaz, 2014).

## Sampling and Participants

A purposive sampling strategy was employed to ensure the selection of information-rich participants who could offer in-depth insights into the academic development of the distance education programs (Creswell & Creswell, 2017). The study targeted tutors from the SoE's network of 27 tutorial centres across Bangladesh, which constitutes the broader population of interest, comprising 400 tutors. From this population, 82 tutors were chosen from eight conveniently located tutorial centres based on their accessibility and operational capacity during the study period. Although tutors were selected through random inclusion within those centres, purposive attention was given to including coordinators of the programs as key informants, which tutors believed was due to their strategic roles in program implementation. Experienced tutors, who were available in the tutorial sessions during the data collection process, were prioritized for selection as sample participants, ensuring that the most knowledgeable and engaged educators informed our study. Moreover, these eight centres were situated in eight divisional cities of Bangladesh and deemed the most significant among the 27 centres for size and geographical familiarity. These eight study centres were the oldest government teacher training colleges in the country and held a leading position among all teacher training centres. Although sampling only from these centres may have restricted generalizability, the selected centres offered appropriately pertinent insights into the phenomena being examined.

The sample size of 82 participants significantly exceeds the commonly cited threshold for grounded theory studies, where 20 to 30 participants are often considered sufficient to achieve theoretical saturation (Creswell & Creswell, 2017). The relatively large sample size thus strengthens the robustness and depth of the data, enhancing the interpretive potential of the findings.

## Data Collection

Data were gathered through a qualitative survey comprising both closed and open-ended questions in the written questionnaire. The survey was designed to elicit tutors' perceptions regarding factors affecting the academic development of the teacher education programs (Bachelor of Education and Master of Education). Specifically, participants were asked questions in three key areas: (a) demographic background; (b) identification of areas requiring improvement; and (c) suggestions for enhancing academic quality in the identified areas.

The open-ended questions were particularly valuable in capturing nuanced and context-specific viewpoints, which aligns with the grounded theory emphasis on inductively building theory from rich qualitative data (Birks & Mills, 2015).

## Data Analysis

Upon gathering data from the participants, all information was meticulously documented and analyzed to discern themes or patterns. Thematic analysis approaches were followed to understand the underlying meaning of the data, and meaningful themes were then employed to analyze the study further, enhancing its systematic nature, rigor, and quantifiability. Classical content analysis was selected over thematic analysis due to its structured, frequency-based coding logic, which aligned with the study's aim to not only interpret emergent themes but also assess their relative salience across a large qualitative sample ($n$ = 82). Unlike thematic analysis, which prioritises deep semantic interpretation without quantitative enumeration,

classical content analysis facilitates the systematic categorisation and quantification of textual data into recurring codes (Leech & Onwuegbuzie, 2007). This method was considered especially suitable given the study's intent to highlight dominant institutional concerns based on tutor consensus.

While the coding process was influenced by grounded theory's iterative logic and constant comparative methods, the study primarily employed classical content analysis to categorise and quantify recurring themes (Charmaz, 2014; Leech & Onwuegbuzie, 2007). The coding process followed a multi-step procedure. First, both authors independently engaged in open coding of a shared subset of responses (25%) to develop an initial codebook. Codes were refined through iterative discussions to achieve conceptual clarity and alignment. Intercoder reliability was established by calculating agreement on the application of codes across three rounds of coding, yielding an average agreement rate of 88%, which exceeded the commonly accepted threshold of 80% for qualitative studies (Miles et al., 2014). The final codebook was then applied to the full dataset using manual coding. Codes were grouped into higher-order categories, which formed the basis for the 11 emergent themes. This analytical rigor has ensured transparency, reproducibility, and validity in capturing tutors' perspectives across diverse institutional settings.

## Trustworthiness and Researcher Reflexivity

To ensure the trustworthiness and credibility of the findings, the researchers engaged in prolonged engagement within the study sites. Time was spent at each tutorial centre to observe contextual dynamics and build rapport with participants, facilitating a richer understanding of the educational phenomena under investigation (Lincoln & Guba, 1985).

Further, reflexive practices were employed to minimise researcher bias, and member checking was performed informally through follow-up communication with selected tutors to verify the accuracy of interpretation and thematic classification. These procedures enhanced the validity and authenticity of the findings, ensuring that they faithfully represented participants' perspectives.

The first author is an academic staff member at the School of Education, Bangladesh Open University (BOU), the institution under study. This insider status provided privileged access to participants, an in-depth understanding of institutional structures, and nuanced interpretation of tutors' experiences. However, acknowledging the risk of role duality and potential bias, several measures were adopted to uphold analytical objectivity. These included maintaining professional-researcher boundaries during data collection, ensuring participant anonymity through written surveys, and adhering to a consistent coding protocol. The second author—who is affiliated with a Canadian university and holds no professional ties to BOU—brought an external, critical perspective to both data analysis and interpretation. The collaborative coding process, undertaken jointly, served as a methodological check to enhance interpretive neutrality. Such reflexive practices are consistent with established qualitative research standards that emphasise positionality awareness and researcher transparency (Berger, 2015).

## Theoretical Integration and Analytical Framing

The constructivist paradigm not only framed the study's epistemological stance but also guided the interpretive process by foregrounding tutors' situated perspectives as co-constructors of knowledge. Grounded theory, used not for theory generation in its purest form but as a methodological guide, allowed

for concepts to emerge organically through systematic data engagement. The authors resisted imposing theoretical models at the outset, instead letting patterns and categories emerge from tutors' voices. This inductive process reflects the bottom-up ontology of grounded theory and constructivism, privileging tutor agency and meaning-making. For example, themes such as professional development and curriculum co-design emerged not merely as procedural gaps but as indicators of unequal power dynamics within ODL governance. These dynamics resonate with broader theoretical critiques of institutional control in distance education (Barrow & Grant, 2012; Tait, 2018), where tutors are often relegated to delivery roles without substantive input into curriculum or policy.

### Ethics Statement

All procedures adhered to the ethical standards established by the Research Evaluation Committee of the School of Education, BOU. The approval (Ref: BOU/SOE/1(64))/13/99) was obtained from the committee, and no risks connected with this research were anticipated or foreseen.

## Findings and Discussion

The thematic categories presented in this section are not just descriptive classifications but reflect deeper socio-institutional dynamics embedded in the ODL context. By applying a constructivist lens, tutors' narratives have been treated as interpretive acts shaped by institutional cultures, personal teaching histories, and pedagogical ideologies. Moreover, the emergent themes provide an empirical basis for theorising power relations, voice, and governance in distance education. For instance, tutors' appeals for curriculum inclusion and supervised practicum mirror tensions between top-down educational design and bottom-up experiential expertise. These findings critique the asymmetry between institutional authority and tutor agency—a theme underexplored in much of the existing ODL literature.

### Demographic Characteristics of Tutors

A total of 82 tutors—50 from the Bachelor of Education and 32 from the Master of Education programs—participated in this study. The majority of tutors were male (76%), with females comprising 24%. Tutors possessed substantial professional experience, with an average of 19 years in the teaching profession and approximately 14 years of affiliation with BOU. This depth of experience provided a valuable lens into the academic and operational strengths and shortcomings of the teacher education programs. Their sustained association with BOU lends credibility to their insights regarding program enhancement and innovation.

### Emergent Themes

To systematically capture tutors' perspectives on the academic development of teacher education programs, qualitative responses were subjected to classical content analysis. Recurring patterns were thematically coded, yielding the frequency-based categories shown in Table 1.

**Table 1**

*Themes Identified by Bangladesh Open University Tutor Participants*

| Theme | Frequency |
|---|---|
| Attendance in tutorial sessions | 21 |
| Curricula and modules | 14 |
| Tutorial sessions | 12 |
| Physical resources | 12 |
| Professional development of tutors | 12 |
| Academic and practice teaching supervision | 12 |
| Role of coordinators | 3 |
| Assignments, research, and assessment | 2 |
| Co-curricular and extracurricular activities | 2 |
| Examination and scheduling | 2 |
| Program duration and structure | 2 |

In the sections that follow, a critical synthesis of the findings based on the tutors' perspectives concerning the academic development of the teacher education programs offered through BOU is presented, organized by theme. Each theme has been elaborated upon with theoretical grounding and contextual insight.

## *Attendance in Tutorial Sessions*

Tutor responses strongly emphasised the need for mandatory attendance in tutorial sessions, which was the most recurrent theme in the dataset ($n$ = 21). Although the ODL model traditionally allows flexibility, tutors underscored that teacher education programs require active engagement through regular in-person interaction to develop pedagogical competence. This demand reflects the practical nature of teaching, where tutorial sessions function as essential spaces for modeling classroom practices. One of the participants said: "In the case of [the Bachelor of Education program], practical learning is not possible without a face-to-face class. Moreover, 70–80% of students did not attend the tutorial sessions. Therefore, the objective of [the Bachelor of Education degree] is not achieved without class attendance. The SoE should look into the matter for the development of the program."

Empirical studies support this claim, suggesting a positive correlation between attendance and academic performance in blended and distance modalities (Nkolo, 2021). However, this position is problematised by research indicating no direct correlation in purely distance contexts (van Zyl et al., 2012). Moreover, Olivier (2016) observed that well-designed instructional modules can sometimes reduce learners' motivation to attend tutorials. Hence, while tutors advocate attendance-based assessment mechanisms, this raises broader questions about balancing learner autonomy with pedagogical accountability in ODL frameworks.

## *Curricula and Modules*

Tutors identified outdated and rigid curricular materials as a barrier to meaningful learning ($n$ = 14). Their suggestions included periodic revision, increased digital integration, and open calls for module development involving practitioners. One participant stated, "Transparency should be incorporated into the writer selection process through an open advertisement when developing modules. The instructional resources contain a significant amount of outdated information that should be removed." Another participant suggested "conducting workshops with concerned tutors of related subjects before developing the modules of that subject. SoE should include tutors from different study centres in the curriculum committee of the programs."

These highlight a tension between top-down curriculum design and the lived realities of those who deliver instruction. Scholars have argued that static printed modules can stifle pedagogical responsiveness (Woo, 2011). In ODL contexts, module quality directly impacts learner outcomes (Sembiring, 2020). Tutors' exclusion from curriculum development processes echoes broader critiques of centralised educational governance that marginalises implementers. The call for decentralised, participatory module design aligns with contemporary instructional design principles emphasising contextual relevance and tutor autonomy (Conole, 2013).

## *Tutorial Sessions*

The limited number of tutorial sessions was viewed as insufficient for supporting learners' pedagogical development ($n$ = 12). Tutors proposed increasing frequency, even suggesting bi-weekly formats. One participant commented:

> In the case of the Bachelor of Education program, the School of Education should enhance the total number of practice teaching sessions. More emphasis must be placed on practical sessions (practice teaching) rather than theoretical ones. It is also necessary to expand the number of tutorial sessions in the Master of Education program.

The tutorial was seen not just as a problem-solving platform but a space for learner-tutor bonding and motivation. This reflects the tutorial's broader pedagogical function, as tutors serve not merely as content facilitators but as affective and cognitive mediators (Segoe, 2014). The findings support Sembiring's (2020) view that tutorial support is a primary driver of academic excellence. However, contrasting findings (McAndrew et al., 2010; Sembiring, 2017) have suggested that learner independence and access to digital resources may offset limited tutorial access. Thus, expanding tutorial frequency should be evaluated alongside complementary strategies such as blended learning and peer-supported environments.

## *Physical Resources*

A consistent concern among tutors was the lack of infrastructural support in tutorial centres ($n$ = 12). Recommendations included developing dedicated academic buildings, libraries, computer labs, and multimedia support. One participant discussed the difficulties of having too many students in a tutorial session: "For the tutors to conduct class in a planned way, the student-teacher ratio shouldn't be greater than 1:40." While modern learners increasingly access e-resources (Olaniran et al., 2017), the absence of physical infrastructure in rural and semi-urban contexts perpetuates inequity. The international literature

has acknowledged that localising resources and decentralising access points are key strategies to managing resource constraints in ODL (Lentell & O'Rourke, 2004). This issue raises broader concerns about digital inequality and institutional readiness in distance education ecosystems.

## *Professional Development of Tutors*

Tutors reported limited institutional support for continuous professional development (n = 12). They advocated for regular workshops, research support, and opportunities to pursue higher education (e.g., PhD or MPhil). This reflects a broader discourse on professionalization in ODL, where tutors often face professional isolation. ODL tutors occupy multiple roles—facilitators, mentors, assessors—and require context-specific training (Segoe, 2014; Xiao, 2016). However, institutional structures often fail to recognise or remunerate this complexity. Becher and Trowler (2001) argued that professional identity in higher education is co-constructed through communities of practice, which are largely absent in ODL environments. Institutional investment in tutor development is thus critical not only for academic outcomes but also for staff retention and morale.

## *Academic and Practice Teaching Supervision*

Supervised practice teaching conducted by learners emerged as a key concern ($n$ = 12). Tutors proposed doubling the number of sessions and ensuring institutional linkage with practice schools. One participant emphasized the importance of practice teaching: "There should be an opportunity to supervise the teaching practice activities of the learners at the schools where they are practicing. This initiative could improve the quality of the learners." The current ad hoc approach, lacking field supervision, undermines one of the core goals of teacher education: developing classroom competence. This concern aligns with studies that highlight the centrality of experiential learning in teacher training (Sahoo & Chandra, 2014). While logistical challenges are real, some institutions have responded by incorporating peer teaching and simulated teaching exercises to compensate (Mokoena, 2017). Without adequate field support, the teacher education program risks becoming overly theoretical and disconnected from practical realities.

## *Role of Coordinators*

Although mentioned by only three tutors, the role of coordinators emerged as strategically significant. Coordinators are essential in aligning tutorial schedules, resolving administrative challenges, and serving as communication bridges between institutional leadership and ground-level tutors. One participant noted, "Coordinators should be empowered to align the academic calendar more realistically." Their marginalisation in current institutional structures undermines responsiveness and continuity in program delivery. Felber (2020) also stressed the importance of mid-level academic actors in improving instructional quality. Enhancing coordinators' capacity through formalised roles and decision-making authority can facilitate more coherent academic planning across tutorial centres.

## *Assignments, Research, and Assessment*

Although mentioned by only two respondents, the issue of assignment and research contextualisation signals deeper pedagogical misalignments. Tutors recommended aligning tasks more closely with learners' instructional realities to foster relevance and engagement. One tutor suggested making assignments "more relevant to students' teaching contexts," a view supported by Jegathesan et al. (2018), who found that

situational tasks improved knowledge application and learner retention in Malaysian ODL institutions. Sembiring (2017) also reported that students perceived assignments as more impactful than exams in promoting reflective learning. These insights collectively point to the need for a restructured assessment model that favours authentic, performance-based tasks and better supports learner agency.

### *Co-Curricular and Extracurricular Activities*

The emphasis by two tutors on co-curricular and extracurricular activities reflects an emerging concern about holistic learner engagement in ODL contexts. One participant suggested the university "assign marks on extracurricular activities and make it mandatory," highlighting a perceived void in learner motivation and community building. Though infrequently mentioned, such suggestions align with contemporary educational theories promoting socio-emotional learning and peer bonding (Rapanta et al., 2020). Incorporating co-curricular initiatives could mitigate learner isolation—an often-cited barrier in ODL—and promote academic perseverance through identity formation and inclusive learning environments.

### *Examination and Scheduling*

Although only two tutors raised concerns about examination management, their feedback highlights administrative challenges that can erode student trust. One respondent suggested "exam schedules should align with weekends and be announced earlier," reflecting the unique needs of working learners enrolled in BOU's ODL programs. These concerns echo Nkolo's (2021) findings, which identified assessment scheduling and communication gaps as significant barriers to student satisfaction. Addressing such logistical issues through learner-informed timetabling and grievance mechanisms can significantly enhance the credibility and transparency of the academic process in ODL systems.

### *Program Duration and Structure*

Tutors recommended extending the current duration of teacher education programs at BOU from 1 year (two semesters) to 2 years (four semesters). One participant stated, "Increase the time duration of both programs to 2 years or four semesters," underscoring the concern that the existing timeframe constrains pedagogical depth and practical exposure. To further investigate this concern, we conducted a comparative analysis of teacher education program durations across a strategically selected set of international ODL institutions. These universities were chosen based on three key criteria: (a) the provision of distance-delivered teacher education programs comparable to BOU's bachelor's and master's degrees in education; (b) geographic and economic diversity, including both Global South and Global North contexts; and (c) institutional maturity and recognition in the field of open and distance education. Institutions such as Indira Gandhi National Open University, Allama Iqbal Open University, and Open University Malaysia represent South and Southeast Asian models that, like BOU, serve large numbers of in-service teachers through blended and distance modes. Meanwhile, universities such as The Open University (UK), Athabasca University (Canada), and Charles Sturt University (Australia) offer relevant insights from more mature ODL systems with longer-established academic frameworks.

Notably, differences in program duration across these institutions are shaped by contextual factors including national teacher certification standards, curriculum design mandates, practicum integration requirements, and broader labour market expectations (Rapanta et al., 2020; UNESCO & International Task Force on Teachers for Education 2030, 2024). For example, while shorter durations are common in

some South Asian contexts due to flexible delivery for working professionals, extended formats in Australia, the UK, and Canada often reflect more intensive practicum demands and layered academic progression. Tutors believe that as a result, while absolute duration varies, the overall trajectory suggests that BOU's 1-year model may lack the necessary structural depth and experiential scope observed internationally—particularly in areas such as field supervision, coursework sequencing, and student support mechanisms.

Comparative analysis (see Table 2) confirms that BOU's program duration is shorter than those at comparable institutions, such as Indira Gandhi National Open University and Charles Sturt University, where 2- to 3-year programs are standard. Longer durations allow for expanded practicum experiences and reflective learning, as supported by Sahoo and Chandra (2014), who emphasised the value of experiential learning in teacher education. Extending the program would also enhance academic parity and facilitate international credit recognition.

**Table 2**

*Comparative Duration of Teacher Education Programs by Institution and Region*

| Region | University | Bachelor of Education program duration (years) | Master of Education program duration (years) |
|---|---|---|---|
| Asia | Indira Gandhi National Open University, India | 2 | 2 |
| Asia | Allama Iqbal Open University, Pakistan | 1.5 | Not offered |
| Asia | Open University Malaysia | Not offered | 2 |
| Africa | Open University of Tanzania | 3 | 2 |
| Europe | The Open University, UK | 6 (part-time) | 3 |
| North America | Athabasca University, Canada | Not offered | 4 (part-time) |
| Oceania | Charles Sturt University, Australia | 3 | 2 |

*Note*. Data sourced from: Indira Gandhi National Open University (2023, n.d.); Allama Iqbal Open University (n.d.); Open University Malaysia (n.d.); Open University of Tanzania (n.d.-a., n.d.-b.); The Open University (n.d.-a., n.d.-b.); Athabasca University (n.d.); and Charles Sturt University (n.d.-a., n.d.-b.).

This comparison suggests that the current duration of the teacher education programs at BOU could benefit from further alignment with international practices. Expanding the program structure has the potential to enhance academic rigour while better supporting learners in achieving comprehensive professional development outcomes.

## Study Implications

The study's findings yield multiple interrelated implications for institutional policy, pedagogical practice, and future reform strategies within ODL-based teacher education. By linking tutors' grounded experiences to systemic structures, the study extends Tait's (2000, 2018) learner support framework. While Tait

categorises tutor roles into cognitive, affective, and systemic support, this study reveals the constraints tutors face in fulfilling these roles due to their marginal position in governance. The results advocate for expanding Tait's model by integrating into it dimensions of power, voice, and institutional participation, which tutors believe will result in a more critical understanding of tutor roles within the governance of ODL. Tutors are not just support agents; they are epistemic actors whose exclusion undermines program effectiveness and equity. The implications of this study are detailed in the sections that follow.

## Rethinking Engagement Policies

While learner autonomy is a hallmark of distance education, the demand for compulsory tutorial attendance reveals an urgent need to balance flexibility with structure. Institutions could consider hybrid attendance policies—combining face-to-face sessions with digital engagement metrics—to ensure pedagogical consistency without undermining learner autonomy.

## Participatory Curriculum Development

Tutors' exclusion from curriculum design suggests a top-down model of instructional governance. Establishing tutor advisory boards, conducting consultative curriculum reviews, and implementing open calls for module contributors would foster shared ownership and contextual alignment of learning materials.

## Institutionalisation of Professional Development

The ad hoc and inequitable nature of tutor training warrants a formalised, institutionally funded professional development framework. This could include regular seminars, pedagogical certifications, and financial incentives for higher studies, enabling tutors to remain academically and technologically agile.

## Structuring Supervised Practicum Models

Unsupervised or poorly managed teaching practice weakens the pedagogical foundation of teacher education programs. BOU should institute supervised practicum models in partnership with local schools and incorporate blended practicum approaches (peer-teaching, video portfolios) to overcome logistical constraints.

## Strengthening Academic Infrastructure

The lack of physical and technological infrastructure—especially in rural tutorial centres—exacerbates inequalities in the ODL system. Policy reforms must allocate decentralised funds for infrastructural development, including information and communication technology (ICT) tools, libraries, and multipurpose academic spaces.

## Feasibility and Implementation Considerations

While these recommendations are theoretically robust, their implementation faces practical challenges. Notably, the enhancement of infrastructure, including ICT resources and library facilities, demands significant financial investment. For BOU, which operates under public funding constraints, such expenditures must be aligned with national budgetary allocations and external donor support. According to the tutors, institutional prioritisation of infrastructure should therefore be phased in and needs-based,

focusing initially on tutorial centres with the most acute resource deficits. Additionally, the proposed extension of program duration and restructuring may require regulatory approvals from oversight bodies, such as the University Grants Commission (UGC) and the Ministry of Education, thereby introducing bureaucratic timelines into academic reform cycles.

### Operationalising Tutor Advisory Boards

To actualise participatory curriculum development, tutor advisory boards must be formalised with transparent criteria for tutor selection—such as years of experience, subject-matter expertise, and regional representation. These boards should convene biannually, both virtually and in person, and be anchored within BOU's existing academic governance structures. Their mandate should include reviewing new module drafts, recommending revisions based on tutorial feedback, and participating in consultative curriculum retreats. Institutional buy-in is essential; thus, advisory board outcomes must feed into SoE's Academic Committee with formal channels for policy consideration and adoption.

### Strategic Program Revision

The comparative analysis makes a compelling case for revising the duration and structure of BOU's teacher education programs. A transition toward 2-year models, aligned with global standards, could improve academic depth, international compatibility, and overall program credibility.

### Role Reconfiguration of Academic Coordinators

Coordinators are crucial academic intermediaries. Their role should be institutionally clarified, formalised, and capacitated with adequate decision-making power, budgetary authority, and academic oversight responsibilities to enhance responsiveness across tutorial centres.

## Conclusion and Recommendations

Tutors serve as the principal academic interface between learners and institutions in ODL environments, particularly in professional teacher education programs. Grounded in a constructivist paradigm and guided by grounded theory methodology, this study explored the perspectives of 82 experienced tutors from BOU to identify key strategies for advancing academic development in its education degree programs. The findings reveal eleven interrelated themes requiring institutional attention—most notably learner attendance, curriculum revision, tutorial session structure, physical infrastructure, tutor professional development, and supervision of practice teaching. Together, these themes expose persistent gaps between policy formulation and ground-level realities.

A critical insight emerging from the study is the inherent tension between ODL's emphasis on learner autonomy and the structured, interactive engagement necessary for professional teacher preparation. This necessitates the recalibration of current policies—for instance, by promoting incentivised tutorial attendance rather than mandatory enforcement, thereby balancing flexibility with pedagogical rigour. Likewise, the call for participatory curriculum design underscores the importance of institutional responsiveness. Enabling tutors to co-create instructional materials and engage in curriculum revision processes not only enhances contextual relevance but also strengthens professional identity and pedagogical alignment.

The practical implications of these findings are multidimensional. Institutions should formalise hybrid attendance models, establish tutor advisory boards with clear selection criteria and periodic meetings, and decentralise curriculum review processes to ensure equity and inclusion. The study also highlights the urgency of institutionalising professional development as a structured, ongoing endeavour through regular workshops, certifications, and incentives for advanced study. These measures would enable tutors to remain pedagogically agile and professionally engaged. Moreover, a supervised practicum model—developed in collaboration with local schools and supported by blended modalities such as peer teaching and video reflections—would reinforce experiential learning in teacher preparation. Decentralised infrastructure funding is equally vital to ensure equitable access to ICT resources, libraries, and tutorial centre facilities across geographically diverse contexts.

Even the themes that emerged with lower frequency—such as assessment redesign, co-curricular activities, and examination scheduling—offer early signals of systemic inefficiencies. Addressing these areas proactively may prevent future erosion of academic quality and learner trust in institutional processes.

Given these concerns, structural reform is indispensable. Revising the current 1-year program model into a 2-year structure, in alignment with international benchmarks, could significantly enhance academic depth, learner support, and global competitiveness. The role of academic coordinators, too, must be clarified and capacitated to function effectively as liaisons between institutional goals and field-level implementation.

Theoretically, this study contributes to the broader discourse on learner support and academic governance in ODL by integrating a constructivist and grounded theory-informed approach to tutor narratives. Tutors' reflections illuminate not only operational insights but also the latent power asymmetries and marginalisation within institutional design processes. This necessitates a shift in how tutors are perceived—not merely as peripheral service providers but as legitimate partners in institutional learning and reform. Such a reconceptualisation calls for future research grounded in critical-pragmatic lenses to interrogate the intersection of pedagogy, policy, and professional agency.

Future research should also adopt longitudinal and comparative designs to assess the impact of tutor-informed reforms. Mixed-method studies could further disaggregate insights across tutor categories (e.g., subject specialists, coordinators) and tutorial centres. Cross-institutional comparisons in the South Asian or Global South contexts may offer valuable lessons for building equitable and contextually attuned ODL systems.

In sum, academic development in ODL is a dynamic, context-sensitive process that demands both top-down vision and bottom-up engagement. Tutors, as frontline academic actors, possess experiential knowledge essential for meaningful reform. Recognising and integrating their perspectives into policy discourse is not only a matter of institutional efficiency but also an ethical imperative for inclusive, sustainable educational development.

# References

Allama Iqbal Open University. (n.d.). *B.Ed. programmes*. Retrieved September 16, 2025, from https://www.aiou.edu.pk/bed-programme

Athabasca University. (n.d.). *Master of Education—Open, digital, and distance education*. Retrieved September 16, 2025, from https://www.athabascau.ca/calendar/graduate/fhss/master-of-education-in-open-digital-and-distance-education.html?ss360SearchTerm=Masters%20of%20Education#programregulations

Barrow, M., & Grant, B. (2012). The "truth" of academic development: How did it get to be about "teaching and learning"? *Higher Education Research & Development*, *31*(4), 465–477. https://doi.org/10.1080/07294360.2011.602393

Bascia, N., Carr-Harris, S., Fine-Meyer, R., & Zurzolo, C. (2014). Teachers, curriculum innovation, and policy formation. *Curriculum Inquiry*, *44*(2), 228–248. https://doi.org/10.1111/curi.12044

Becher, T., & Trowler, P. (2001). *Academic tribes and territories*. McGraw-Hill Education (UK).

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research*, *15*(2), 219–234. https://doi.org/10.1177/1468794112468475

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide* (2nd ed.). Sage Publications.

Charles Sturt University. (n.d.-a.). *Bachelor of Educational studies*. Retrieved September 16, 2025, from https://study.csu.edu.au/courses/bachelor-educational-studies

Charles Sturt University. (n.d.-b.). *Master of Education (with specialisations)*. Retrieved September 16, 2025, from https://study.csu.edu.au/courses/master-education

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Sage Publications.

Conole, G. (2013). *Designing for learning in an open world*. Springer.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

de Metz, N., & Bezuidenhout, A. (2018). An importance–competence analysis of the roles and competencies of e-tutors at an open distance learning institution. *Australasian Journal of Educational Technology, 34*(5). https://doi.org/10.14742/ajet.3364

Felber, S. (2020). Adjunct faculty participation in the centralized design of online courses. *Online Journal of Distance Learning Administration*, *23*(1), 1–17. https://ojdla.com/articles/adjunct-faculty-participation-in-the-centralized-design-of-online-courses

Govender, P. (2018). E-tutors' pedagogical practices in a selected open and distance learning university in South Africa. *Progressio: South African Journal for Open and Distance Learning Practice*, *40*(1–2). https://doi.org/10.25159/0256-8853/4706

Hofer, S. I., Nistor, N., & Scheibenzuber, L. (2021). Online teaching and learning in higher education: Lessons learned in crisis situations. *Computers in Human Behavior,* 121, 106789. https://doi.org/10.1016/j.chb.2021.106789

Indira Gandhi National Open University. (n.d.). *Admission: Information bulletin*. Retrieved September 16, 2025, from https://ignouadmission.samarth.edu.in/index.php/site/programme-detail?id=e56b6f82533f742f53e5a6079086db7627c3f868cf5d59ee01b4a314dd7ceebc1224

Indira Gandhi National Open University. (2023). *Programme guide: Bachelor of Education (B.Ed.)*. https://www.ignou.ac.in/viewFile/SOE/programmeguide/BED-2023.pdf

Jegathesan, R., Noryati, A., Amar Hisham, J., & Wan Nordiana, W. H. (2018). Learners' satisfaction and academic performance in open and distance learning (ODL) universities in Malaysia. *Global Business and Management Research: An International Journal*, *10*(3), 511–523. http://ur.aeu.edu.my/id/eprint/459

Johnson, J. A. (2001). Principles of effective change: Curriculum revision that works. *The Journal of Research for Educational Leaders*, *1*(1), 5–18. https://www2.education.uiowa.edu/archives/jrel/fall01/Johnson_0101.htm

Joubert, Y., & Snyman, A. (2020). The contribution of the e-tutor mode in an open distance learning higher education institution: The perspective of the e-tutor. *The Independent Journal of Teaching and Learning, 15*(1), 6–21. https://hdl.handle.net/10520/EJC-1d66798530

Klement, B. J., Paulsen, D. F., & Wineski, L. E. (2017). Implementation and modification of an anatomy-based integrated curriculum. *Anatomical Sciences Education*, *10*(3), 262–275. https://doi.org/10.1002/ase.1676

Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, *22*(4), 557–584. https://doi.org/10.1037/1045-3830.22.4.557

Leibowitz, B., Bozalek, V., van Schalkwyk, S., & Winberg, C. (2015). Institutional context matters: The professional development of academics as teachers in South African higher education. *Higher Education*, *69*, 315–330. https://doi.org/10.1037/1045-3830.22.4.557

Lentell, H. (2001, October). The importance of the tutor in open and distance learning (ODL). In R. Mills & A. Tait (Eds.), *Supporting the student in open and distance learning: Collated conference papers. Proceedings of the 9th Cambridge International Conference on Open and Distance Learning*. (pp. 110–118). The Open University.

Lentell, H., & O'Rourke, J. (2004). Tutoring large numbers: An unmet challenge. *The International Review of Research in Open and Distributed Learning*, *5*(1). https://doi.org/10.19173/irrodl.v5i1.171

Li, S., Zhang, J., Yu, C., & Chen, L. (2017). Rethinking distance tutoring in e-learning environments: A study of the priority of roles and competencies of open university tutors in China. *The International Review of Research in Open and Distributed Learning, 18*(2), 189–212. https://doi.org/10.19173/irrodl.v18i2.2752

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

McAndrew, P., Scanlon, E., & Clow, D. (2010). An open future for higher education. *Educause Quarterly*, *33*(1). https://er.educause.edu/articles/2010/3/an-open-future-for-higher-education

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage Publications.

Mokoena, S. (2017). Student teachers' experiences of teaching practice at open and distance learning institution in South Africa. *Turkish Online Journal of Distance Education*, *18*(2), 122–133. https://doi.org/10.17718/tojde.306564

Mori, Y., Harland, T., & Wald, N. (2022). Academic developers' professional identity: A thematic review of the literature. *International Journal for Academic Development, 27(4)*, 358–371. https://doi.org/10.1080/1360144X.2021.2015690

Muazam, S., Majeed, F., & Khattak, S. M. (2021). Benefits of near-peer tutoring in small group discussions: A pilot study. *Journal of Islamabad Medical & Dental College*, *10*(1), 23–30. https://doi.org/10.35787/jimdc.v10i1.577

Nkolo, T. (2021). *Impact of student support services on academic performance of Bachelor of Education students in Open and Distance Learning (ODL) mode: The case of Botswana Open University* [Unpublished doctoral dissertation]. Botho University.

Ntuli, C. H. S., & Gumbo, M. T. (2019). Tutors' views on the integrated tutor model in open distance learning. *Perspectives in Education*, *37*(2), 53–66. https://hdl.handle.net/10520/EJC-1d90346e26

Olaniran, S. O., Duma, M. A. N., & Nzima, D. R. (2017). Assessing the utilization level of e-learning resources among ODL based pre-service teacher trainees. *Electronic Journal of e-Learning*, *15*(5), 385–395. https://academic-publishing.org/index.php/ejel/article/view/1846

Olivier, B. (2016). The impact of contact sessions and discussion forums on the academic performance of open distance learning students. *The International Review of Research in Open and Distributed Learning, 17*(6), 75–88. https://doi.org/10.19173/irrodl.v17i6.2493

Open University Malaysia. (n.d.). *Master of Education*. Retrieved September 16, 2025, from https://www.oum.edu.my/programmes-offered-at-oum/postgraduate/master-of-education/

Open University of Tanzania. (n.d.-a.). *Bachelor of Education (Adult and distance learning)*. Retrieved February 11, 2025, from https://www.out.ac.tz/ba-eduaddlearning/

Open University of Tanzania. (n.d.-b.). *Masters of education in open distance learning (M. Ed. ODL)*. Retrieved February 11, 2025, from https://www.out.ac.tz/medodl/

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online university teaching during and after the COVID-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education, 2*(3), 923–945. https://doi.org/10.1007/s42438-020-00155-y

Sahoo, P. K., & Chandra, S. (2014). A study of the relationship between students' learning styles and instructional inputs in a teacher education programme of IGNOU. *Asian Association of Open Universities Journal*, *9*(1), 17–34. https://doi.org/10.1108/AAOUJ-09-01-2014-B003

Sánchez-Elvira Paniagua, Á., & Simpson, O. (2018). Developing student support for open and distance learning: The EMPOWER project. *Journal of Interactive Media in Education, 2018*(1), Article 9. https://doi.org/10.5334/jime.470

Segoe, B. (2014). Tutor support in an open distance learning (ODL) environment for upgrading teachers. *Journal of Human Ecology*, *48*(1), 161–169. http://krepublishers.com/02-Journals/JHE/JHE-48-0-000-14-Web/JHE-48-1-000-14-Abst-PDF/JHE-48-1-161-14-2682-Segoe-B-A/JHE-48-1-161-14-2682-Segoe-B-A-Tx[17].pdf

Sembiring, M. G. (2017). Exploratory study of academic excellence associated with persistence in ODL setting. *Asian Association of Open Universities Journal*, *12*(2), 125–136. https://doi.org/10.1108/AAOUJ-01-2017-0015

Sembiring, M. G. (2020). Academic excellence as a determinant of self-confidence among graduates of ODL programs. *Asian Association of Open Universities Journal*, *15*(3), 411–423. https://doi.org/10.1108/AAOUJ-09-2020-0068

Shelley, M., White, C., Baumann, U., & Murphy, L. (2006). "It's a unique role!" Perspectives on tutor attributes and expertise in distance language teaching. *The International Review of Research in Open and Distributed Learning*, *7*(2). https://doi.org/10.19173/irrodl.v7i2.297

Sutherland, K. A. (2018). Holistic academic development: Is it time to think more broadly about the academic development project? *International Journal for Academic Development*, *23*(4), 261–273. https://doi.org/10.1080/1360144X.2018.1524571

Tait, A. (2018). Education for development: From distance to open education. *Journal of Learning for Development, 5*(2), 101–115. https://doi.org/10.56059/jl4d.v5i2.294

Tait, A. (2000). Planning student support for open and distance learning. *Open Learning*, *15*(3), 287–299. https://doi.org/10.1080/713688410

The Open University. (n.d.-a.). *BA (Honours) Education Studies (Primary)*. Retrieved September 16, 2025, from https://www.open.ac.uk/courses/education/degrees/ba-education-studies-primary-q94

The Open University. (n.d.-b.). *Masters degree in education*. Retrieved September 16, 2025, from https://www.open.ac.uk/postgraduate/qualifications/f70

UNESCO & International Task Force on Teachers for Education 2030. (2024). Valuing teacher voices: Towards a new social contract for education. UNESCO. https://teachertaskforce.org/sites/default/files/2024-10/1530.24_WTD2024_final_0.pdf

van Zyl, J. M., Spamer, E. J., & Els, C. J. (2012). Effect of contact class attendance on the academic success of open distance learning students in Advanced Certificate in Education programs. *Contemporary Educational Technology, 3*(3), 166–183. https://doi.org/10.30935/cedtech/6076

Walker, R. (2015). Peer observation for online distance learning tutors: Creating the conditions for effective peer exchange. *European Journal of Open, Distance and E-Learning*, *18*(1), 34–51. https://doi.org/10.1515/eurodl-2015-0003

Woo, T. K. (2011). Developing quality learning materials for effective teaching and learning in an ODL environment: Making the jump from print modules to online modules. *Asian Association of Open Universities Journal*, *6*(1), 51–58. https://doi.org/10.1108/AAOUJ-06-01-2011-B006

Xiao, J. (2016). Who am I as a distance tutor? An investigation of distance tutors' professional identity in China. *Distance Education*, *37*(1), 4–21. https://doi.org/10.1080/01587919.2016.1158772


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Online Learning and Opera: The Rise in Digital Vocal Master Classes and Workshops

**Weiyi Li**
*College of Chinese & Asean Arts, School of Music and Dance, Chengdu University, Cheng Du, China*

## Abstract

Professionalism in opera singing depends on the quality of education. This study examines the teaching of opera singing online via master classes and seminars. Based on the present findings, master classes are effective in improving the voice technically and range-wise, whereas seminars are effective in helping opera singers achieve emotionality. These instructional methods were compared in terms of content variety, balance between theoretical and practical components, and their effect on performance artistry and technical skills. Those learning to sing opera through master classes reported experiencing improvements in memory (24%) and problem-solving (22%). Students who attended the online seminars reported improvements in concentration (28%) and memory (26%). Challenges include limited control over assignments, attendance, and assessment. This study highlights effective tools for online opera training, with potential for future comparisons of methods in developing vocal and instrumental skills.

*Keywords*: cognitive abilities, content variety, distance education, expression, vocal range

# Introduction

To produce a stable and beautiful operatic sound, vocalists must possess a deep understanding of the piece they are reproducing. Opera's aesthetic revolves around emotional expressivity, which demands a certain level of skill from the vocalist (Aaberg, 2023). Given the growing variety of digital learning technologies, it becomes increasingly important to identify which solutions are most effective in opera skills development.

In China, opera is of particular importance as a tool for preserving vocal and theatrical traditions. Learning the art of operatic singing implies mastering the vocal technique and adapting one's voice to different pieces of music. Vocal artists should also know how the music they reproduce was created and under what historical circumstances it came to life (Almqvist & Werner, 2024; Silas et al., 2024). A common approach to teaching opera singers in China involves using traditional music to explore the distinctive features of the national culture. The integration of artistic elements into opera singing teaching practice is typical of the Chinese education, but other educational systems in the world also apply this approach to teaching, as it allows for the realization of a theatrical performance (Joseph & Lennox, 2021; Y. Li, 2024). When educating operatic artists, the emphasis should be placed on phonetics, breathing, and singing technique (de Villiers & Gillmer, 2024). With Chinese opera, special attention is paid to pronunciation, which also has its own intonation that is different from the existing melodic line (Wang, 2024b). Because Chinese opera combines singing, dramatic acting, acrobatics, and martial arts, the teaching process also focuses on unveiling the vocalist's individuality and potential (Chang, 2022).

The integration of online platforms into opera singing education is not yet a standardized practice; however, it offers significant potential for implementing distance learning aimed at enhancing professional competencies (Crisosto-Alarcón et al., 2025). The advantage of distance learning in the context of operatic training lies in its capacity to provide an individualized approach to information processing and to engage instructors from various countries. This facilitates the development of professional skills through exposure to diverse pedagogical approaches and performance techniques of vocal coaches and opera singers. Although distance learning predominantly emphasizes independent study, it allows for the repeated review of educational materials and critical analysis of vocal performance, thereby contributing to higher levels of professional mastery (Liming & Sartjinpong, 2025). Online master classes and workshops may focus on a single skill (e.g., breath control or articulation) that supports the in-depth remediation of specific gaps in operatic performance (Y. Li, 2024).

Even though operatic training requires careful preparation, it can be implemented using a distance approach. An advantage of online learning is that it allows students to learn the art of singing regardless of their location (Nikolai et al., 2023). In these conditions, teacher–student communication depends on the level of technological organization. The choice of digital technology affects the ability to conduct high-quality seminars and master classes, the format of which makes it difficult to introduce musical accompaniment for students to use while learning (Chen, 2023; Longlong & Luen, 2023). Distance education is associated with the use of video to deliver content; hence, an interactive system that will transmit high-quality sound must be selected (Cuervo et al., 2023). Surely, some digital devices allow for singing with accompaniment and even enable vocalists to tailor the accompaniment to their skill level (Shu, 2023). Today's students resort to online master classes and

seminars because they break up the monotony of learning, improving engagement (P.-P. Li & Wang, 2023).

The academic relevance of digital vocal master classes and seminars lies in their capacity to support the professional development of musicians. This is largely attributable to the abundance and accessibility of educational materials available in the digital space (Fan et al., 2025). Master classes and seminars often adopt nontraditional formats for delivering content that tend to engage students and foster motivation for participation. These educational formats promote the continuous expansion of musical horizons by introducing updated approaches to vocal performance (Ilxomovna, 2025). They also contribute to a broader understanding of practical techniques for melodic interpretation across various musical styles. Through such instructional methods, students are encouraged to cultivate independence beyond the confines of conventional academic curricula (Yuhan & Sartjinpong, 2025).

Challenges in exploring the specifics of operatic training stem from the lack of a clearly defined connection between the potential of vocal master classes and seminars and their practical implementation. Although these formats hold considerable promise and align with emerging models of vocal education, more attention must be paid to the current opportunities for online opera training. In particular, focus should be placed on vocal projection and the acoustic refinement of timbre—elements that are often overlooked in the design of distance learning courses. One of the central difficulties of remote instruction is the absence of a coherent strategy for monitoring task completion, which can negatively impact the quality of vocal training. Additionally, distance learning presents limitations in assessing a singer's technical execution, especially in areas such as body engagement and the establishment of proper breath support. These limitations may be addressed through the selection of targeted instructional methodologies designed to meet specific objectives for developing an operatic voice.

## Literature Review

The initial focus of the theoretical inquiry was to identify the factors that contribute to the enhancement of operatic singing and to explore the potential of achieving these improvements through the use of digital tools. The practical application of digital technologies in vocal training can be realized through the selection of specific repertoire, the use of automated analysis tools, and the progressive performance of compositions with a focus on individual musical fragments. The online curriculum must be theoretically justified and incorporate instructions prepared by competent teachers. The development of vocal skills is facilitated through the continuous refinement of practical abilities. An understanding of the technical and physiological aspects of singing shapes the accumulation of performance experience. This integrated approach enables students to internalize the nuances of each musical piece, thereby enhancing the overall quality of their performance (Jansson & Balsnes, 2021). In synchronous online education, teachers must first deal with any technical problems that may arise when establishing a dialogue with their students and then address the cultural dimension of the curriculum to improve musical cultural literacy and attract a geographically distinct audience. Instruction should be oriented toward the integration of intercultural dimensions, particularly through the selection of musical repertoire and the inclusion of diverse genres. This approach enables the development not only of the technical aspects of operatic singing but also of expressive performance, which is often shaped by the traditional and cultural characteristics inherent in individual compositions (Løkke Jakobsen et al., 2023). Opera teaching specifically begins with the

establishment of a theoretical foundation (an understanding of the principles of breath support, singing within specific vocal registers, the execution of musical ornamentation, and the coordination between breath control and sound production techniques). When training opera singers, it is important to eliminate students' dependence on subjective perception and to cultivate performance skills that are based on multidimensionality. This can be achieved by using existing online courses as a foundational framework for instruction (Luo & Leung, 2022).

Further analysis was aimed at determining the format of remote learning of opera singing and possible advantages. Digital technologies make learning more flexible and imaginative, serving as effective tools for developing vocal skills (Biasutti et al., 2022). Digital technologies enhance the flexibility and creativity of the learning process, enabling more effective planning of the educational curriculum and facilitating the identification of tools for developing vocal skills. The online format allows for the adoption of the Feldenkrais method (i.e., learning through movement), which was reported to increase awareness of singing and enhance practice. The Feldenkrais Method is associated with the use of hand, body, and head movements during vocal performance, which contribute to achieving greater expressiveness. The process of self-discovery provides individualized benefits for vocal performance (Paparo, 2022). The traditional system of education transformed during the COVID-19 pandemic, with many learners moving online. Music educators shifted to using online applications (apps) to teach music theory and instruments. Muyu is but one example of such an app; it promotes collaborative learning and student engagement through prerecorded videos, effectively motivating users to improve their music knowledge. The practical use of the app is associated with the facilitation of collaborative learning and the engagement of students using prerecorded video materials (Ng et al., 2022). Encouraging students to write songs can help them understand the various components of music and explore the depth of their creativity. This approach allows for the development of improvisation techniques (Wang, 2024a).

The next group of scientific works is aimed at evaluating the advantages of online learning for developing students' cognitive abilities. Online courses that provide opportunities to complete a large number of assignments have the potential to carry out the same function that educators have, encouraging learners to nurture their own thinking skills. The practical component of the learning process can be implemented through students' completion of a substantial number of online tasks. Increasing students' self-efficacy contributes to the formation of musical habits (Vaizman & Harpaz, 2023). Emergency remote learning encompasses the process of interaction between instructional teams and is aimed at automating the educational process. This is facilitated by interactive music applications that enable the analysis of students' singing and provide corresponding feedback and recommendations. This approach enables learners to focus on individual elements of a vocal performance that make it exude grace and subtlety, be it vocal color or delivery (Rucsanda et al., 2021). Distance learning can also facilitate personal growth: online vocalists have been reported to gain greater confidence when training independently (Smith et al., 2022).

The literature analysis also focused on potential challenges arising from the implementation of distance learning formats in operatic singing instruction. Vocal training in an online environment may present specific difficulties, often associated with insufficient digital preparedness and the lack of effective pedagogical strategies. In particular, the most pressing issues are linked to suboptimal sound quality, which hinders the perception of nuanced vocal details. Technical problems with the software employed can negatively affect student motivation and limit the possibility of synchronous musical performance (Ngobeni, 2024). Furthermore, the challenges of remote music education may stem from

a shortage of educational resources and specialized music software, which complicates classroom management and diminishes the quality of information assimilation (Wang, 2025).

The review of published studies revealed a predominant focus on the general characteristics of distance learning with a lack of detailed rationale for the implementation of specialized online seminars and master classes. The application of online approaches to opera singing instruction has been addressed only superficially, with limited attention to the practical techniques involved, and it has been primarily oriented toward teaching theoretical content. Existing research tends to emphasize general music education through remote modalities rather than the specifics of operatic training. Enhancing the practical relevance of such studies requires a substantiated exploration of methods for achieving high-quality operatic performance through online seminars and vocal training courses.

# Problem Statement

The online learning format, favored by modern vocal performers, is transforming the conventional method of vocal training. The growing popularity of online singing master classes and seminars is associated with the use of atypical approaches to teaching and learning that allow vocalists to discover and showcase their individuality. When it comes to opera music, the uniqueness of one's performance is important, for the personal interpretation of the piece by a vocalist distinguishes one composition from another, and it allows vocalists to represent a certain culture. In operas, the plot is sometimes based on adequately presented historical events, requiring the vocalist to possess skills at the higher end of the spectrum of expertise. Hence, online singing courses must be constructed with special care, for opera music is filled with various artistic flourishes that add to the expressiveness of the vocal performance, and vocalists have to study all or nearly all of them to be able to deliver competitive, high-quality performances. This study aims to investigate the effectiveness of online singing master classes and seminars. The objectives of the study were (a) to identify specific factors that contribute to a high-quality operatic performance and whether online master classes and seminars can help in developing thereof, (b) to compare the strengths of specific master classes (Juilliard Extension, the Royal College of Music) and online seminars (Living Opera, Angel's Music Academy) designed for opera singers, (c) to measure the cognitive abilities of individuals exposed to online master classes and seminars, and (d) to identify potential problems that students may face when learning how to sing opera via online master classes and seminars.

# Materials and Methods

## Study Design

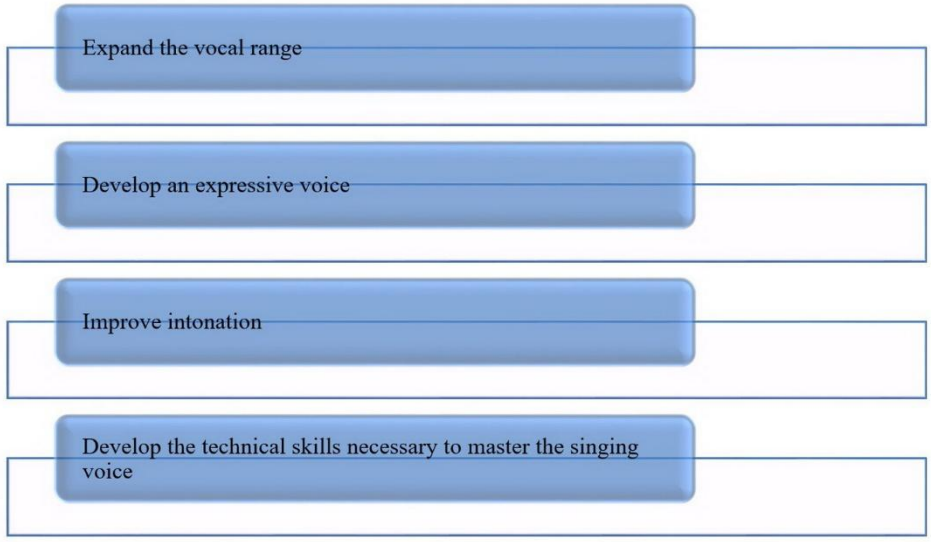
The study consisted of three stages: (a) identifying which direction to take in developing opera singing skills, (b) comparing different master classes and seminars for online opera singers in terms of effectiveness, and (3) carrying out an 8-week intervention and examining its effect on the cognitive skills of novice opera singers. The first stage entailed analyzing a total of 200 opera performances to establish which aspects of opera singing must be covered in online training. The analysis of operatic performances involved an examination of both the acoustic and technical features of singing, as well as stylistic characteristics. Particular attention was given to evaluating the performative and theatrical delivery techniques employed by professional vocalists during live presentations. This approach focused on assessing the consistency and flexibility of vocal production, exploring subtle nuances and intonation, and investigating the relationship between breath control techniques and pitch accuracy.

Such a methodology enabled a more precise understanding of operatic forms and the expressive emphases used by performers to achieve individualized and emotive interpretations.

Special emphasis during the analysis was placed on the use of unconventional vocal techniques that contributed to shifts in the stylistic rendering of musical compositions. The identification of common evaluative criteria guided the selection of the most influential features, which were subsequently incorporated into the study. The analysis of operatic performances included both video evaluation and spectrographic analysis conducted using Praat software, version 6.4.31 (Boersma & Weenink, 2025). Praat enabled a detailed acoustic analysis, including frequency distribution of melodic lines, assessment of resonance characteristics, and examination of dynamic and intensity parameters in vocal execution. In addition, we analyzed the content of more than 100 master classes and seminars, as well as student feedback. Master classes selected for the analysis are available in the 2024 catalogs offered by Juilliard Extension and the Royal College of Music, whereas the chosen seminars are part of the 2024 series of online courses released by Living Opera and Angel's Music Academy. All materials are publicly available. The analysis encompassed an evaluation of thematic content, its comprehensiveness, alignment with university curricula in operatic vocal training, the potential for offering individualized instructional approaches, and the balance between theoretical and practical components within the course structure. The content analysis revealed four elements that online master classes and seminars for novice opera singers should focus on: (a) expanding the vocal range, (b) developing an expressive voice, (c) improving intonation, and (d) developing the technical skills necessary to master the singing voice (see Figure 1).

**Figure 1**

*Four principles for developing opera singing skills*




Source: Developed by the authors

The core principles behind an excellent vocal performance in opera singing were obtained using the Mann–Whitney *U* test and Cliff's delta (δ). The *U* statistic reflects the relationship between two independent samples. The test requires that values from the two populations do not correlate with each other to ensure that no patterns will occur when a random variable increases or decreases

(Almqvist & Werner, 2024). The Mann–Whitney *U* test does not assume a normal distribution; therefore, the emphasis in the analysis was placed on the ordinal values of the selected criteria. This nonparametric coefficient enabled the identification of the relative prominence of certain criteria over others through a rank-based comparison of master classes and online seminars. The calculations were conducted at a significance level of $\alpha = 0.05$. Cliff's delta is a nonparametric effect size estimate that assesses the difference between two observations without using the standard deviation (Zeng, 2023). The calculation of Cliff's delta primarily involves the assessment of sample size, which reflects the effect size of specific criteria in relation to the development of operatic singing skills, depending on the selected course or master class. Cliff's delta represents a nonparametric measure of effect size, accounting for asymmetric distributions associated with the distinctive characteristics of various instructional methods. A value of 0 indicates no difference between groups; a value of 1 signifies that the influence of the first type of instruction dominates over the second, while a value of –1 indicates the opposite—that the second type of instruction has a stronger influence than the first.

Overall, the significance of vocal criteria for high-quality operatic performance was assessed through an analytical approach that involved the examination of operatic works to identify relevant evaluation criteria, as well as the study of master classes and online seminars aimed at implementing these criteria. The analysis of operatic performances was based on video recordings and live broadcasts, which allowed for an assessment not only of vocal quality but also of performers' stage behavior. The selection of evaluation criteria was guided by impact on intonational accuracy and expressive delivery—elements that enhance audience engagement.

As previously outlined, the analysis of master classes and online seminars focused on the thematic content and pedagogical strategies employed. The primary assessment was carried out by the authors of the study; however, in examining the specific features of the instructional formats, student feedback from those who had completed the training was also taken into account. To this end, over 400 reviews were collected, with particular attention given to well-reasoned responses containing concrete examples and direct feedback from the organizers of the master classes and online seminars.

A linear rating scale ranging from 1 (negative) to 5 (positive) was employed. Evaluations were assigned by the authors based on three criteria: the importance of a given vocal parameter for operatic development, the extent to which it can be cultivated through master classes and online seminars, and the content of student feedback. The resulting data were aggregated to obtain mean values and subjected to Mann–Whitney *U* tests and Cliff's $\delta$ calculations. To minimize bias, measures were taken to ensure objectivity, including the use of standardized evaluation criteria (e.g., vocal range development, expressive sensitivity, vocal intonation, and technical vocal control) and a rigorous analysis of the instructional content. Each author independently rated the data based on comparative review analysis, enabling cross-validation of the results and ensuring their reliability.

The second stage entailed analyzing the selected seminars and master classes with respect to the four aspects of opera singing established in the first step. Methods for developing the skills needed for an excellent operatic performance were chosen from the traditional repertoire and were used by online instructors throughout the selected classes. The selection of criteria for evaluating effectiveness was also informed by the elements used in sampling online seminars and master classes for inclusion in the study. The chosen effectiveness indicators encompassed both pedagogical aspects (such as lesson variety and the balance between theoretical and practical content) and vocal dimensions (including artistic expression and vocal technique). These indicators were selected to assess the impact of master

classes and online seminars on the development of professional competencies. The effect of an online master class or seminar was quantified using the following formula (Equation 1) (Lowery, 2023):

$$V = \frac{s}{x}. \quad (1)$$

where $s$ represents the standard deviation of the variable from its accepted norm, and $x$ is conditional significance.

The coefficient of variation is employed to compare the ratio of a specific indicator to a predefined reference value. In this context, the standard deviation of a given parameter from the mean value was taken into account. The application of the coefficient of variation enabled an assessment of the degree of alignment between the analyzed training program and the effectiveness criteria, with a focus on the prominence of specific indicators within the study.

For a more compelling argument, the Mann–Whitney *U* statistic was additionally calculated. The Mann–Whitney *U* test was conducted to compare overall program effectiveness (as determined by the mean effectiveness score) with each individual quality criterion of the master classes and online seminars, including lesson diversity, content richness (theoretical and practical materials), potential for achieving artistic expressiveness, and the development of technical skills. Table 2 presents the average Mann–Whitney *U* values, calculated with measures taken to minimize potential errors.

Independent experts carried out the evaluation of master class and online seminar effectiveness—15 vocal pedagogy instructors—who had access to the materials and directly observed participant sessions. These experts were recruited from prominent Chinese institutions (Central Conservatory of Music, Wuhan Conservatory of Music, Sichuan Conservatory of Music) as well as European institutions (Universität Mozarteum Salzburg, Royal Academy of Music, Hochschule für Musik, und Theater München). Importantly, the invited experts had no involvement in the design or delivery of the evaluated master classes and seminars, which helped eliminate potential conflicts of interest or bias toward any particular instructional program.

A standardized 5-point Likert scale (ranging from 1 to 5) was used to assess the effectiveness of each program according to clearly defined criteria: diversity of lessons, richness of theoretical and practical content, potential for achieving artistic expressiveness, and development of technical skills. These standardized criteria were readily identifiable based on the program structure, thus minimizing the risk of subjective interpretation. Communication with the independent experts was conducted via WeChat, version 8.0.44 (Tencent Holdings Limited, 2023).

The 8-week intervention involved exposing four groups of learners to the selected seminars and master classes. Learners in groups 1 and 2 participated in Juilliard Extension's and Royal College of Music's online classes, whereas learners in groups 3 and 4 took online seminars from Living Opera and Angel's Music Academy, respectively. The chosen materials follow the underlying principles of excellent operatic performance outlined in Figure 1 and help learners meet the vocal requirements to perform opera music. After the intervention ended, the impact of online instructional methods on the acquisition of cognitive skills was assessed. Data for master class groups and online seminar groups were obtained separately. Effectiveness was assessed based on the respondents' development of cognitive skills aimed at achieving high-quality performance through enhanced neural activity. The evaluation of cognitive skills focused on their influence on the comprehension of operatic performance

processes, depending on the instructional approach selected by the students. All calculations were done using the Mann–Whitney *U* test and Cliff's δ.

Thus, the assessment of students' cognitive abilities was conducted through observational methods by independent experts who had previously participated in the second stage of the study. The evaluation criteria included students' approaches to task interpretation, attentional stability, response patterns to assignments, and execution strategies. In addition, standardized indicators such as information processing speed, memory retention, and self-regulation were taken into account. The level of cognitive skill development was assessed across all respondents who received instruction through master classes and seminars. This assessment involved providing a general evaluation to students who were taught via master classes and those who engaged in online seminar-based learning, without further subgroup differentiation.

In addition, participants were asked about the challenges they had encountered during the intervention period. For this, an open survey was administered, in which respondents were asked to name and describe a specific problem related to learning via online master classes and seminars. Students were required to either agree with the stated issue or refute it. The use of Cronbach's alpha coefficient enabled the assessment of the internal consistency of all student responses, aimed at eliminating internal contradictions. The resulting Cronbach's alpha value of 0.78 confirmed the validity and reliability of the collected responses. Responses were collected via WeChat within 1 day.

## Participants

A total of 216 respondents with vocal singing skills but no experience with opera participated in the intervention. The study population included instrumental musicians, singers specializing in genres other than opera, teachers, and choreographers recruited via ads on Douyin (i.e., the Chinese short video app). Individuals responding to the recruitment efforts were asked to sing different pieces of vocal music from the repertoire proposed by the researchers to determine their level of vocal music skill. Those eligible were divided into four groups of 54 participants; the groups were then randomly assigned to receive one of the four online instructions. The age of the respondents ranged from 19 to 24 years.

## Statistical Processing

All calculations were done in Microsoft Excel. This program allowed us to maintain the sequence of the selected criteria while splitting the collected data into groups.

## Ethical Considerations

This study follows the ethical guidelines prepared by the Norwegian National Committee for Research Ethics in Science and Technology (2024). The content of this paper is original and has not been published previously. All information taken from other sources as part of the theoretical framework has been inserted with proper citation.

## Limitations

Opera singing is the sole focus of the current study. Future research should include other genres of vocal music to assess the effectiveness of online seminars and master classes within a broader area of application.

# Results

Compared to other vocal genres, opera is a complex yet exceptionally beautiful art form that is both technically and physically demanding. The content analysis of multiple master classes and seminars revealed that when teaching novice opera singers online, it is best to focus on the following (Table 1): a person's vocal range, expressivity of the vocal performance, accuracy of intonation, and technical capabilities of the voice. The calculations were conducted using the Mann–Whitney *U* test and Cliff's δ.

**Table 1**

*The benefits of using master classes and online seminars for opera skills development*

| Effects | Master classes | | | Online seminars | | | Difference between instructional methods | | |
|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *Mdn* | *M* | *SD* | *Mdn* | *U* | δ | *p* |
| Vocal range expansion | 8.71 | 2.43 | 9 | 6.15 | 1.20 | 6 | 167 | 0.316 | 0.003[a] |
| Emotional expressivity | 7.20 | 1.38 | 8 | 7.56 | 1.50 | 8 | 132 | 0.221 | 0.211 |
| Intonation development | 7.59 | 1.82 | 8 | 7.34 | 1.46 | 8 | 141 | 0.203 | 0.213 |
| Technical mastery of the voice | 8.76 | 2.47 | 9 | 4.0 | 0.98 | 4 | 185 | 0.319 | 0.002[a] |

[a] Represents values that do not correlate at $p > 0.05$.

A wide vocal range is important for opera singers since the ability to comfortably produce a broad span of notes without straining one's voice affects the quality of the operatic sound. When expanding their vocal range, singers should focus on maintaining the purity and beauty of the reproduced sound. The challenge is that opera pieces tend to weave high and low notes together, and vocalists are expected to seamlessly move between them. Hence, the vocal routine should additionally enhance the voice's strength and timbre. This can be achieved through vibrato singing exercises.

An expressive voice is a way to express the emotional depth of opera music. Expressive singing is the ability to reproduce vocally different emotive qualities of the vocal music piece or set the emotional tone of the passage by manipulating one's tone of voice. A touching, captivating performance requires the inclusion of appropriate melodic embellishments and ornamentation.

Intonation in opera singing is the ability to use the rising and falling vocal pitch and sound combinations to convey the personality and intentions of the character or to emphasize the importance of the scene. Accurate intonation contributes to the purity of the reproduced sound and proper coordination. Only intonation allows singers to convey the essence of the operatic work.

Opera repertoire demands technical mastery. To build a solid technical foundation, vocalists should adopt various techniques and vocal exercises that will improve their breath control and articulation.

A comparison of distance master classes and online seminars based on the authors' analysis of their content, topics used, content of classes, and analysis of feedback from students who attended these training formats showed that online master classes allowed learners to greatly improve their vocal potential and technique through demonstrations. Online seminars, on the other hand, focused more on explaining theory. The present study assessed the potential of online seminars and master classes

to cultivate opera singing skills. Numerical values were obtained using the coefficient of variation. The results are presented in Table 2.

**Table 2**

*Effectiveness of online master classes versus online seminars for novice opera singers*

| Programs | Content variety | Theory / practice balance | Performance artistry | Technical skills | Mean effectiveness | Mann–Whitney *U* | *p* |
|---|---|---|---|---|---|---|---|
| Master classes | | | | | | | |
| Juilliard Extension | 0.92 | 0.93 | 0.91 | 0.92 | 0.92 | 131 | 0.127 |
| Royal College of Music | 0.87 | 0.89 | 0.91 | 0.89 | 0.89 | 134 | 0.135 |
| Seminars | | | | | | | |
| Living Opera | 0.83 | 0.74 | 0.95 | 0.79 | 0.82 | 120 | 0.936 |
| Angel's Music Academy | 0.89 | 0.88 | 0.93 | 0.87 | 0.89 | 132 | 0.129 |

Based on the responses of independent experts, it was established that Living Opera offered seminars that focused more on the artistry of an operatic performance rather than a singer's technical arsenal. The proposed series of eight lectures prepared listeners for participation in an opera performance. The training process involved learning how to overcome the fear of performing in front of other people and achieve psychological stability, choosing one's repertoire, developing a unique interpretation of operatic pieces, and learning how to conduct auditions. Experts established that the lectures were oversaturated with theoretical knowledge that did not allow learners to develop performance-related skills and that the lessons were largely identical in terms of content.

The expert ratings used for the coefficient of variation showed Angel's Music Academy seminars were geared toward helping singers work on their technique. More specifically, the proposed video course helped students hone both the technical and artistic skills of vocal performance while informing them about the distinct ways of aesthetic expression in the operatic stage performance. In addition, the course allowed those who took it to focus on studying the breadth of the vocal score and the specifics of their own timbre. The way in which the content was presented contributed to the development of critical thinking.

Juilliard Extension's master classes were all-encompassing and well-balanced experiences; not only did they help learners work on their technique and ear, but the combination of theory and practice led to a better understanding of different singing techniques. Diversity in the curriculum allowed singers to select and expand their musical repertoire according to their technical and creative capabilities. The same was true for the Royal College of Music's master classes. Led by experienced operatic artists, the master classes taught approaches to opera as a performance and ways of conducting oneself while on stage. The results are presented in Table 3, based on calculations using the Mann–Whitney *U* test and Cliff's δ.

**Table 3**

*Cognitive abilities after exposure to different instructional methods*

| Cognitive ability | Master classes | | | | Seminars | | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | % of respondents | *M* | *SD* | *Mdn* | % of respondents | *M* | *SD* | *Mdn* | *U* | δ | *p* |
| Memory | 24 | 8.65 | 2.40 | 9 | 26 | 8.69 | 2.41 | 9 | 147 | 0.305 | 0.214 |
| Concentration | 18 | 8.22 | 8.13 | 9 | 28 | 8.72 | 2.42 | 9 | 127 | 0.291 | 0.196 |
| Problem-solving | 22 | 8.61 | 2.37 | 9 | 11 | 6.64 | 1.82 | 6 | 183 | 0.372 | 0.004[a] |
| Ability to replicate unfamiliar music after hearing the piece | 17 | 7.91 | 2.01 | 8 | 13 | 7.21 | 1.95 | 7 | 124 | 0.361 | 0.08 |
| Thinking | 19 | 8.56 | 2.31 | 9 | 22 | 8.61 | 2.38 | 9 | 125 | 0.284 | 0.195 |

[a] Represents values that do not correlate at $p > 0.05$.

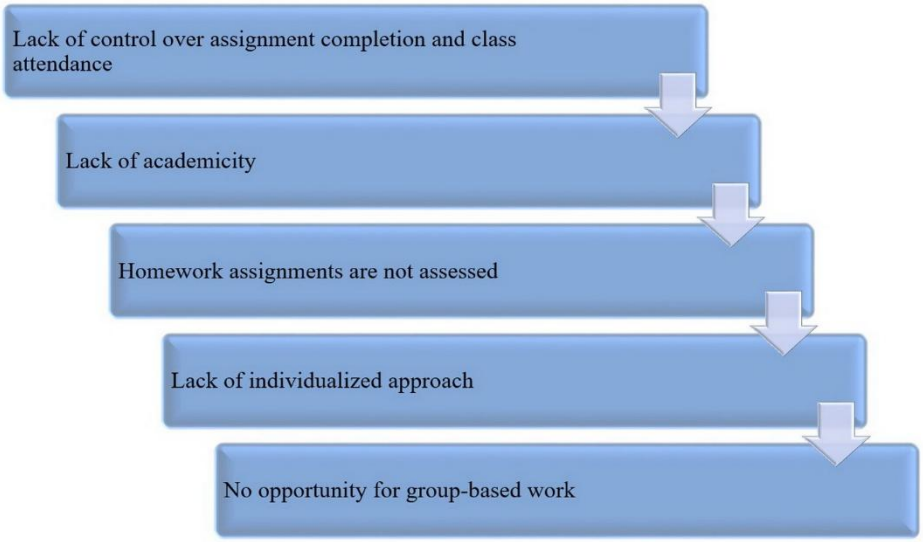
Learning through online seminars and learning through online master classes yielded different results in the cognitive department (the modality effect), as shown in Table 3. The observation of students showed exposure to master classes facilitated the development of memory, problem-solving skills, and thinking. The singers' memory was reinforced as a result of applying the newly acquired knowledge and skills in a practical context. A better memory enabled students to adapt to the operatic repertoire and find solutions to specific problems, such as complex fragments and transitions between pitches, which was related to the students' ability to reproduce a fragment of a musical work, allowing up to two intonation or structural errors. At the same time, students were able to interpret the opera work a day after listening. By developing their thinking skills, participants improved their ability to interpret the operatic piece and increase the quality of the operatic performance. Advanced thinking skills implied students' ability to engage in music-theoretical analysis, including the examination of musical texture. This was also reflected in cognitive flexibility, which enabled the adoption of innovative approaches to musical performance and the interpretation of musical imagery.

In contrast to master classes, seminars placed greater emphasis on theory. Exposure to online seminars fostered the development of concentration, memory, and thinking, with development of concentration being the result of the constant repetition of the learning material. The development of attentional concentration was facilitated through repeated exposure to learning materials, which enabled students to focus on previously studied elements. This improvement was reflected in learners' accurate performance of complex musical passages, ability to shift efficiently between different musical tasks, and reduction in number of errors during the repetition of musical pieces. Teacher support played a crucial role in the development of these cognitive abilities.

Aside from the benefits the given modalities provide, learning via online seminars and master classes comes with its challenges, as shown by an open survey via WeChat (see Figure 2).

**Figure 2**

*Challenges faced by respondents when learning opera singing skills through online seminars and master classes*




Source: Developed by the authors

According to the respondents, the main problems of learning via online master classes and seminars were related to the lack of academicity. To achieve higher results, students had to be independent and work on their technique without anyone's help. The lack of control over assignment completion and class attendance hindered the development of accurate vocal techniques. Such conditions required students to be psychologically prepared to spend time independently figuring out individual components of their singing technique.

## Discussion

Online courses are widely adopted as an alternative means of teaching music. An innovative format of content delivery that integrates a variety of educational resources related to the art of singing and instructions guiding one toward discovering their own performance style has been associated with higher achievement (G. Li, 2024). Singers develop their vocal skills through practice and accumulate experience in a manner that requires an understanding of the technical and physiological dimensions of singing. This approach allows vocalists to "feel" the piece and exploit this auditory and theoretical knowledge to improve the quality of the performance (Jansson et al., 2024). Previous research has explored the possibility of developing vocal skills in a manner that focuses on individual parameters, such as breathing, expressiveness, and performance style. The present study has looked into ways of achieving an expressive performance in opera specifically. The basic recommendation involves building a wide vocal range, working on expressivity and intonation, and improving the technical skills necessary to control the vocal apparatus. Unlike previous studies, this work examined which instructional method (i.e., online master class and online seminar) is best for helping students throughout this journey.

The Internet provides consumers with the proper means to improve their learning while removing the potential limitations that can otherwise hinder their progress. The online courseware is usually

prerecorded and/or presented through online simulation in virtual classrooms. When it comes to music education, the effectiveness of this content delivery method depends on which tool is used for audio processing, as teachers have to accurately transmit the tone and timbre of the vocal music piece for students to understand and correct their mistakes as vocal performers (Rui, 2024). Digital technologies are effective and accessible tools for teaching and learning music that allow for better knowledge assimilation, higher engagement, and the development of creative skills. With their help, teachers can save on musical resources while enhancing the vocal abilities of their students (Rexhepi et al., 2024). Online courses require the use of instructional design tools and assessments; however, the course material embedded in the online course must be tailored to individual student capabilities. Audio content must be preprocessed with MIDI to improve the depth and quality of learning. More specifically, a well-constructed MIDI dataset can help teachers improve students' music literacy, as reported in previous research (Shi, 2024). Previous research has explored the effectiveness of innovative technologies in music teaching. The findings suggest combining theory and practice as a means of achieving better results. This study considered the possibility of using certain master classes (Juilliard Extension: https://www.juilliard.edu/stage-beyond/juilliard-extension; Royal College of Music: https://www.rcm.ac.uk/) and online seminars (Living Opera: https://www.livingopera.org/for-singers; Angel's Music Academy: https://angelsmusicacademy.com/) to develop opera singing skills. The effectiveness of each was tested with respect to content variety, balance, and effect on performance artistry and technical skills.

For the beauty of the singing voice to be unveiled, the teaching–learning process requires clear contextualization (Lim et al., 2024), and if the learning takes place online, the instructor should address the dimension of nonverbal communication to help learners understand the subtleties of stagecraft. Role-playing exercises, for example, are good for increasing students' confidence, leading to a greater sense of autonomy and a higher level of professionalism (Urbaniak & Mitchell, 2024). In published studies, analyses have primarily focused on the development of vocal skills related to the refinement of vocal performance. In contrast, this article emphasizes the development of cognitive skills associated with intentional operatic interpretation and the potential for achieving greater expressive nuance in vocal delivery.

Online learning can stimulate conventional content delivery, but it also allows for the incorporation of innovative ideas. In an online environment, the key to effective vocal training is to maintain a certain level of academicity with different teaching strategies. It is also important to keep learners motivated throughout the online course, which requires the use of specific design choices (Q. Li et al., 2024). Listening to music can be instrumental in developing musical skills. Through systematic reflection on music practices, active listeners can improve their knowledge and achieve autonomy (Esteve-Faubel et al., 2024).

Previous research has outlined the advantages of online learning, but the details of educational approaches have not been specified. This study focused on two delivery modalities (i.e., master classes and online seminars) and their effectiveness in developing the foundational abilities necessary to sing opera, and it informs us about their effect on cognitive skills and the potential challenges associated with developing opera singing skills through master classes and online seminars.

# Conclusions

The present findings suggest that individuals with prior vocal knowledge and skills can potentially use online seminars and master classes to learn the art of opera singing. The conducted analysis of the specifics of master classes and online seminars showed that master classes were found to be effective in expanding vocal range, teaching the technical skills needed to master the voice, and developing singing intonation. Seminars, on the other hand, can help vocalists achieve an emotional operatic performance and improve intonation. Master classes held by Juilliard Extension were superior to those from the Royal College of Music in terms of balance between the theoretical and practical material ($V$ = 0.93) and content variety ($V$ = 0.92). Master classes in general were found to be more helpful than online seminars for learners working out the operatic technique. The advantage of master classes lies in their capacity to provide individualized vocal training, allowing for consideration of the singer's physiological characteristics and enabling vocal adjustments in a remote format. This is facilitated through the provision of vocal demonstrations by the instructor, which enables real-time observation and analysis of vocal technique and dynamics. In contrast, seminars are predominantly oriented toward theoretical instruction, which limits students' exposure to the practical techniques essential for operatic performance. As regards online seminars, courses from Angel's Music Academy were considered better due to a more balanced and meaningful content mix ($V$ = 0.88).

Through observations, experts established that respondents who participated in master classes were able to develop their memory (24%), a skill that enables vocalists to construct a more thoughtful approach to operatic performance. This was associated with the analysis and assimilation of a substantial amount of practical information, which was subsequently applied during vocal performance. Online seminars were good for concentration (28%) but less effective at enhancing students' ability to perform unfamiliar compositions (13%). The results obtained were attributed to the lack of mechanisms for reinforcing the theoretical content during the seminar itself.

The significance of this study lies in its exploration of the effectiveness of distance education methods for teaching opera singing. Future research can assess the possibilities of using master classes and online seminars to teach a cohort of schoolchildren to sing opera.

# References

Aaberg, M. (2023). Virtual teaching in the applied string studio during the COVID-19 pandemic. *String Research Journal*, *13*(1), 51–71. https://doi.org/10.1177/19484992221121759

Almqvist, C. F., & Werner, A. (2024). Maintaining and challenging conservative teaching and learning culture in conservatories: The need for holistic pedagogy in educational fields of tension. *Research Studies in Music Education*, *46*(2), 257–270. https://doi.org/10.1177/1321103X231187766

Biasutti, M., Antonini Philippe, R., & Schiavio, A. (2022). Assessing teachers' perspectives on giving music lessons remotely during the COVID-19 lockdown period. *Musicae Scientiae*, *26*(3), 585–603. https://doi.org/10.1177/1029864921996033

Boersma, P., & Weenink, D. (2025). *Praat* (Version [6.4.31]) [Computer software]. Softonic. https://praat.en.softonic.com/

Chang, Z. (2022). The use of online vocal training programs as a means to develop creative thinking and vocal prowess. *Interactive Learning Environments*, *31*(10), 7214–7225. https://doi.org/10.1080/10494820.2022.2064514

Chen, Y. (2023). On the development of vocal music literature courses in colleges and universities: Taking Chinese art songs as an example. *International Journal of Education and Humanities*, *7*(2), 174–177. https://doi.org/10.54097/ijeh.v7i2.5572

Crisosto-Alarcón, J., Maldonado-Delgado, P., & Rivas-Campos, M. (2025). Conceptual metaphors about voice in the teacher–student interaction system during Western classical singing lessons. *Journal of Voice*. Advance online publication. https://doi.org/10.1016/j.jvoice.2025.02.021

Cuervo, L., Bonastre, C., Camilli, C., Arroyo, D., & García, D. (2023). Digital competences in teacher training and music education via service learning: A mixed-method research project. *Education Sciences*, *13*(5), Article 459. https://doi.org/10.3390/educsci13050459

de Villiers, A. C., & Gillmer, R. (2024). An assessment framework for contemporary commercial music (CCM) in higher education. *Journal of Applied Research in Higher Education*, *16*(4), 977–987. https://doi.org/10.1108/JARHE-11-2022-0337

Esteve-Faubel, J. M., Botella-Quirant, M. T., & Esteve-Faubel, R. P. (2024). Use of animated stories to improve music education practices of trainee primary school teachers in Spain. *British Journal of Music Education*, *41*(1), 87–102. https://doi.org/10.1017/S0265051723000244

Fan, X., Chen, R., Zheng, J., & Hu, Y. (2025). Research on the teaching value and promotion of virtual simulation platform for digital human vocal music education. *Frontiers in Interdisciplinary Educational Methodology*, *2*(2), 108–123. https://doi.org/10.71465/2tmpra62

Ilxomovna, A. C. (2025). Methodology for improving academic singing performance skills for music education students. *International Journal of Pedagogics*, *5*(1), 63–65. https://doi.org/10.37547/ijp/Volume05Issue01-18

Jansson, D., & Balsnes, A. H. (2021). Choral conducting education: The lifelong entanglement of competence, identity and meaning. *Research Studies in Music Education*, *43*(3), 347–365. https://doi.org/10.1177/1321103X19863184

Jansson, D., Døving, E., & Balsnes, A. H. (2024). Serious musicking: Choral pedagogy in "low-threshold" choirs. *Music Education Research*, *26*(4), 417–431. https://doi.org/10.1080/14613808.2024.2365722

Joseph, D., & Lennox, L. (2021). Twists, turns and thrills during COVID-19: Music teaching and practice in Australia. *Music Education Research*, *23*(2), 241–255. https://doi.org/10.1080/14613808.2021.1906852

Li, P.-P., & Wang, B. (2023). Artificial intelligence in music education. *International Journal of Human–Computer Interaction*, *40*(16), 4183–4192. https://doi.org/10.1080/10447318.2023.2209984

Li, Q., Tian, Z., Li, Z., Han, J., Zhang, Z., & Zhang, Y. (2024). Enhancing the online learning experience of performing arts students in the post-pandemic era: A large-scale empirical study in Hong Kong. *Sage Open*, *14*(2). https://doi.org/10.1177/21582440241243185

Li, G. (2024). Internet of Things and cloud computing integration of opera elements into vocal singing teaching practice exploration. *EAI Endorsed Transactions on Scalable Information Systems*, 11(1), 1–10. https://doi.org/10.4108/eetsis.4233

Li, Y. (2024). How does the music that students listen to affect their level of musical literacy? Comparative analysis of Chinese students' musical literacy formed with modern online technologies in the context of music and non-music academic majors. *Interactive Learning Environments*, *32*(8), 4688–4702. https://doi.org/10.1080/10494820.2023.2205895

Lim, E. H. C, Kan, R. Y. P, & Seng, N. S.-R. (2024). Courage to create: A new music education degree in instrumental and vocal teaching. *Education Sciences*, *14*(3), Article 315. https://doi.org/10.3390/educsci14030315

Liming, T., & Sartjinpong, N. (2025). The current situation of speech level singing method in the popular singing major at Xinghai Conservatory of Music, Guangzhou City. *Journal of Buddhist Education and Research (JBER)*, *11*(2), 590–598. https://so06.tci-thaijo.org/index.php/jber/article/view/283892

Løkke Jakobsen, M., Hebert, D. G., & Ørngreen, R. (2023). Synchronous online instrumental music teaching in cross-cultural learning contexts. *International Journal of Music Education*, *43*(2), 288–309. https://doi.org/10.1177/02557614231201916

Longlong, C., & Luen, L. C. (2023). New trend in online pedagogy for college vocal music education in the digital era. *International Journal of Academic Research in Progressive Education and Development*, *12*(4), 923–931. https://ijarped.com/index.php/journal/article/view/852

Lowery, C. L. (2023). The role of education in the public and its problems: A Deweyan perspective on political literacy. *Education and Culture*, *39*(1), 3–34. https://muse.jhu.edu/pub/60/article/923156/

Luo, Y., & Leung, B.-W. (2022). Proposing an assessment framework for Cantonese operatic singing after reviewing the current practices in Hong Kong and Guangdong, China. *Music Education Research*, *25*(1), 102–117. https://doi.org/10.1080/14613808.2022.2156490

Ng, D. T. K., Ng, E. H. L., & Chu, S. K. W. (2022). Engaging students in creative music making with musical instrument application in an online flipped classroom. *Education and information Technologies*, *27*(1), 45–64. https://doi.org/10.1007/s10639-021-10568-2

Ngobeni, N. C. (2024). Investigating virtual teaching experiences: Perspectives of undergraduate student teachers in singing and choral techniques. *International Journal of Learning, Teaching and Educational Research*, *23*(2), 184–202. https://doi.org/10.26803/ijlter.23.2.9

Nikolai, H., Kisiel, M., & Peng, Y. (2023). Vocal education in the context of intercultural communication: Experience of teaching Chinese students. *International Journal of Chinese Education*, *12*(1). https://doi.org/10.1177/2212585X231169741

The Norwegian National Committee for Research Ethics in Science and Technology. (2024, November 13). *Guidelines for research ethics in science and technology*. https://www.forskningsetikk.no/en/guidelines/science-and-technology/guidelines-for-research-ethics-in-science-and-technology/

Paparo, S. A. (2022). Singing with awareness: A phenomenology of singers' experience with the Feldenkrais method. *Research Studies in Music Education*, *44*(3), 541–553. https://doi.org/10.1177/1321103X211020642

Rexhepi, F. G., Breznica, R. K., & Rexhepi, B. R. (2024). Evaluating the effectiveness of using digital technologies in music education. *Journal of Educational Technology Development and Exchange, 17*(1), 273–289. https://doi.org/10.18785/jetde.1701.16

Rucsanda, M. D., Belibou, A., & Cazan, A. M. (2021). Students' attitudes toward online music education during the COVID 19 lockdown. *Frontiers in Psychology*, *12*, Article 753785. https://doi.org/10.3389/fpsyg.2021.753785

Rui, Y. (2024). Simulation of e-learning in vocal network teaching experience system based on intelligent Internet of things technology. *Entertainment Computing*, *50*, Article 100711. https://doi.org/10.1016/j.entcom.2024.100711

Shi, W. (2024). Diversified teaching strategies for ethnic instrumental music courses in colleges and universities in the information age. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 1–24. https://doi.org/10.2478/amns-2024-1126

Shu, L. (2023). Voice recognition and video vocal music teaching based on 5G wireless sensing technology. *International Journal of System Assurance Engineering and Management*. Advance online publication. https://doi.org/10.1007/s13198-023-02132-7

Silas, S., Müllensiefen, D., & Kopiez, R. (2024). Singing ability assessment: Development and validation of a singing test based on item response theory and a general open-source software environment for singing data. *Behavior Research Methods*, *56*(5), 4358–4384. https://doi.org/10.3758/s13428-023-02188-0

Smith, A. M., Kleinerman, K., & Cohen, A. J. (2022). Singing lessons as a path to well-being in later life. *Psychology of Music*, *50*(3), 911–932. https://doi.org/10.1177/03057356211030992

Tencent Holdings Limited. (2023). *WeChat* (Version [8.0.44]) [Mobile application]. https://www.wechat.com/

Urbaniak, O., & Mitchell, H. F. (2024). "Imagine you're in the Opera House …": Learning nonverbal communication for the concert stage. *Music Education Research*, *26*(2), 112–126. https://doi.org/10.1080/14613808.2024.2319581

Vaizman, T., & Harpaz, G. (2023). Retuning music teaching: Online music tutorials preferences as predictors of amateur musicians' music self-efficacy in informal music learning. *Research Studies in Music Education*, *45*(2), 397–414. https://doi.org/10.1177/1321103X221100066

Wang, Y. (2024a). Vocal creativity: Analyzing students song making processes in blended learning. *Interactive Learning Environments*, *32*(5), 2196–2206. https://doi.org/10.1080/10494820.2022.2141267

Wang, Y. (2024b). Vocal education in higher educational institutions in China: Student motivation and creativity. *Interactive Learning Environments*, *32*(3), 813–823. https://doi.org/10.1080/10494820.2022.2098778

Wang, Y. (2025). Challenges in music education in Chinese colleges and universities. *Journal of the Knowledge Economy*, *16*(2), 7934–7958. https://doi.org/10.1007/s13132-024-02191-6

Yuhan, F., & Sartjinpong, N. (2025). Research on the current situation of Chinese popular vocal music major by interview at Shenyang Conservatory of Music. *Journal of Buddhist Education and Research (JBER)*, *11*(2), 580–589. https://so06.tci-thaijo.org/index.php/jber/article/view/283891

Zeng, C. (2023). How to cultivate practice-oriented talents in higher vocational music teaching. *Contemporary Education and Teaching Research*, *4*(1), 26–29. https://doi.org/10.47852/bonviewCETR2023040107

**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Mobile Learning in Mathematics: Benefits, Challenges, Strategies, and Proponents' and Opponents' Views

**Rusen Meylani**
*Ziya Gokalp Faculty of Education, Dicle University, Türkiye*

## Abstract

Integrating mobile learning technologies in mathematics education in light of the pandemic, natural disasters, and recent technological developments is a significant shift in practice. This research applied the systematic review approach to assess the impact of mobile learning aids on students' learning attitudes, motivation, and performance in mathematics. Based on a synthesis of peer-reviewed articles published from 2010 to 2024, this paper focuses on five themes: (a) student and educator experience and engagement, (b) modification in instructional practices, (c) learning and achievements, (d) opportunities, challenges, and strategies, and (e) mobile learning perceptions. The review reveals that mobile learning enhances students' essential and higher-order thinking skills by establishing interactivity and learning through inquiry. However, other factors, such as the unequal distribution of resources and technology, still pose a challenge. To address these issues, it is suggested that future policies should focus on improving infrastructure, digital competency, and professional learning for instructors. Further, it is essential for institutions to engage faculty members in collaborative platforms to enhance their practice and have secure measures for data protection. This systematic review emphasizes the need for the researcher to conduct these evaluations over time to gauge the long-term impact of mobile learning and to look at the effects of this form of learning on different groups of students. It also concludes that future studies need to pay attention to ethical concerns and ways to provide equal opportunities in integrating technology into mathematics education for all students.

*Keywords*: mobile learning, mathematics, critical thinking, engagement, digital knowledge, professional development

# Introduction

This modern science-focused world requires people to be equipped with fundamental knowledge and skills. In the 21st century, instructors are responsible for teaching mathematics in a way that allows students to acquire practical and effective problem-solving skills (Szabo et al., 2020). As one of the most essential pillars of contemporary society, mathematics is inevitable in routine decision-making and in a modern globalized society; without sufficient knowledge of mathematics, progress can sometimes be challenging (Rizki & Priatna, 2019).

Math literacy is an essential requirement in the 21st century, yet it continues to pose challenges in elementary education (Tang et al., 2023). Although students understand its significance, many do not like the subject. Both instructors and students find mathematics boring, and one of the main reasons why they dislike it is the complexity of understanding complicated formulas. Therefore, math instructors should find the causes of math anxiety and use efficient classroom strategies to enable students to overcome these challenges (Gravemeijer et al., 2017).

In the past few years, the integration of mobile learning has received increased attention in mathematics education (Novita & Herman, 2021). *Mobile learning* can be defined as learning on mobile devices, including smartphones and tablets. It allows access to educational resources and participation in learning activities from any place at any time. It facilitates flexible, personal, and interactive learning not confined to a classroom learning environment (Krouska et al., 2024). Mobile learning technologies offer new approaches to enhancing students' activities and interest in mathematics (Ninghardjanti & Dirgatama, 2021). Nashrullah et al. (2023) undertook a review to establish the effects of incorporating mobile learning in enhancing Realistic Mathematics Education. They revealed that mobile learning can improve students' performance and thinking abilities (Nashrullah et al., 2023).

Mobile technologies in the mathematics teaching–learning process have, therefore, brought about a significant shift in instructional practices and demand the assessment of their qualitative impact on mathematics learning. Instructors have introduced mobile technologies using interactive math applications, digital whiteboards, and game-based learning platforms, which help to personalize the education process, give feedback to the students, and involve them in the process more actively. Alabdulaziz (2021) revealed that the COVID-19 pandemic stimulated the process of realizing digital technologies in math education. His research shows that over 98% of the participants considered the pandemic a driver of digital learning, and they reported using online platforms more often after the pandemic. Mobile devices, touchscreens, pen tablets, and digital libraries were identified among the main tools used to teach mathematics during COVID-19 (Alabdulaziz, 2021).

Mobile mathematics education closely resonates with distance education theory, which focuses on learning free from time constraints and providing flexible access to educational content and resources. This theory, introduced by Börje Holmberg in the 1960s, teaches that mobile devices are powerful instruments for providing mathematics instruction and resources beyond the classroom (Moore, 1991; Velmurugan et al., 2025). With mobile learning, students can learn various interactive math applications and access online libraries and virtual lessons anytime and anywhere, making learning easier and more customized. Distance education theory also emphasizes the necessity of interaction and communication between a learner and an instructor, which can also be provided with the help of mobile learning. Mobile technologies allow networked communication through social media and discussion

forums, as well as real-time feedback systems, which give students access to other students and tutors and can provide far more depth to the learning process (Vankúš, 2021).

It is important to understand the impact of mobile learning on mathematics for several reasons. These specifically include the effects of technological learning inclusion on enhancing the students' motivation, engagement, and learning outcomes. Tang et al. (2023) conducted a systematic review on 52 articles. Their findings offer four key aspects regarding the inclusion of mobile learning into mathematics education: (a) the significance of the learning design for the research purpose, (b) the application of the technology into the mathematical curriculum, (c) the global nature of mobile learning research, and (d) the widespread use of mobile devices in secondary education settings (Tang et al., 2023).

The review discussed the influence of mobile technology on education, focusing on mathematics. It highlighted how mobile learning has transformed traditional teaching methods, enhancing learning outcomes. The COVID-19 pandemic and natural disasters have underlined the need for flexible and accessible education solutions. Mobile learning became crucial during lockdowns, allowing students to access learning materials from home. The review emphasized the importance of mobile technologies in making education accessible to all students, especially those in underserved areas (Tang et al., 2023).

The research gap of this review (Tang et al., 2023) lies in the limited exploration of the implementation of mobile learning in mathematics. The review was performed to determine the impact of mobile learning from 2010 to 2024. The current review aims to establish how mobile learning can help transform mathematics education. It seeks to demonstrate how the use of mobile technologies could be incorporated in such a way as to foster a more facilitative and engaging learning process. In addition, the review provides insights into the challenges faced when adopting a mobile learning solution and the measures that can be taken to make education more robust and available.

## Research Objectives

The objectives of this research were to determine the following:

- How does mobile learning integration in mathematics impact students' experience and engagement?
- How does integrating mobile learning in mathematics modify the educator's instructional practices?
- How does learning mathematics through mobile methods affect students' learning outcomes and achievements?
- What are the opportunities and challenges of using mobile learning in mathematics, and which strategies can be applied to mitigate challenges?
- In what way does the perception of mobile learning in mathematics education differ between proponents and opponents in its implementation

# Methodology

This paper reports on a systematic review that involved searching relevant studies, selecting those meeting specific criteria, and synthesizing the findings. This section discusses the search strategy, the inclusion and exclusion criteria, and the data extraction and analysis methods.

## Search Strategy

A search strategy is a set of keywords that are used in such a way as to provide the relevant records in the target database. Formulating a search strategy is the first process in systematically searching the literature in the area of interest (Finfgeld-Connett & Johnson, 2013). To make the literature search more extensive, ERIC, Scopus, Web of Science, and Google Scholar databases were used to find the relevant academic papers. Furthermore, Boolean operators were used to help further narrow down the search results to include only the needed research studies while excluding unnecessary entries. Table 1 displays the search terms used to find the relevant keywords.

**Table 1**

*Search Strategies*

| Serial. No | Search strategy |
|---|---|
| 1. | ("Mobile Learning" OR "M-learning" OR "Mobile Education" OR "Mobile Technologies") AND ("Mathematics" OR "Math Education" OR "Math Instruction") |
| 2. | ("Student Engagement" OR "Learning Outcomes" OR "Academic Performance" OR "Motivation") AND ("Mobile Learning" OR "Mobile Tools") AND ("Mathematics") |
| 3. | ("Teacher Practices" OR "Instructional Strategies" OR "Teaching Methods") AND ("Mobile Learning" OR "Digital Tools") AND ("Mathematics") |
| 4. | ("Challenges" OR "Barriers" OR "Limitations") AND ("Mobile Learning") AND ("Math" OR "Education") |
| 5. | ("Perceptions" OR "Attitudes" OR "Views") AND ("Students" OR "Instructors") AND ("Mobile Learning") AND ("Math") |
| 6. | ("Gamification" OR "Augmented Reality" OR "Interactive Apps" OR "Educational Technology") AND ("Mathematics Education") AND ("Mobile Use" OR "Mobile Platforms") |
| 7. | ("Critical Thinking" OR "Problem-Solving" OR "Cognitive Skills") AND ("Mobile Learning") AND ("Math Education") |

## Review Selection Criteria

The term *selection criteria* relates to the general principles defining the range of research articles and studies to be included in the review (Gentles et al., 2016). Table 2 depicts the inclusion and exclusion criteria used in this review.

**Table 2**

*Inclusion and Exclusion Criteria*

| Inclusion criteria | Exclusion criteria |
|---|---|
| Articles that investigated mobile learning interventions in mathematics education | Articles not concentrating on mathematics education through mobile learning |
| Articles written explicitly in the English language | Articles written in any language other than English |
| Articles published from 2010 to 2024 | Articles that were published before 2010 |
| Articles published in peer-reviewed journals or reputable conference proceedings | Articles needing more rigorous methodological frameworks |

## Data Analysis

The data from the selected studies were coded thematically with MAXQDA 24 (Verbi, 2025); inductive coding was employed to elicit patterns and themes. These themes were cross-mapped across studies, clustered, and synthesized to present an overall picture of the influence of mobile learning on math learning. This systematic review analyzed existing research on the impact of mobile learning in mathematics education, focusing on five key themes: student and educator engagement, instructional practices, learning outcomes, challenges and strategies, and perceptions. Findings were synthesized from multiple studies to highlight the benefits and limitations of mobile technologies, offering insights into how they shape teaching methods, student achievement, and equitable access to learning.

# Findings and Results

The findings of this review provide a comprehensive understanding of mobile learning's role in mathematics education. They are organized into five main themes: (a) student and educator experience and engagement, (b) modification in instructional practices, (c) learning and achievements, (d) opportunities, challenges, and strategies, and (e) mobile learning perceptions. Table 3 shows the themes, subthemes, and number of papers involved in each theme, along with the references.

## Table 3

*Main Themes, Subthemes, and Supporting References*

| Theme | Subtheme | References | No. of references |
|---|---|---|---|
| Student and educator experience and engagement | Students' interest and engagement | Al Mulhem & Almaiah (2021); Kang (2024); Krouska et al. (2022); Li et al. (2024); Svela et al. (2019); Vankúš (2021); Wong et al. (2022) | 7 |
| | Students' and instructors' narratives and experiences | Al Omoush & Mehigan (2023); Chen & Zainal Abidin (2023); Dawodi et al. (2023); Edmonds & Smith (2017); Zhang et al. (2023); Ibrahim (2024) | 6 |
| Modification in instructional practices | Educator's mobile learning instructional strategies | Alabdulaziz (2021); Borba et al. (2016) | 2 |
| | Educator's adaptations and professional development needs | Kaliisa & Michelle (2019); Papadakis et al. (2021); Rohadi & Yavani (2023); Taleb et al. (2015); Tang et al. (2023) | 5 |
| Learning and achievements | Understanding of students' learning and their achievements | Song & Cai (2024); Suyatmo et al. (2023) | 2 |
| | Enhancing problem-solving and critical thinking skills | Huda et al. (2022); Ninghardjanti & Dirgatama (2021) | 2 |
| | Improving higher-order thinking skills and cognitive abilities | Astuti et al. (2023); Sucilestari & Arizona (2020) | 2 |
| Opportunities, challenges, and strategies | Technological, pedagogical, and contextual challenges | Dawodi et al. (2023); Ghoulam et al. (2024); Kaliisa & Michelle (2019) | 3 |
| | Strategies to solve barriers | Barana & Marchisio Conte (2023); Ghoulam et al. (2024); Upadhayaya (2023) | 3 |
| Mobile learning perceptions | Proponent views | Burke et al. (2022); Drigas & Pappas (2015); Marques & Pombo (2021); Taleb et al. (2015) | 4 |
| | Opponent views | Astuti et al. (2023); Chen et al. (2022); Deswita et al. (2024); Ghimire (2023) | 4 |

The literature review findings highlight that mobile learning improves students' participation in lessons, changes instructors' approaches toward instruction, and positively affects learning. They also stress the implementation issues and challenges and showcase differing opinions regarding the efficiency of mobile learning. Furthermore, they reveal the strengths, weaknesses, opportunities, and threats of adopting technology into the system.

## Theme 1: Student and Educator Experience and Engagement

Mobile learning's influence on students' academic achievement and learning experience has become a research focus in mathematics education. Including flexible technologies through mobile methods enhances students' overall learning experience and motivation as students use these technologies daily. Enhanced lessons are developed for mobile devices to implement interactive media and gaming elements in mathematics learning. By analyzing students' motivation, attitudes, and cognitive abilities, several authors have observed that the inclusion of mobile learning is associated with a positive learning experience and engagement (Al Mulhem & Almaiah, 2021; Krouska et al., 2022; Svela et al., 2019). This theme explored the impact of mobile learning on students and instructors, focusing on their experiences and levels of engagement.

### *Students' Interest and Engagement*

Several past studies have recognized the role of mobile learning in enhancing students' learning interest and motivation in mathematics. Mobile technologies offer instructors the flexibility to create effective lesson plans that fulfil this generation of students' needs. They are proficient in using mobile devices and tablets (Krouska et al., 2022; Svela et al., 2019). The adoption of mobile games tied to learning has been shown to improve students' engagement and performance (Al Mulhem & Almaiah, 2021). According to Vankúš (2021), mobile learning captures students' attention and motivates them to learn more, primarily through gaming with appropriate learning approaches.

Research findings (Kang, 2024) have pointed out that game-based learning boosts students' motivation, engagement, attitudes, pleasure, and state of flow in mathematics education. Mobile learning implementation also increases student interaction with their instructors, as well as learning materials, develops and maintains positive attitudes, and increases self-generated motivation and learning outcome effectiveness (Li et al., 2024). Furthermore, research has established that the use of games in supporting education in the digital age effectively boosts students' learning engagement, enhancing their motivation to learn (Wong et al., 2022).

### *Students' and Instructors' Narratives and Experiences*

The educational applications of augmented reality and gamification in mathematics have gained much popularity due to their efficiency in reinforcing students' motivation, engagement, and knowledge improvement (Al Omoush & Mehigan, 2023). In addition, educational robotics has also improved students' access to, attention to, and learning interest in mathematics through better and more playful interactive tools (Edmonds & Smith, 2017). Furthermore, location-based mobile learning activities have been found to present students with active, enjoyable, and realistic learning experiences, thus enabling them to engage with mobile content, different places, and other students (Zhang et al., 2023).

To enhance students' learning experience in mathematics education, the relationship between students' and instructors' motivation, engagement, and performance should be studied (Ibrahim, 2024). Students mentioned that using interactive tools gave them autonomy and the willingness to provide

mathematical material with peers and instructors as part of the learning process. Instructors also recounted that mobile learning has improved classroom interaction and helped them employ differentiated instruction (Chen et al., 2022).

Afghan students have exemplified mobile learning platforms as essential tools for bridging gaps in education, particularly where conventional learning resources have been limited or unavailable. Students noted that mobile technologies provided them with flexibility and accessibility in cases of other external challenges, such as environmental issues. Instructors observed that such tools enabled them to reach distant students in widespread locations, thereby being inclusive (Dawodi et al., 2023). Thus, instructors and students repeatedly emphasized how using such technologies increased motivation, achieved greater depth of learning, and resulted in a more engaging classroom atmosphere.

## Theme 2: Modification in Instructional Practices

Implementing mobile learning technologies in mathematics learning has required new teaching strategies. Instructors are seeking new ways of embedding mobile learning devices into practice to improve students' learning outcomes. This means that teaching methodologies must be modified and materials must be created to inform instructors on how to employ mobile technologies in teaching math. This theme explores the professional learning and pedagogical approaches required of instructors to teach through mobile learning.

### *Educators' Mobile Learning Instructional Strategies*

Mobile learning has impacted the delivery of mathematics education through modifications in instructional methods, prompting teachers to improve their practices and seek professional development. Instructors have embraced mobile technology in the classroom to expand mathematics learning in this era of technology. According to Borba et al. (2016), to prepare instructors to incorporate mobile technologies, digital libraries, and collaborative learning into mathematics, they must be trained through blended learning.

Instructors have shifted from traditional teaching approaches to adapt mobile devices in teaching mathematics, whereby new approaches are incorporated to capture students' attention and enhance their learning. Alabdulaziz (2021) organizes the use of mobile devices in mathematics education into categories, including applications for teaching and learning, practical uses for mobile devices, and the education of mathematics instructors, underlining the importance of instructors' acceptance of technologies in their teaching practices.

### *Educator's Adaptations and Professional Development Needs*

Instructors have been encouraged to search for effective instructional strategies in teaching mathematics that can be implemented in collaboration with the functionalities of mobile devices, such as interactive simulations, augmented reality, and gamification. Following the adoption and incorporation of mobile technologies in learning institutions, Tang et al. (2023) noted that educational researchers have deemed this new form of learning "mobile learning." Mobile learning has brought vitality and reform to conventional models of learning (Tang et al., 2023). Furthermore, mobile learning has empowered instructors to deliver personalized instructions, integrate students' learning styles, and create engaging and innovative learning solutions. Taleb et al. (2015) emphasize that technology plays a vital role in reshaping teaching methods and increasing students' interest in learning. In particular,

mobile learning has been shown to be especially effective in motivating students to engage with mathematics.

According to Papadakis et al. (2021), instructors have explored mobile learning to implement Realistic Mathematics Education in classrooms in an engaging manner to improve student learning outcomes. Papadakis et al. (2021) noted that better teaching practices are key to adopting mobile technologies in mathematics to design learning activities that are contextually related to students' interests and real-life situations in mathematics.

Instructors have realized the importance of engaging in professional development to update their practices while implementing change regarding mobile learning technologies in mathematics education in the learning environment. Rohadi and Yavani (2023) recognized that teacher training and continued professional development make up one way that instructors may ensure the use of mobile learning technologies in online mathematics courses. They emphasized the importance of embracing contingency in instructional practices for digital students (Rohadi & Yavani, 2023).

Kaliisa and Michelle (2019) have identified different mobile learning policies and strategies, focusing on Education for All, specifically in higher learning institutions. To address inequalities, instructors must develop a clear understanding of the context and realities within their locality (Kaliisa & Michelle, 2019). Thus, it can be concluded that mobile learning technologies have influenced instructors by shifting their conventional teaching practice, requiring them to adopt different behaviors and engage in professional development to integrate technologies into the teaching–learning process. Instructors must provide positive and productive learning climates that support students and improve mathematical learning through mobile technologies in the classroom.

## Theme 3: Learning and Achievements

Mobile learning technologies enhance math knowledge by developing students' mental abilities, critical thinking, and problem-solving. They ensure the students' active involvement with the content by providing interactive, engaging, and customized experiences. This theme deals with how such tools assist students in applying math concepts to real-life problems.

### *Understanding Students' Learning and Their Achievements*

Mobile learning technologies have given instructors new tools that assist students in enhancing their thinking skills and developing a deeper understanding of mathematical concepts. Incorporating mobile learning with inquiry-based learning models improves students' learning achievements and develops their problem-solving skills (Suyatmo et al., 2023). According to Song and Cai (2024), mobile learning applications have improved students' critical thinking abilities in mathematics education. Instructors using mobile learning to create effective communication platforms help students develop their critical thinking and problem-solving skills by effectively applying, assessing, and analyzing math concepts in various practice situations (Song & Cai, 2024).

### *Enhancing Problem-Solving and Critical Thinking Skills*

The effective adoption of mobile devices enhances critical thinking among students as it promotes problem-solving and self-learning skills. Instructors can help students develop analytical and evaluative skills by incorporating critical thinking tasks into mobile-based interactive learning media. Tasks involving technology also incorporate exploration and inquiry skills, which are necessary in

mathematics (Ninghardjanti & Dirgatama, 2021). Huda et al. (2022), in their research, mentioned that the use of mobile devices in mathematics helped students to develop inquiry-based skills. Including mobile devices promotes analytical thinking, communication, and interpersonal skills to help students overcome complicated problem-solving challenges in mathematics (Huda et al., 2022).

### *Improving Higher-Order Thinking Skills and Cognitive Abilities*

The adoption of inquiry-based learning models supported by mobile technologies has effectively enhanced cognitive processes and critical thinking skills among students. The inquiry-based model developed on mobile devices is based on different problem-solving, critical thinking, and reflective tasks that provide interactive learning experiences. This model enhances students' understanding of a particular subject, such as math, and allows them to handle daily life problems (Sucilestari & Arizona, 2020). Astuti et al. (2023) indicated that including mobile learning helps develop rigorous mathematical thinking that supports higher-level students' mathematical-level concepts. Therefore, this theme proves that mobile learning enhances the critical thinking and problem-solving skills of students. Using mobile technologies helps instructors design engaging and rich learning environments that foster students' critical thinking skills and enable them to succeed in mathematics education.

## Theme 4: Opportunities, Challenges, and Strategies

Mobile learning technologies have the potential to enhance mathematics education, but their use also comes with challenges, such as issues with device compatibility and network connection. In this context, instructors need to find ways to make fair distributive decisions and address issues of diversity and inequality. This theme highlights strategies to overcome the mobile learning integration challenges and improve effectiveness.

### *Technological, Pedagogical, and Contextual Challenges*

Instructors and institutions face several pedagogical, technological, and contextual challenges while integrating mobile learning into mathematics education. The most common technological challenges that instructors face are related to the compatibility of devices, network access, and the presence of effective software. The successful implementation of mobile application software for mathematics education can be hindered by ineffective software, which disrupts seamless communication across various devices (Dawodi et al., 2023). Furthermore, in the context of mobile learning, data security and privacy issues require robust protocols and measures (Ghoulam et al., 2024).

According to Kaliisa and Michelle (2019), instructors who adopt a mobile technological approach to teaching may need help designing content that fully makes use of mobile devices to provide students with an interactive learning experience. Research suggests adequate planning for instructors to guarantee they use mobile devices efficiently to produce math material that engages students (Kaliisa & Michelle, 2019).

### *Strategies to Solve Barriers*

There are many challenges related to mobile learning, including resource and culture differences, as well as institutional limitations. Technological inequalities in resources, inadequate use of mobile devices, and poor infrastructure negatively impact effective mobile learning in mathematics. To address these barriers, applying systemic approaches that guarantee mobile learning for all students is essential (Kaliisa & Michelle, 2019).

Instructors and institutions may effect several measures. According to Ghoulam et al. (2024), implementing technical support and training enables instructors to enhance their understanding of mobile technology for teaching. Moreover, data protection policies should ensure that student rights are not violated and regulatory laws are followed (Ghoulam et al., 2024). Upadhayaya (2023) noted that instructors should collaborate to effectively share knowledge regarding appropriate mobile learning practices. Furthermore, instructors should tailor their teaching to their students' needs and promote continuous improvement to provide compelling mathematical learning experiences (Upadhayaya, 2023).

For effective mobile learning to enhance socioeconomic equity, specific formative assessment tools, such as marking criteria, must be appropriately employed in education systems to ensure all students are treated equally (Barana & Marchisio Conte, 2023). Thus, overcoming challenges for mobile learning in mathematics requires a holistic approach with an aim to empower instructors to create engaging and inclusive learning environments.

## Theme 5: Mobile Learning Perception

The use of mobile learning in mathematics education has elicited positive and negative responses. Its proponents claim that creativity and adaptability increase students' engagement and learning. The opponents, alternately, are concerned about factors that may hinder the effectiveness of such learning strategies, including distractions, reduction in face-to-face learning, and issues related to the digital divide. This theme covers both perspectives to help instructors create fair and efficient strategies for integrating mobile technologies into learning environments that are welcoming to students with a wide range of needs.

### *Proponent Views*

Proponents of mobile learning in mathematics education have cited that its use improves learning achievement, motivation, and interaction. Taleb et al. (2015) confirmed that using mobile technologies enhances students' motivation to learn as it sparks their interest and enthusiasm (Taleb et al., 2015). In their research, Drigas and Pappas (2015) emphasized that mobile learning activities increase critical thinking in students, and therefore, more mobile learning activities should be implemented. Furthermore, according to Burke et al. (2022), integrating technology for educational purposes can enhance students' perceptions of learning outcomes, especially in mathematics, by providing them with technology-enhanced collaborative and realistic activities. By reviewing instructors' experiences with incorporating mobile augmented reality games, Marques and Pombo (2021) provided insight into the impact of mobile learning on students and teaching approaches.

### *Opponent Views*

Opponents of mobile learning in mathematics education have voiced concerns and criticism, focusing on its negative side. Some critics argue that mobile learning can lead to diversions, decreased face-to-face communication, and decreased student attention toward learning (Astuti et al., 2023). Furthermore, as per Deswita et al. (2024), mobile learning can cause a digital divide, privacy issues, and overreliance on technology.

Chen et al. (2022) performed qualitative research investigating how improper mobile phone use affects students' anxiety and self-confidence in mathematics and proposed prospective challenges in and drawbacks of using it in mathematics learning. They noted that mobile learning causes math anxiety

and a lack of self-confidence in problem-solving, adversely affecting cognitive skills (Chen et al., 2022). Moreover, Ghimire's (2023) research has highlighted concerns about equity of access to mobile gadgets for all students and the effects of using mobiles for learning purposes on the psychosocial health of students.

# Discussion

This systematic review has discussed the implications of mobile technologies on mathematics learning, basing its findings on students' and instructors' experiences. It has reviewed major themes that highlight increasing momentum toward using digital tools to stimulate student engagement, motivation, and performance in mathematics. Across the literature, mobile technologies are noted to act as key drivers in shaping adaptive, interactive, and student-focused learning environments.

Five themes were identified: (a) student and educator experiences and engagement, (b) modification in instructional practices, (c) learning and achievements, (d) opportunities, challenges, and strategies, and (e) mobile learning perceptions.

The findings related to the first theme affirm that mobile learning significantly improves students' and instructors' experience because it is more engaging, flexible, and motivating when it comes to mathematical learning (Al Mulhem & Almaiah, 2021; Krouska et al., 2022; Svela et al., 2019). Mobile-based game elements and interactive content attracted students and made them mentally active (Al Mulhem & Almaiah, 2021). Instructors discovered the usefulness of mobile technologies in developing adaptive lesson plans with the learner as the focal point (Vankúš, 2021). This implies that mobile learning not only increases academic performance but also creates a more comfortable and animated learning process (Kang, 2024).

The second theme reveals that mobile learning integration in mathematics changes teaching practices. It requires teachers to make efforts to acquire new approaches to teaching mathematics and to participate in professional development on an ongoing basis. These findings are consistent with previous research on the significance of teacher training and how mobile technologies should be contextualized to students' interests and learning needs to promote motivation and better achievements (Borba et al., 2016; Taleb et al., 2015). Furthermore, changes related to traditional pedagogy and the introduction of technology-enhanced and individualized education indicate one of the shifts in the educational sector that necessitate adaptive teaching behavior to develop engaging and supportive learning scenarios (Papadakis et al., 2021; Rohadi & Yavani, 2023). Successful implementation of mobile learning cannot be achieved solely through the acceptance of technology; it also requires experts in educator development.

The third theme reveals that mobile learning technologies can help significantly improve the mathematical literacy skills of students by cultivating higher-order thinking skills and abilities associated with critical thinking, problem-solving, and inquiry-based learning models. Mobile tools make possible the development of individual experiences that are personal and interesting, as they help students to effectively apply math concepts to real-life situations. Research findings are consistent about the impact of mobile devices in math learning, which assumes an integrative approach and prompt availability of material that propagates analytic, critical, and contemplative thinking, of which the former is fundamental to intuitive learning and achievements in academics (Astuti et al., 2023; Song & Cai, 2024; Suyatmo et al., 2023). This highlights how mobile learning can enrich the learning environment with more effective learning empowerment of students and instructors.

The fourth theme reveals that mobile learning in mathematics education presents tremendous opportunities but faces technological, pedagogical, and contextual challenges such as issues in device compatibility, network connectivity, and data protection. These need to be addressed through systemic responses and strategies including technical assistance, specific instructor training, and balanced resource allocation for assured equal access (Dawodi et al., 2023; Ghoulam et al., 2024; Kaliisa & Michelle, 2019). Furthermore, teacher cooperation and ongoing adjustments in instructional techniques are required to battle the aforementioned problems and develop effective, interesting learning spaces (Barana & Marchisio Conte, 2023; Upadhayaya, 2023).

Lastly, the fifth theme reveals that there are differing perceptions of the concept of mobile learning in mathematics education. Proponents state that by encouraging creative activities and collaboration, mobile learning promotes students' learning by enhancing motivation, critical thinking, learning, and engagement (Drigas & Pappas, 2015; Taleb et al., 2015). On the other hand, opponents discuss the possibility of distractions, diminished face-to-face communication, and equity concerns regarding the digital divide, in addition to the suggestions that mobile use may develop math anxiety as well as decreased confidence in students (Astuti et al., 2023; Chen et al., 2022; Ghimire, 2023). An equilibrium between these views is essential for comprehensive and viable mobile learning strategies.

This review's findings align with the distance education theory because they underscore a flexible and learner-centered setting with the help of mobile technologies that surpass conventional classroom borders and foster connectivity and personalized learning (Velmurugan et al., 2025). The findings imply that the translation of mobile learning may require institutions to invest in sound infrastructure, instructors' ongoing professional development, and policies of inclusion in order to capitalize fully on some of the benefits of mobile learning and resolve some of the equity and access issues presented by distance education environments.

## Recommendations

The following are recommendations for the execution of mobile learning in educational institutions:

- Authorities and educational leaders should ensure that every learner has access to mobile learning tools and a stable Internet connection; this will remove the technological and situational aspect of accessibility that hinders equity in mathematics learning opportunities (Dawodi et al., 2023; Kaliisa & Michelle, 2019).

- Governments should ensure that instructors receive continued training that integrates technical skills and pedagogical approaches toward the successful integration of mobile technology that fosters differentiation and engagement among students (Borba et al., 2016; Rohadi & Yavani, 2023; Taleb et al., 2015).

- Learning institutions must create professional learning communities for teachers to share best practices, issues, and innovations among each other and so they can continue adapting and improving when using mobile learning in learning institutions (Upadhayaya, 2023).

## Limitation and Future Implications

The present review's limitation is that the findings are limited to the mathematics classroom only; the results may not be generalizable to other contexts or have the same relevance in the future. Nevertheless, important conclusions can be made regarding the application of mobile learning in

teaching and students' interactions. It also emphasizes the lack of literature and gives recommendations for trainers and policymakers implementing mobile learning for students.

## Conclusion

Mobile learning technologies promise to improve student engagement, learning, and teaching activities in mathematics education, but they come with some concerns that must be addressed. Thus, a careful approach, constant assessment, and adherence to equality principles are critical for achieving the intended positive outcomes without experiencing significant negative consequences. In this way, mobile learning will benefit students and instructors by improving the efficiency of educational processes and, more importantly, providing a better chance for all students to receive the educational opportunities they need and deserve.

# References

Alabdulaziz, M. S. (2021). COVID-19 and the use of digital technology in mathematics education. *Education and Information Technologies*, *26*(6), 7609–7633. https://doi.org/10.1007/s10639-021-10602-3

Al Mulhem, A., & Almaiah, M. A. (2021). A conceptual model to investigate the role of mobile game applications in education during the COVID-19 pandemic. *Electronics*, *10*(17), Article 2106. https://doi.org/10.3390/electronics10172106

Al Omoush, M. H., & Mehigan, T. (2023, June 8–10). *Leveraging robotics to enhance accessibility and engagement in mathematics education for vision-impaired students*. 2023 5th International Congress on Human–Computer Interaction, Optimisation and Robotic Applications (HORA), Istanbul, Turkiye. https://doi.org/10.1109/HORA58378.2023.10156748

Astuti, P., Febrian, F., Pratiwi, W. D., Hayati, H. S., Aruny, Y. D., & Diniati, S. (2023). Developing rigorous mathematical thinking approach-based learning materials for sequences and series concept at high school level in Tanjungpinang. *BIO Web of Conferences*, *79*, Article 05005. https://doi.org/10.1051/bioconf/20237905005

Barana, A., & Marchisio Conte, M. (2023). Promoting socioeconomic equity through automatic formative assessment. *Journal on Mathematics Education*, *15*(1), 227–252. https://doi.org/10.22342/jme.v15i1.pp227-252

Borba, M. C., Askar, P., Engelbrecht, J., Gadanidis, G., Llinares, S., & Aguilar, M. S. (2016). Blended learning, e-learning and mobile learning in mathematics education. *ZDM*, *48*, 589–610. https://doi.org/10.1007/s11858-016-0798-4

Burke, P. F., Kearney, M., Schuck, S., & Aubusson, P. (2022). Improving mobile learning in secondary mathematics and science: Listening to students. *Journal of Computer Assisted Learning*, *38*(1), 137–151. https://doi.org/10.1111/jcal.12596

Chen, S., Jiang, H., & Guan, J. (2022). Exploring the influence of problematic mobile phone use on mathematics anxiety and mathematics self-efficacy: An empirical study during the COVID-19 pandemic. *Sustainability*, *14*(15), Article 9032. https://doi.org/10.3390/su14159032

Dawodi, M., Baktash, J. A., & Dawodi, S. M. R. (2023). Leveraging mobile learning platforms for flexible education delivery: Bridging educational gaps in Afghanistan. *Journal of Electrical Electronics Engineering*, *2*(4), 539–547. https://doi.org/10.33140/JEEE.02.04.18

Deswita, H., Japar, M., & Solihatin, E. (2024). Integrating collaborative mobile learning into the mathematics classroom: An exploration of student'meaningful learning activities using technology. *Proceedings of International Conference on Education*, *2*(1), 562–571. https://doi.org/10.32672/pice.v2i1.1414

Drigas, A. S., & Pappas, M. A. (2015). A review of mobile learning applications for mathematics. *International Journal of Interactive Mobile Technologies*, *9*(3), 18–23. https://doi.org/10.3991/ijim.v9i3.4420

Edmonds, R., & Smith, S. (2017). From playing to designing: Enhancing educational experiences with location-based mobile learning games. *Australasian Journal of Educational Technology*, *33*(6), 41–53. https://doi.org/10.14742/ajet.3583

Finfgeld-Connett, D., & Johnson, E. D. (2013). Literature search strategies for conducting knowledge-building and theory-generating qualitative systematic reviews. *Journal of Advanced Nursing*, *69*(1), 194–204. https://doi.org/10.1111/j.1365-2648.2012.06037.x

Gentles, S. J., Charles, C., Nicholas, D. B., Ploeg, J., & McKibbon, K. A. (2016). Reviewing the research methods literature: Principles and strategies illustrated by a systematic overview of sampling in qualitative research. *Systematic Reviews*, *5*, Article 172. https://doi.org/10.1186/s13643-016-0343-0

Ghimire, S. P. (2023). Use of mobile/smart phones and students' mathematics learning: A case of basic schools in Nepal. *Pragyaratna प्रज्ञारत्न*, *5*(1), 108–118. https://doi.org/10.3126/pragyaratna.v5i1.59277

Ghoulam, K., Bouikhalene, B., Babori, A., & Falih, N. (2024). Exploring the impact of mobile devices in electronics e-learning: A case study evaluating the effectiveness of mobile learning applications in the field of electronics and sensors. *Advances in Mobile Learning Educational Research*, *4*(2), 1058–1072. https://doi.org/10.25082/AMLER.2024.02.001

Gravemeijer, K., Stephan, M., Julie, C., Lin, F.-L., & Ohtani, M. (2017). What mathematics education may prepare students for the society of the future? *International Journal of Science and Mathematics Education*, *15*, 105–123. https://doi.org/10.1007/s10763-017-9814-6

Huda, S., Mawaddah, S., & Elvierayani, R. R. (2022). Design of an inquiry-based mathematical literacy learning model to encourage critical thinking. *DIDAKTIKA: Jurnal Pemikiran Pendidikan*, *28*(2), 18–28. https://doi.org/10.30587/didaktika.v28i2(1).4408

Ibrahim, U. (2024). Assessing the impact of mobile applications on student engagement in ICT and computer science education. *International Journal of Applied and Advanced Multidisciplinary Research*, *2(3)*, 221–232. https://doi.org/10.59890/ijaamr.v2i3.1533

Kaliisa, R., & Michelle, P. (2019). Mobile learning policy and practice in Africa: Towards inclusive and equitable access to higher education. *Australasian Journal of Educational Technology*, *35*(6), 1–14. https://doi.org/10.14742/ajet.5562

Kang, J. (2024). Benefits and challenges of mobile-learning brought to student learning outcomes in higher education: A systematic review from 2014–2023. *International Journal of Academic Research in Progressive Education and Development*, *13*(1), 2107–2123. https://doi.org/10.6007/IJARPED/v13-i1/20698

Krouska, A., Troussas, C., Kabassi, K., & Sgouropoulou, C. (2024). An empirical investigation of user acceptance of personalized mobile software for sustainability education. *International Journal of Human–Computer Interaction*, *40*(19), 5876–5883. https://doi.org/10.1080/10447318.2023.2241614

Krouska, A., Troussas, C., & Sgouropoulou, C. (2022). Mobile game-based learning as a solution in COVID-19 era: Modeling the pedagogical affordance and student interactions. *Education and Information Technologies*, *27*(1), 229–241. https://doi.org/10.1007/s10639-021-10672-3

Li, Y., Chen, D., & Deng, X. (2024). The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PLoS One*, *19*(1), Article e0294350. https://doi.org/10.1371/journal.pone.0294350

Marques, M., & Pombo, L. (2021). Teachers'experiences and perceptions regarding mobile augmented reality games: A case study of a teacher training. In L. G. Chova, A. L. Martínez, & I. C. Torres (Eds.), *INTED2021 proceedings: 15th International Technology, Education and Development Conference* (pp. 8938–8947). International Academy of Technology, Education and Development (IATED). https://doi.org/10.21125/inted.2021.1865

Moore, M. G. (1991). Editorial: Distance education theory. *American Journal of Distance Education*, *5*(3), 1–6. https://doi.org/10.1080/08923649109526758

Nashrullah, F. R., Rochmad, R., & Cahyono, A. N. (2023). Mathematical critical thinking abilities of students in terms of self-regulated learning in realistic mathematics education assisted by mobile learning. *Mathline: Jurnal Matematika dan Pendidikan Matematika*, *8*(3), 1035–1056. https://doi.org/10.31943/mathline.v8i3.469

Ninghardjanti, P., & Dirgatama, C. H. A. (2021). Building critical thinking skills through a new design mobile-based interactive learning media knowledge framework. *International Journal of Interactive Mobile Technologies*, *15*(17), 49–68. https://doi.org/10.3991/ijim.v15i17.23801

Novita, R., & Herman, T. (2021). Digital technology in learning mathematical literacy, can it helpful? *Journal of Physics: Conference Series*, *1776*, Article 012027. https://doi.org/10.1088/1742-6596/1776/1/012027

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2021). Teaching mathematics with mobile devices and the Realistic Mathematical Education (RME) approach in kindergarten. *Advances in Mobile Learning Educational Research*, *1*(1), 5–18. https://doi.org/10.25082/AMLER.2021.01.002

Rizki, L., & Priatna, N. (2019). Mathematical literacy as the 21st century skill. *Journal of Physics: Conference Series*, *1157*(4), Article 042088. https://doi.org/10.1088/1742-6596/1157/4/042088

Rohadi, T., & Yavani, Z. (2023). Exploring the use of mobile-assisted language learning (MALL) for EFL learning at a junior high school in cirebon. *Foreign Language Instruction Probe*, *2*(2), 128–142. https://doi.org/10.54213/flip.v2i2.307

Song, H., & Cai, L. (2024). Interactive learning environment as a source of critical thinking skills for college students. *BMC Medical Education*, *24*(1), Article 270. https://doi.org/10.1186/s12909-024-05247-y

Sucilestari, R., & Arizona, K. (2020). The impact of inquiry-based learning on students' critical thinking skills. In *Proceedings of the 2nd International Conference on Islam, Science and Technology (ICONIST 2019)* (pp. 172–175). Atlantis Press. https://doi.org/10.2991/assehr.k.200220.031

Suyatmo, S., Yustitia, V., Santosa, T. A., Fajriana, F., & Oktiawati, U. Y. (2023). Effectiveness of the inquiry based learning model based on mobile learning on students' creative thinking skills: A meta-analysis. *Jurnal Penelitian Pendidikan IPA*, *9*(9), 712–720. https://doi.org/10.29303/jppipa.v9i9.5184

Svela, A., Nouri, J., Viberg, O., & Zhang, L. (2019). A systematic review of tablet technology in mathematics education. *International Journal of Interactive Mobile Technologies (iJIM)*, *13*(8), 139–158. https://doi.org/10.3991/ijim.v13i08.10795

Szabo, Z. K., Körtesi, P., Guncaga, J., Szabo, D., & Neag, R. (2020). Examples of problem-solving strategies in mathematics education supporting the sustainability of 21st-century skills. *Sustainability*, *12*(23), Article 10113. https://doi.org/10.3390/su122310113

Taleb, Z., Ahmadi, A., & Musavi, M. (2015). The effect of M-learning on mathematics learning. *Procedia—Social and Behavioral Sciences*, *171*, 83–89. https://doi.org/10.1016/j.sbspro.2015.01.092

Tang, D. M., Nguyen, C. T. N., Bui, H. N., Nguyen, H. T., Le, K. T., Truong, K. L. G., Tran, N. T., Vo, N. K., & Nguyen, T. T. (2023). Mobile learning in mathematics education: A systematic literature review of empirical research. *EURASIA Journal of Mathematics, Science and Technology Education*, *19*(5), Article em2268. https://doi.org/10.29333/ejmste/13162

Upadhayaya, P. R. (2023). Information communication technology in education: Bringing innovation in classroom. *Ganeshman Darpan*, *8*(1), 96–110. https://doi.org/10.3126/gd.v8i1.57335

Vankúš, P. (2021). Influence of game-based learning in mathematics education on students' affective domain: A systematic review. *Mathematics*, *9*(9), Article 986. https://doi.org/10.3390/math9090986

Velmurugan, P. R., Swadhi, R., Varshney, K. R., Regins, J. C., & Gayathri, K. (2025). Creating engaging and personalized learning experiences in distance education: AI and learning analytics. In H. S. Mamede & A. Santos (Eds.), *AI and learning analytics in distance learning* (pp. 103–126). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-7195-4.ch005

Verbi. (2025). *MAXQDA* (Version 24) [Software]. https://www.maxqda.com/new-maxqda-24

Wong, J., Bayoumy, S., Freke, A., & Cabo, A. (2022). Augmented reality for learning mathematics: A pilot study with WebXR as an accessible tool. In *SEFI 50th annual conference of the European Society for Engineering Education: "Towards a new future in engineering education, new scenarios that european alliances of tech universities open up"* (pp. 1805–1814). Universitat Politècnica de Catalunya. https://doi.org/10.5821/conference-9788412322262.1216

Zhang, Y., Yang, X., Sun, X., & Kaiser, G. (2023). The reciprocal relationship among Chinese senior secondary students' intrinsic and extrinsic motivation and cognitive engagement in learning mathematics: A three-wave longitudinal study. *ZDM—Mathematics Education*, *55*(2), 399–412. https://doi.org/10.1007/s11858-022-01465-0


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Insights From an Umbrella Review of Flipped Learning in Higher Education

**Hebatullah ElGamal and Olaf Zawacki-Richter**
*Carl von Ossietzky Universität Oldenburg, Germany*

## Abstract

There is a noticeable growth in the number of systematic reviews published in open, distance, and digital education (ODDE), with a growing focus on flipped teaching and learning, particularly in higher education, emphasizing the need to consolidate evidence and findings under one comprehensive review. This umbrella review aims to thoroughly understand the current state of flipped learning in higher education and pinpoint research gaps, analyzing 23 systematic reviews published between 2018 and 2022 from three international databases: Web of Science, Education Source, and Scopus. It delves into publication and authorship patterns while synthesizing key insights. The thematic scope of the reviews reveals that many were focused on the effectiveness of flipped learning and teaching interventions, as well as learning design. The review explored theories guiding practice and research, instructional design considerations, and the application of flipped classrooms in various fields of study. It also examined the reported challenges of the flipped classroom model. As there are a scarcity of theoretical frameworks and a lack of detailed information on the pedagogical challenges of this model, recommendations are presented to enhance research and practice of flipped teaching and learning. The results of this umbrella review provide valuable insights to guide research in future and improve the quality of systematic reviews in the field of ODDE in general and flipped teaching and learning in particular.

*Keywords*: flipped learning, higher education, umbrella review, systematic review

# Introduction

According to Xiao (2023), open distance and digital education (ODDE) has gained widespread acceptance as a teaching and learning method across all levels of education. This has led to numerous reviews in ODDE to support evidence-based practice, as digital teaching methods are relatively new in mainstream education compared to the decades of research and practice in open and distance education (Zawacki-Richter & Naidu, 2016). In particular, there has been a significant surge in flipped classroom approach research, with several systematic reviews (SRs) recently published (e.g., Karabulut-Ilgu et al., 2018; Lo, 2020; Şensöz & Erdemir, 2022; ElGamal, 2022). This surge in research underscores the growing importance of the flipped (or inverted) classroom approach for educators, researchers, and practitioners, specifically in higher education (HE; Farmus et al., 2020).

The flipped classroom (FC) model is a student-centered teaching method. It prioritizes learners' needs by allowing them to learn and review course materials before attending a class. Bergman and Sams (2012) highlighted that this approach enables students to apply and reinforce their knowledge during interactive sessions. The model extends beyond videos and homework, aiming to create a learning environment focusing on the student's learning process. It provides meaningful activities and encourages higher-order thinking and problem-solving skills (Akçayır & Akçayır, 2018). The flexibility of this blended learning model allows for adaptation based on the teacher, technology, content, and digital resources used in different classrooms, ensuring that the needs of students are always at the forefront. Educators who embrace this model demonstrate a deep understanding and empathy for their students' learning needs (Bagley, 2020).

Flipped learning advocates for a constructivist approach (Eppard & Rochdi, 2017) to teaching and learning, in which the role of educators is not only necessary but crucial. To create a constructivist learning environment, it is essential to employ various teaching methods, including learning tools, group work, and active learning. However, the key to this approach lies in feedback and facilitation, given the fundamentally learner-centric nature of the FC approach (Hwang et al., 2019). These elements ensure that students construct their knowledge and understanding of the subject matter (see Jonassen et al., 1995). The goal of instruction is to provide guided experiences within the zone of proximal development to support learners' progress, making educators, researchers, and practitioners integral in this process. Innovative educational technology has enabled the flipped learning approach to empower teachers to seamlessly incorporate a variety of digital resources, such as online learning platforms, online discussion tools, and video-watching tools, into their teaching plans. The flipped learning method extends class time for active learning activities such as discussions and peer collaborations. The FC method also enables teachers to design pre-class and in-class content, activities, and assignments tailored to their students' unique intellectual, physical, and cultural needs (Pulley, 2014).

Previous discussions about the FC have mainly focused on its impact on academic performance, as demonstrated by studies such as Park et al. (2021) and Arslan (2020). However, it is worthwhile to recognize that this approach has the potential to support a variety of 21st-century learning goals beyond just academic success (Zhou & Li, 2018). These goals generally refer to higher-order thinking skills falling into three categories: information and communication skills, thinking and problem-solving skills, and interpersonal and self-directional skills (Sahin, 2009).

The number of SRs on the flipped approach has been increasing exponentially (Lo, 2020). Staying informed about the latest findings can take time and effort. Previous SRs (e.g., Akçayır &Akçayır, 2018;

Lundin et al., 2018) have found that research on the FC is widely dispersed. Different perspectives in each review may lead to varying results and new conclusions due to the selection of various studies. Each review has its unique focus, scope, and level of comprehensiveness. For instance, some reviews (e.g., Karabulut-Ilgu et al., 2018) have examined the effect of flipped learning on teaching and learning, while others have provided a comprehensive overview of research spanning a significant period (e.g., Birgili et al., 2021; ElGamal, 2022), indicating the need for a more comprehensive review of the evidence on this method (Kapur et al., 2022). Many different names refer to syntheses of existing SRs, one of which is an *umbrella review* (Aromataris et al., 2014) or a *review of systematic reviews*, according to The Joanna Briggs Institute (2014). In a recent umbrella mapping review covering 576 SRs in the field of ODDE published between 2018 and 2022 (Zawacki-Richter et al., 2025), flipped learning was one of the most popular research topics. Based on their content analysis of SRs in ODDE, the authors concluded:

> Leading topics [in ODDE systematic reviews] are AI in education, game-based learning, virtual and augmented reality, blended learning, computer-assisted language learning, flipped learning, digital competencies of learners and teachers, OER, and MOOCs. They account for almost half of the reviews included. These subject areas offer a critical mass of studies which, based on this mapping, would warrant in-depth synthesis of the findings in follow-up umbrella reviews. (p. 19)

This umbrella review delivers a detailed analysis and synthesis of SRs on flipped learning based on Zawacki-Richter et al.'s umbrella mapping review (2025). The aim was to explore the FC approach from new perspectives beyond academic achievements and student views. The review also addresses research gaps associated with the FC method. By compiling and analyzing findings from flipped learning reviews, this umbrella review demonstrates how flipped-related research has evolved and exposes the interconnections needed among different disciplines to advance the field. It synthesizes SR data, highlighting information about flipped research's progress and pattern, review types used, and topics covered by the SRs. This approach aims to identify topics related to underexplored flipped learning and provide new insights for future research. This can be particularly useful for readers, aspiring researchers, and policymakers who need well-informed decisions. The study also explored the theoretical models guiding the flipped approach, outlining design considerations relevant to the FC in higher education, and differentiating the various implementation approaches of the FC model across different subject areas, providing perceptions of the everyday challenges encountered. This review addressed the following questions:

RQ1: What are the authorship and publication patterns in systematic reviews related to flipped classrooms?

RQ2: What is the (main) research focus of systematic reviews on flipped classrooms?

RQ3: What theories guide practice and research on flipped learning and teaching?

RQ4: What are the flipped classroom's instructional design considerations?

RQ5: What distinguishes the application of flipped classrooms in various fields of study?

RQ6: What are the reported challenges of the flipped classroom?

# Method

A systematic review aims to answer specific questions using an explicit, systematic, and replicable search strategy (Gough et al., 2017; Zawacki-Richter et al., 2020). Umbrella reviews follow a procedure similar to SRs (Aromataris et al., 2014). Therefore, the best approach to conducting an umbrella review is to adhere to the steps in the systematic process (Lee et al., 2022): (a) formulate the review question and protocols, (b) define inclusion and exclusion criteria, (c) create the search strategy and identifying information sources, (d) screen the articles based on the inclusion/exclusion criteria, (e) present the results of the search strategy in a flowchart, (f) extract relevant descriptive data from the included studies, (g) assess the quality of the included studies, and (h) synthesize the evidence collected (Zawacki-Richter et al., 2020). This systematic and transparent process helps ensure that a SR is reproducible. Grant and Booth (2009) defined an umbrella review as compiling evidence from multiple reviews into one accessible document. The focus is on a broad condition or problem with competing interventions, highlighting reviews addressing these interventions and their outcomes. As such, the umbrella review aims to tackle the issue of heterogeneity and provide a summary of the existing evidence and an in-depth synthesis.

## Search Strategy

Building upon the umbrella mapping review conducted by Zawacki-Richter and colleagues published in 2025, this umbrella review extended the same search protocol. This protocol, which included a RIS-file containing bibliographic information for all 576 articles, had been registered and published on the Open Science Framework (OSF) platform (URL_1). Zawacki-Richter et al., (2025) developed an extensive search string (see Table 1) to retrieve SRs on open, distance, and digital education (ODDE). The eligibility criteria (see Table 2) pertained to five areas for inclusion and exclusion: publication year, language, type, education level, and research methodology. Articles were collected from three international databases: Education Source, Scopus, and Web of Science. Despite concerns about peer-review processes in the scientific community (see Smith, 2006), only articles published in peer-reviewed journals were included due to their general trustworthiness in academia and rigorous review processes (Lee et al., 2022; Nicholas et al., 2015). To manage the immense growth in the number of SRs, the search was limited to 2018 to 2022, covering the most recent 5 years at the time of the study. As English is the common working language, the search was limited to publications in English. The targeted studies were those in the field of ODDE that explicitly stated in their titles or abstracts that their methodological approach was a systematic review. A subset of this included retrievals, accounting for 23 SR studies, focused on the FL method, and these were the review studies selected for synthesis.

**Table 1**

*Search String Used in This Study*

| Topic | Search terms |
|---|---|
| Context | (distan* OR online OR open OR technology-enhanc* OR digital) W/3 (educat* OR learn* OR teach*) |
| AND | |
| Review type | systematic W/2 review |

## Search Procedure

The initial search by Zawacki-Richter et al. (2025) across the above-specified databases for titles, abstracts, and keywords occurred in November 2022. It was updated in January 2023 to encompass all records from 2022. Initially, 4,449 records were identified and imported into reference management software (See Figure 1). Following the removal of 1,359 duplicates, a basic Python code was created to filter out studies that did not include the term "systematic" in their title or abstract. A total of 474 papers were excluded using the code and were subsequently manually reviewed to confirm that no papers were mistakenly rejected. The titles and abstracts of the remaining 2,616 articles were screened based on the inclusion and exclusion criteria outlined in Table 2. During the initial selection process for titles and abstracts, our priority was to maximize sensitivity rather than specificity (i.e., to include rather than exclude; Zawacki-Richter et al., 2020).

**Table 2**

*Inclusion and Exclusion Criteria Applied to This Systematic Review Search*

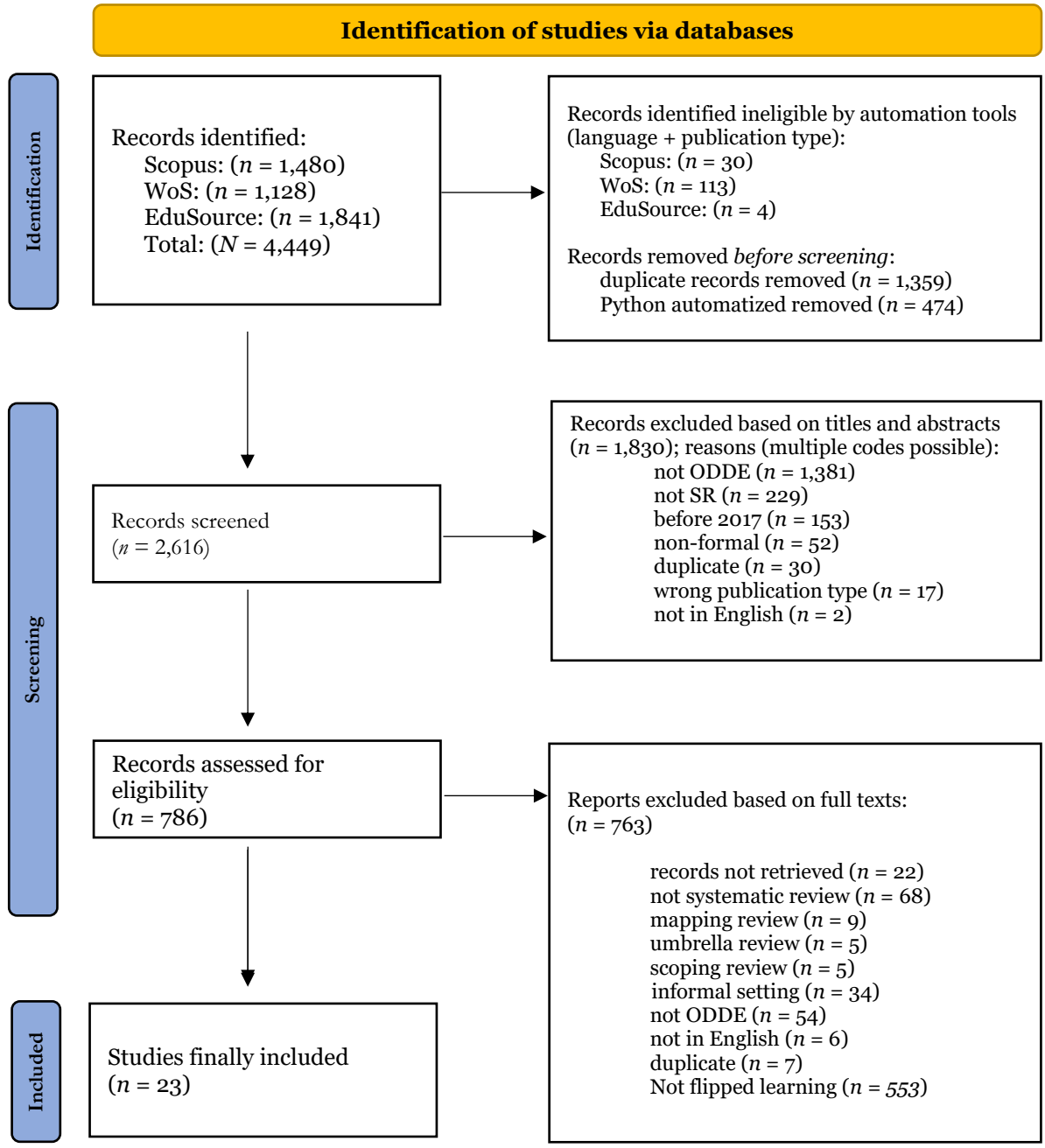
| Criteria | Inclusion | Exclusion |
|---|---|---|
| Publication year | 2018–2022 | before 2018 |
| Language | English | Not in English |
| Education level | Any level in ODDE, including K–12, HE, LLL, TVET | Not ODDE, informal, non-formal |
| Methodology | SR[1] | Not SR |
| Publication type | Peer-reviewed academic journal article indexed in Scopus, WoS, or Education Source | Not journal article (e.g., book, editorial, note) |

*Note.* ODDE = open, distance, and digital education; HE = higher education; LLL = lifelong learning; TVET = technical and vocational education and training; SR = systematic review; WoS = Web of Science.

[1] Papers that claim to conduct a systematic review in the title or abstract were all included.

**Figure 1**

*PRISMA Diagram Showing Identification, Screening, and Inclusion Process of This Study*




*Note.* PRISMA = preferred reporting items for systematic reviews and meta-analyses; WoS = Web of Science; ODDE = open, distance, and digital education; SR = systematic review. Adapted from "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," by D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and PRISMA Group, 2009, *PLoS Med*, *6*(7), p. 8 (https://doi.org/10.1371/journal.pmed.1000097). CC BY 4.0.

*Inter-rater reliability*. Inter-rater reliability among coders was evaluated using Fleiss' kappa (κ). The coding consistency for including or excluding the 60 articles among the three coders was κ = .77, indicating good agreement.

## Data Analysis

From the Zawacki-Richter et al. (2025) umbrella mapping review, 23 SRs dealing with FL were included for further examination and synthesis out of 576 reviews. We used descriptive and analytical techniques to address the review questions and analyze these SRs. Initially, we established a protocol to classify the general authorship characteristics, including review type, number of included studies, educational setting, country, journal, and discipline. The data was organized using an electronic data extraction table in Microsoft Excel.

As a method for researchers to construct observations of others by expressing their views, an analytical strategy was deemed relevant to the analysis and synthesis of the findings in this umbrella review. The review included studies using different theories and empirical approaches (quantitative and qualitative), so the analysis was configurational and exploratory. The procedure used thematic and content analysis to identify, compare, and report valuable findings to education professionals. Through thorough reading and coding, themes and sub-themes were identified (Braun & Clarke, 2006) using QDA Lite software (https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/). Each included SR was used as a unit of analysis. This approach was deemed suitable for summarizing and synthesizing the reviews' findings.

### *Limitations of the Data*

The review was limited to studies published between 2018 and 2022 and only included peer-reviewed studies in English. Other languages were excluded. The review focused only on FL rather than macro strategies, such as blended learning with FL as a subtopic.

# Results and Synthesis

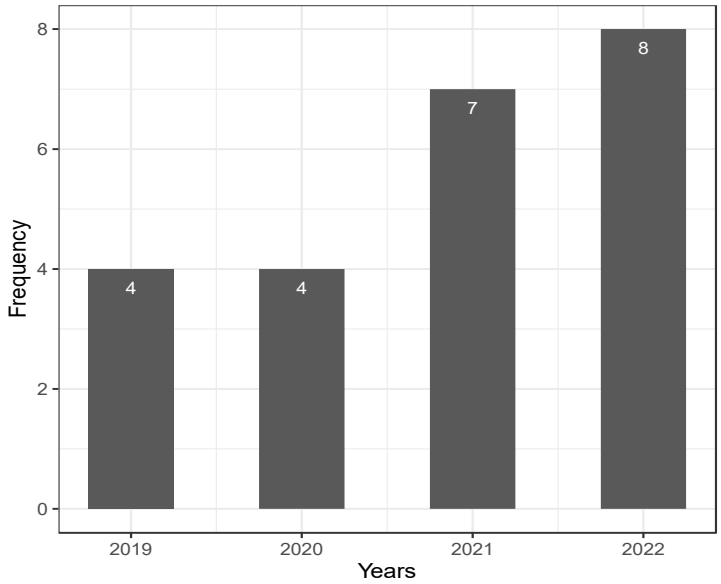
## RQ 1: Publication and Authorship Patterns

### *Published Systematic Reviews per Year*

The number of papers published from 2019 onwards significantly increased. The number of included SRs on flipped learning per year grew steadily from four in 2019 and 2020 to eight in 2022 (see Figure 2).

**Figure 2**

*Number of Included Articles Per Year (N = 23)*




## *Number of Authors*

Systematic reviews are time- and labour-intensive. On average, it takes 67 weeks to conduct and publish a SR (Borah et al., 2017). It is therefore recommended to work in a team and divide the steps in a SR, especially the process of screening and coding, to achieve valid and reliable results. This is also apparent in the SRs included here: the median number of authors per paper is three. One SR was published by a team of nine authors (Xu et al., 2019).

## *Journals*

The 23 SRs included in the corpus were published in 22 different journals. Two SRs on flipped learning in higher education were published in the journal *Nurse Education Today* (see Table 3). The majority of SRs on FC ($n$ = 8) were published in educational technology and open, distance, and digital education (ODDE) journals such as the *International Journal of Educational Technology in Higher Education*, the *Turkish Online Journal of Distance Education,* and the journal of *Computer Assisted Language Learning*.

It stands out that six journals came from the field of medicine and health sciences (e.g., *Nurse Education Today*, *BMC Medical Education*, and the *Journal of Dental Education*), which indicates that the SR methodology is frequently applied in these fields and underscores the fact that it has its roots here (Hammersley, 2020).

**Table 3**

*Number of Included Articles by Journal*

| Journal | *n* |
|---|---|
| *Nurse Education Today* | 2 |
| *BMC Medical Education* | 1 |
| *BMC Nursing* | 1 |
| *Computer Assisted Language Learning* | 1 |
| *Education & Information Technologies* | 1 |
| *Education Sciences* | 1 |
| *Educational Sciences: Theory and Practice* | 1 |
| *Eurasia Journal of Mathematics, Science and Technology Education* | 1 |
| *International Journal of Educational Technology in Higher Education* | 1 |
| *International Journal of Emerging Technologies in Learning* | 1 |
| *International Journal of Information and Education Technology* | 1 |
| *Journal of Dental Education* | 1 |
| *Journal of Language and Linguistic Studies* | 1 |
| *Journal of Professional Nursing* | 1 |
| *Journal of Statistics Education* | 1 |
| *Mathematics* | 1 |
| *Nurse Education in Practice* | 1 |
| *Nurse Educator* | 1 |
| *South African Journal of Higher Education* | 1 |
| *Sustainability (Switzerland)* | 1 |
| *Turkish Online Journal of Distance Education* | 1 |
| *ZDM—Mathematics Education* | 1 |
| Total | 23 |

## *Countries*

The first author's country of origin was considered to analyze the country-wise distribution of articles ($n$ = 14 countries). Almost 50% of the SRs on flipped learning come from only three countries: China, Canada, and Turkey.

**Table 4**

*Distribution of Systematic Review Articles by Country and Cumulative Percentages*

| Rank | Country | *n* | Cum % |
|---|---|---|---|
| 1 | China | 5 | 21.7 |
| 2 | Canada | 3 | 34.8 |
| 2 | Turkey | 3 | 47.8 |
| 3 | Germany | 2 | 56.5 |
| | New Zealand | 2 | 65.2 |
| 4 | Croatia | 1 | 69.6 |
| | Indonesia | 1 | 73.9 |
| | Iran | 1 | 78.3 |
| | Korea | 1 | 82.6 |
| | Malaysia | 1 | 87.0 |
| | South Africa | 1 | 91.3 |
| | Spain | 1 | 95.7 |
| | USA | 1 | 100.0 |
| | Total | 23 | |

## *Reviews' Mapping: Types, Number of Included Studies, Educational Setting*

Research reviews are essential for making academic and non-academic decisions, as they offer different perspectives (Rees & Oliver, 2012). When conducting a review, it is important to consider the research review question, assumptions, and whether the results are aggregated or configured (Gough et al., 2012). Regardless of the approach taken, various methods can synthesize the information. While the difference between quantitative and qualitative research may not always be clear-cut, it's important to note that the aggregate/configure framework has a more specific meaning regarding the synthesis logic.

According to Gough et al. (2012), many reviews contain aggregation and configuration elements. Some focus on combining and adding up findings from similar studies, while others aim to organize the findings. Aggregative reviews concentrate on specific questions using pre-specified quantitative methods to test theories using empirical observations, which is a deductive method. Reviews with a configurative approach are more likely to ask open questions, answered with qualitative data and iterative methods. They interpret specific examples to address questions about experiences and meaning.

Our review found 21 configurative reviews and two reviews that included elements from both configurative and aggregative types. While all the reviews addressed FCs in higher education, only three addressed FCs in higher education and K–12 settings. Table 5 presents a summary of these review studies, organized by review type, number of studies included, and educational setting.

**Table 5**

*Studies Included in the Umbrella Review by Type, Number, and Setting*

| Author | Year | Review type | Studies, ***n*** | Educational setting |
|---|---|---|---|---|
| Arslan, A. | 2020 | C | 78 | K–12, HE |
| Baltaci | 2022 | C | 20 | HE |
| Banks & Kay | 2022 | C | 12 | HE |
| Cevikbas & Kaiser | 2022a | C | 97 | HE |
| Cevikbas & Kaiser | 2022b | C | 41 | HE |
| Divjak et al. | 2022 | C | 18 | HE |
| Ekici | 2021 | C | 22 | K–12, HE |
| Evans et al. | 2019 | C | 24 | HE |
| Farmus et al. | 2020 | C and A | 10 | HE |
| Fernández-Martín et al. | 2020 | C | 10 | HE |
| Fung et al. | 2021 | C | 12 | K–12, HE |
| Gerber & Eybers | 2021 | C | 167 | HE |
| Gianoni - Capenakas | 2019 | C | 8 | HE |
| Hendrik and Hamzah | 2021 | C | 32 | K–12, HE |
| Kazeminia et al. | 2022 | C | 26 | HE |
| Koh | 2019 | C | 51 | HE |
| Lo & Hew | 2022 | C | 33 | HE |
| Özbay & Cinar | 2021 | C | 24 | HE |
| Park et al. | 2021 | C | 10 | HE |
| Rasheed et al. | 2020 | C | 14 | not mentioned |
| Xu et al. | 2019 | C and A | 22 | HE |
| Youhasan et al. | 2021 | C | 27 | HE |
| Zou et al. | 2020 | C | 34 | HE |

*Note*. C = configurative; A = aggregative; HE = higher education.

## RQ2: Thematic Focus of Reviews

We used the 3M Framework (Zawacki-Richter, 2009; Zawacki-richter & Bozkurt, 2023) framework to classify reviews into macro-level research (systems, theory, methods, and global perspectives), meso-level research (management and organization of institutions), and micro-level research (teaching, learning, individual learners, and teachers).

The analysis revealed an interesting result: 91.3% of the reviews focused on micro-level research, categorized into three subtopics (see Table 6). However, none of the reviews addressed the macro- or meso-level, and othersdid not specify the focus.

At the micro level, three main research areas were identified: (a) evaluating the effectiveness of the FC on academic performance and attitudes, (b) examining its design and related positive outcomes, and (c) identifying specific factors that impact its effectiveness. The reviews concluded that the FCs positively impacted learning compared to traditional learning. Most review studies focused on academic performance, satisfaction, engagement, and motivation. Few reviews ($n$ = 8) examined the pedagogical and instructional design of the FC approach focusing on its role in enhancing the effectiveness of both learning and teaching. For example, Park et al. (2021) critically evaluated the effectiveness of pedagogical strategies for implementing FL in health professions education in South Korea. Only one review, Koh (2019), focused solely on the pedagogical design of flipped classrooms. In addition, the reviews looked at influential factors impacting the positive effects of FC on learning. These factors included subject area, duration of implementation, and the use of specific technologies to aid delivery and facilitate student learning. The results showed that the effectiveness of flipped learning varied across different subject areas.

Distinctively, few reviews within the three subcategories reflected on the implementation of the FC during the COVID-19 pandemic (e.g., Cevikbas & Kaiser, 2022a; Divjak et al., 2022; Gerber & Eybers, 2021; Lo & Hew, 2022).

**Table 6**

*Topical Overview of the Reviews at the Micro Level (n = 21)*

| Topic | Description | *n* | Reviews | Sample and focus |
|---|---|---|---|---|
| FC's outcomes | Academic performance, such as test scores, course grades, and knowledge scores<br>Attitude (satisfaction, engagement, motivation, etc.)<br>Both | 6 | Baltaci, 2022; Banks & Kay, 2022; Fernández-Martín et al., 2020; Gianoni-Capenakas et al., 2019; Özbay & Cinar, 2021; Xu et al., 2019. | Baltaci (2022): benefits and challenges of FL instruction on student achievement and attitudes in Turkey. |
| FC's design | Pedagogical and instruction design features<br>Design features that enhance the outcomes | 8 | Arslan, 2020; Cevikbas & Kaiser, 2022a; Cevikbas & Kaiser, 2022b; Hendrik & Hamzah, 2021; Kazeminia et al., 2022; Koh, 2019; Park et al. 2021; Youhasan, 2021. | Cevikbas & Kaiser (2022a): using innovative technology to personalize FC instruction and improve learning outcomes. |

| Topic | Description | *n* | Reviews | Sample and focus |
|---|---|---|---|---|
| Unique factors enhancing experiences in FCs | Affordance to a specific discipline<br>Scope/mode of implementation (e.g., online FCM) | 7 | Ekici (2021); Evans et al., 2019; Farmus et al., 2020; Fung et al., 2021; Gerber and Eybers, 2021; Lo & Hew, 2022; Rasheed et al., 2020. | Ekici (2021): effectiveness of gamification on FL implementation. |

*Note*. FC = flipped classroom; FL = flipped learning; FCM = flipped classroom model.

## RQ3: Theoretical Foundations

It has been said that "nothing is as practical as a good theory" (Lewin, 1952, p. 169). One approach to differentiate research is by assessing the degree to which it is focused on constructing, examining, or validating theory. Limited evidence supports the positive impact of implementing the FC on students' academic performance (Bishop & Verleger, 2013; Lo & Hew, 2017; Zuber, 2016). This limitation stems from the lack of employment of theoretical frameworks and inconsistency among methods and class activities (Lin & Hwang, 2019; Zuber, 2016). Thus, we assessed how the reviews discussed the integration of theoretical frameworks in the studies they examined, whether to analyze their research findings or guide the implementation of the FC method.

Regrettably, out of the 23 SRs, only seven discussed the theoretical frameworks reported in the studies they reviewed. Five focused on theoretical frameworks guiding flipped classrooms' related practices and implementation approaches, stating the theoretical framework employed for the FC design. The other two incorporated theoretical frameworks to synthesize research findings and assess the impact of the FC practice. The lack of discussion on theory in the FC literature raises concerns.

The umbrella review's coding and analysis were based on the information provided in the included SRs. Although some SRs discussed theoretical frameworks, most did not give the number of studies that used a theoretical framework. For instance, a SR by Baltaci (2022) revealed that only seven out of 20 reviewed articles explicitly mentioned their theoretical frameworks. This type of information was uncommon among the reviews, resulting in limited capacity for data aggregation regarding theoretical frameworks.

The theories that have been reviewed encompassed various approaches for understanding the specific phenomena exhibited in Table 7. Within the field of FC research, the dominant epistemological theory was constructivism. However, studies in this area have employed a range of theoretical approaches, with some emphasizing Bloom's taxonomy for FC design. Other theoretical frameworks included active learning, cognitive load, and self-determination theories. Surprisingly, the theory of differentiated instruction was not central to discussions about the FC model, although it can significantly support teachers in optimizing the student-centered approach.

**Table 7**

*Overview of the Reviewed Theories and Models*

| Author(s) | Theories and models |
|---|---|
| Baltaci (2022) | Bloom's taxonomy<br>Constructivism<br>Self-directed or self-regulated learning |
| Ekici (2021) | Gamified learning<br>Self-determination<br>Cognitive load theory<br>Flow theory |
| Gerber & Eybers (2021) | Gamified learning<br>Self-determination<br>Cognitive load theory |
| Koh (2019) | Self-directed learning |
| Lo & Hew (2022) | Revised Community of Practice (RCOI) |
| Park et al. (2021) | Analysis, design, development, implementation, and evaluation (ADDIE) model |
| Zou et al. (2020) | Self-regulated learning |

In this section, we present examples of SRs that discuss the design of FC learning along with the associated theoretical framework. Gerber and Eybers (2021) thoroughly examined educational theories, including constructivist theory, active learning theory, self-determination theory, cognitive load theory, and the theory of gamified learning. These foundational theories played an influential role in shaping the design considerations derived from their findings. By integrating these theories, the researchers could outline essential design considerations for implementing flipped classrooms. Gerber and Eybers (2021) and Eciki (2021) also emphasized the self-determination theory, which focuses on human motivation when the needs for competence, connection, and autonomy are met. They also elaborated on the cognitive load theory, suggesting that learning experiences should minimize working memory load. The flow theory was also highlighted, emphasizing total concentration achieved when individuals balance perceived challenges and skills. Furthermore, the theory of gamified learning was identified, focusing on using mechanisms to promote student engagement and motivation, which has gained popularity in various contexts for encouraging specific behaviours. Meanwhile, in Baltaci (2022), it was highlighted that Bloom's taxonomy was the most commonly used framework for designing and implementing flipped instruction. This finding is consistent with earlier reviews (e.g., Bishop & Verleger, 2013) that have supported student-centered learning and its associated theories, such as peer-assisted and collaborative learning. Similarly, Park et al. (2021) introduced the analysis, design, development, implementation, and evaluation (ADDIE) model for flipped learning, advocating for pre-learning, in-class activities, and post-class evaluation and self-assessment as the principal teaching strategies.

On the other hand, Lo and Hew (2022) identified various practical elements for online FL practice. They organized them based on the revised community of inquiry (RCOI) framework's components: cognitive, social, teaching, and learner presence. At the same time, Zou et al. (2020) thoroughly reviewed 34

published articles on flipped classrooms in language teaching, considering various aspects, including theoretical foundations.

In summary, the results in this section help map the reviewed theoretical frameworks and suggest the theoretical base to steer insights about the pedagogical offerings of FCs that could enhance the learning outcomes. For instance, since the FC design is predominantly based on social constructivism, almost all literature reports a positive impact on cognitive learning outcomes, regardless of the underlying theoretical framework in designing the FC. However, regarding the effects of other learning outcomes, SRs that include cognitive load theory (e.g., Baltaci, 2022; Ekici, 2021) reported a positive effect on student involvement, enthusiasm, and motivation. On the other hand, SRs with active learning and gamified learning theoretical frameworks (e.g., Gerber & Eybers, 2021) reported a positive impact on student engagement and effective use of study time.

## RQ4: Learning Design Considerations for Flipped Classrooms

A standard operational FC model was identified in the reviewed literature, comprising three main components (Lo, 2023; Youhasan et al., 2021): pre-class, in-class, and post-class activities. While not all reviews focused solely on the design of the flipped classroom, 18 of them touched on general and specific design features and pedagogical considerations, and almost all 23 reviews described their active learning components. Several reviews explored effective design frameworks for this method (e.g., Arslan, 2020; Gerber & Eybers, 2021; Youhasan et al., 2021), and they found a number of common features. For example, the FC model involves instructional design elements and delivery tools. The primary teaching methods of the FC involve active participation, interaction, collaboration, projects, and case-based learning. Instructional design elements for flipping a course include designing content, organizing flexible physical space, creating activities to orient students, motivating an active student mindset, scaffolding, and chunking course content (Arslan, 2020). Popular delivery tools include video lectures for pre-class courses, while the learning management system (LMS) provides a platform for sharing resources and data analytics. Popular post-class strategies include instructor feedback, self-evaluation journals, and peer evaluation.

### *Learning Design Elements and Delivery Tools*

We used the four FLIP pillars—flexible environment, learning culture, intentional content, and professional educator—developed by the Flipped Learning Network in 2014 to classify the design principles and considerations (refer to Table 8) identified in the reviewed SRs. Several design recommendations stand out in specific studies. For example, Gerber and Eybers (2021) suggested breaking the FC model into mini-curricula to prevent student overwhelm and facilitate procedural fluency. In addition, Gianoni-Capenakas et al. (2019) recommended organizing materials through content chunking, which involves direct instruction and active learning strategies.

Another interesting finding by Divjak et al. (2022) was the use of a holistic learning design, aligning teaching and learning strategies with the desired learning outcomes. This should involve a combination of approaches that best match the intended learning outcomes and the needs and characteristics of the student body. Combining the FC with other innovative teaching and learning approaches such as work-based learning (WBL), problem-based learning (PBL), the use of massive open online courses (MOOCs), and game-based learning can have a more significant impact on student learning and satisfaction.

Similarly, Fung et al. (2021) examined the design and planning mechanisms of course materials and activities that assist students in overcoming their blind spots, organizing content materials, and expanding their problem-solving approaches. Instant feedback and scaffolding were also key elements extensively reviewed in most studies. Park et al. (2021) also emphasized using a self-evaluation journal to follow up on pre-class activities.

On the other hand, Ekici (2021) assessed how engagement can be fostered through a gamified FC and discovered that half the studies mentioned the use of teaching aids such as Kahoot, Socrative, and quizzes to gamify instruction.

**Table 8**

*Overview of the Design Principles and Considerations for Flipped Classrooms*

| Category | Subcategory | *n* | Sample review |
|---|---|---|---|
| Flexible environment | Learners' space and pace | 10 | Gianoni-Capenakas et al. (2019) |
| | Duration of the intervention | 1 | Gianoni-Capenakas et al. (2019) |
| | Periodic adjustments of class design | 2 | Gerber & Eybers (2021) |
| Learning culture | Active participation | 22 | Youhasan et al.(2021) |
| | Self-evaluation | 4 | Park et al. (2021) |
| | Learner-centered engaging activities | 15 | Koh (2019) |
| | Personalized learning | 2 | Cevikbas & Kaiser (2022b) |
| Intentional content | Holistic design | 2 | Divjak et al. (2022) |
| | Chunking content | 11 | Fung et al. (2021) |
| | Design of visual material | 18 | Inayah et al. (2023) |
| | Gamified teaching aids | 18 | Ekici (2021) |
| Professional educator | Scaffolding and feedback | 17 | Baltaci (2022) |
| | Monitoring and data recording | 2 | Gianoni-Capenakas et al. (2019) |

## *Descriptive Design Factors*

Understanding a FC intervention involves several factors, from design to implementation. Lo and Hwang (2018) stressed the need for well-documented studies to enhance our understanding of the flipped learning approach. They developed a descriptive framework comprising four components: research background, course design, course activities, and outcome of interventions. While these four components are important, other factors also need consideration.

Unfortunately, out of the analyzed SRs, only four explicitly reported the duration of the flipped intervention, while others mentioned durations as "long" or "short." Similarly, only six SRs reported on the sample size, while others referred to "small samples" without defining them. Moreover, the implementation strategy was generally inferred rather than explicitly explained. There were two main approaches: a partial FC, where the flipped approach was used for only several sessions, and an entire FC, implemented for all or most of the meetings, that was slightly more dominant (see Hendrik & Hamzah, 2021). Evans et al. (2019) also noted variation in the implementation duration, with some

studies using the approach for just one unit in a semester, while others adopted it for an entire semester or longer. This inconsistency in reporting on the pedagogical elements of the FC has led to confusion.

Table 9 presents the duration of the flipped interventions mentioned in four SRs.

**Table 9**

*Duration of Flipped Interventions as Described in Four Systematic Reviews*

| Systematic review | Studies reviewed (*N*) | Studies describing duration (*n*) | Duration |
|---|---|---|---|
| Ekici (2021) | 22 | 17 | 1–4 months |
| | | 3 | < 1 month |
| | | 2 | = 6 months |
| Park et al. (2021) | 10 | 10 | 1–4 months |
| Evans et al. (2019) | 24 | Not specified | 1 unit of content/semester |
| | | Not specified | 1 academic semester |
| Gianoni-Capenakas et al. (2019) | 8 | Not specified | 5–16 sessions |

*Note*. The variation in reporting duration among the included studies indicates the need for standardized measurement units.

The variability in reporting may be linked to the diverse research objectives of the SRs and their studies, suggesting that our findings should only be considered indicative. However, given heterogeneity and risk of bias, future studies should include a large sample size and high-quality research to confirm effectiveness (Xu et al., 2019). Therefore, ithis review recommends that research methods be thoroughly documented and high quality data collected, as Banks and Kay (2022) advocated. To improve data reporting in SRs related to flipped learning, the following factors would need to be included in all reviews:

- Context (educational setting/country)
- Subject area (e.g., nursing)
- Sample size
- Duration of intervention (e.g., semester)
- Approach of the intervention (partial or full), i.e., a blend of traditional and flipped classes or all classes entirely flipped.

## RQ5: The Flipped Classroom Across Disciplines

The impact of the FC can vary depending on the subject area (Giannakos et al., 2014). For example, Lin and Hwang (2019) explored the research trends regarding the use of FCs in medical courses and found that using in-class activities was inconsistent across the studies. In FCs for medical courses, discussions,

exercises, problem-based activities, and group project activities reign as the most popular in-class activities (Lin & Hwang, 2019). This indicates the significance of developing customized flipped learning designs and strategies that address the unique requirements of each subject area and, ultimately, foster a student-centered learning environment. For example, the SR by Park et al. (2021) highlighted the importance of adopting tailored approaches to flip the traditional classroom teaching model for specific courses to centre the learner in the educational environment.

After conducting an in-depth analysis of multiple SRs, it has been revealed that the FC model has been applied in various disciplines. Studies in STEM subject areas (science, technology, engineering, and mathematics), which included computing, programming, mathematics education, and information systems, were the most common areas where FL practices were conducted ($n$ = 8). The application of the FL method in healthcare education was also found to be quite prevalent with ($n$ = 7) reviews on medical education, dental education, and nursing being the primary focus areas. Language teaching ($n$ = 3) was another area where flipped learning methodology was frequently implemented. Several reviews have proposed design recommendations for effectively implementing the FC model in specific fields. We have included the findings related to those three fields below.

## *STEM Disciplines*

Systematic reviews highlighting specific design recommendations were found in some studies ($n$ = 8). For instance, the SR by Farmus et al. (2020) found that "partial flipping" may be sufficient for STEM courses such as statistics. The approach would involve breaking the content into smaller parts and facilitating applied and advanced learning during in-class discussions. However, their research also found that instructors who already incorporated active learning components into their lectures or included lab time may find the FC model innovative. Thus, they recommended flipping only the most challenging parts of the content to ease the burden of flipping a class.

Another SR by Gerber and Eybers (2021) emphasized using case-based scenarios in STEM in-class activities through which the topic content would be discussed and applied with various examples to illustrate concepts, and learners would be expected to contribute to those discussions. The research recommended post-class activities that reinforce learning through application assignments that specifically focus on aspects of the advanced in-class concepts and activities, which is part of the module's summative assessment.

Along the same line, in a recent SR of mathematics education, Cevikbas and Kaiser (2022a) discovered that most FC instructors favoured the use of lecture explanatory videos as their primary instructional method. Nevertheless, alternative methods, such as reading texts and presentations, podcasts, lecture notes, and infographics, are also viable options. The review affirmed the widespread use of lecture videos in flipped mathematics classes, particularly those created by the instructors, and over one-third of the reviewed studies in this SR indicated that the FC pedagogy improved students' conceptual understanding and facilitated active and continuous learning. These findings highlight the significant positive impact of the FC on students' academic development in mathematics, with videos and quizzes emerging as the most effective elements of implementation.

Another SR conducted by Cevikbas and Kaiser (2022b) highlighted that the FC model was especially effective for tailoring engineering education to individual needs and for creating adaptive learning experiences. Various digital tools and platforms, as well as learning analytics, were used to personalize the instruction. Furthermore, the review highlighted the potential of gamification and emerging

technologies, such as virtual reality (VR), machine learning, and cloud technology, in effectively personalizing instruction, especially in STEM disciplines.

On the one hand, while popular activities such as readings, online modules, self-guided exercises, assignments, and interactive tutorials can be done outside the classroom, Hendrik and Hamza (2021) have expressed concerns about the overreliance on videos for knowledge transfer in a FC. Their study suggested that readings and assignments should be used to help students develop lifelong learning skills. They stress the importance of this approach, especially in software development, where developers need to be proficient in reading e-books, manuals, and documentation related to rapidly changing technology.

Finally, other SRs delved into organizational recommendations that have the potential to yield positive outcomes. Fung et al. (2021), for example, found among their included studies that distributing videos earlier, rather than scheduling them to be accessible only on the experiment date, increased student motivation.

### *Medical and Health Sciences*

Generally, systematic reviews in medical and health sciences (n = 7) viewed the FC method as an effective method. The review by Youhasan et al. (2021) demonstrated that the FC learning approach enhanced nursing students' knowledge, skills, and attitudes. The success was attributed to the students' unrestricted access to pre-classroom materials and the interactive in-class environment. Effectively implementing flipped pedagogy requires following an instructional system design at both macro (curriculum and course level) and micro (topics or sessions) levels. Additionally, the success of the FC approach depends on consistent and stable implementation planning (Liu et al., 2024). Youhasan et al. (2021) also found that four main in-class activities optimized outcomes: dividing students into small groups, conducting diagnostic assessments, micro-teaching, and continuing integrative student-centered instructions.

Still, in the nursing field, one interesting recommendation of the SR conducted by Evans et al. (2019) was that the detailed processes used to train faculty for consistent content delivery impacted the quality of teaching in flipped nursing classes. As per research conducted by Divjak et al. (2022), the FC combined with other approaches, such as problem-based learning (PBL), game-based learning, bring your own device (BYOD), and massive open online courses (MOOCs), was shown to be successful in specific study fields, particularly in medical education. Additionally, the review highlighted that in one of the included studies, teachers expressed the opinion that the online FC would be suitable for theory-based lessons but challenging for the practical lab sessions in medical and science disciplines.

Lastly, multiple SRs (e.g., Kazeminia et al., 2022; Park et al., 2021) have documented using a fundamental motivation model in flipped classrooms for theoretical and practical learning in various medical courses. An exciting feature of this model involves prompting students to engage in a WSQ approach (watching, summarizing, and questioning) for pre-class assignments to enhance their comprehension of the theoretical material.

### Language Education

In the language education discipline, SRs (n = 3) referred to specific benefits of the FC method. Based on a SR by Arslan (2020), more than half the studies suggested that using flipped learning in teaching English as a Foreign Language (EFL) or English as a Second Language (ESL) effectively enhanced

students' language skills. Several studies emphasized the role of flipped learning in improving vocabulary, writing, listening, reading, and speaking skills. Researchers also argued that adopting a flexible learning approach that promotes a dynamic learning environment for students would be highly recommended in language teaching. This approach would allow students to thoroughly understand the content by enabling them to replay videos and review activities until they fully grasp the concepts. Flipped learning both enhances students' English language proficiency through video activities and provides immediate feedback and promotes positive collaboration.

Baltaci (2022) elaborated on the importance of feedback through his review, explaining that instant assessment tools providing timely and constructive feedback can significantly enhance learners' writing skills. Such tools can be highly beneficial in language education as they enable students to promptly identify their weaknesses and areas for improvement, allowing them to focus on those areas and make progress quickly.

## RQ6: The Challenges of the Flipped Classroom Model

According to Crompton et al. (2017), failed outcomes with the FC model are useful for understanding the situation better. Their research emphasized the significance of scrutinizing unsuccessful attempts to improve future outcomes. Nine SRs in the corpus discussed the challenges faced using the FC model. We used a categorization scheme developed by Betihavas et al. (2016) that classified the challenges of FC into three categories: student-related, faculty-related, and operational. This categorization was used to code the challenges related to FC model. Table 10 displays the list of coded challenges and sample review studies.

**Table 10**

*Overview of the Challenges of Flipped Classrooms*

| Challenge | Subcategory | Reviews (n) | Example review |
|---|---|---|---|
| Student-related | Not completing pre-class tasks | 7 | Baltaci (2022) |
| | Disorientation | 3 | Ekici (2021) |
| | Resistance and lack of motivation | 6 | Gianoni-Capenakas et al. (2019) |
| | Time consumption and workload | 8 | Gerber & Eybers (2021) |
| Faculty-related | Effort and workload | 8 | Youhasan et al. (2021) |
| | Teacher's preparation | 4 | Park et al. (2021) |
| | Resistance to role change | 2 | Koh (2019) |
| Operational | Technical problems | 5 | Lo & Hew (2022) |
| | Requiring adjustment time | 3 | Divjak et al. (2022)<br>Fung et al. (2021)<br>Inayah et al. (2023)<br>Ekici (2022) |

### *Student-Related Challenges*

Implementing the FC model can be challenging due to various factors hindering its success. One of the main obstacles is the increased workload and time required for preparation from both students and faculty. Review studies such as those conducted by Gerber & Eybers (2021), Baltaci (2022), and Divjak et al. (2022) show that the FC requires more effort and workload from the students than traditional classroom models. Some students struggle with pre-class individual learning (Cevikbas & Kaiser, 2022a). Therefore, adequate planning and preparation are necessary to ensure success. The SR by Gerber and Eybers (2021) also highlighted that some students may need help accessing flipped materials at home, hindering the practice of the approach. Another significant challenge is students' lack of engagement with the model and failure to complete assignments. Teachers must, therefore, provide proper guidance and support to help students adapt to the new instructional model. Baltaci (2022) also emphasized that implementing an FC model requires assigning time for flipped instruction outside the classroom. This can be challenging for students with other commitments.

### *Faculty-Related Challenges*

The FC workload is stressful for educators (Lo & Hew, 2022) as they plan lessons and create or adapt content to help students learn and prepare for class (Cevikbas & Kaiser, 2022a). Divjak et al. (2022) identified a further disadvantage of FCs, stating that teachers had to invest much time in preparation and communication. For example, carefully designing asynchronous and synchronous stages, preparing rich resources, and planning time to answer questions require substantial time from teachers. Baltaci (2022) mentioned also the lack of immediate student assistance and the need for extensive resources and material development as drawbacks of flipped instruction. This places an extra burden on teachers.

In addition, teachers' competence and preparation are among the frequently reported faculty-related challenges. Even if a teacher is highly competent in face-to-face delivery, the quality of the video lessons he/she prepares may be poor due to insufficient instructional or technical skills, as highlighted by Gerber and Eybers (2021). They recommended using MOOCs in FCs to support with pre-class content development, especially in resource-challenged countries, keeping in mind that the use of MOOCs or online video content may pose a challenge in ensuring alignment with the overall curricula, module objectives, and outcomes. Also, technical issues could arise due to inadequate teacher training (Ekici, 2021). Some teachers may also resist the change in their teacher-centric roles, hindering the learner centrism at the core of the FC model (Koh, 2019).

### *Operational Challenges*

Implementing flipped learning has brought challenges, mainly related to technology and infrastructure (Baltaci, 2022), including inadequate resources, limited Internet access, and insufficient technical support (Arslan, 2020; Lo & Hew, 2022). As a result, teachers and students may have difficulty accessing and sharing materials online. Ekici (2021) pointed out that the challenges of flipped learning are compounded by the fact that some students may need more equipment to participate in online learning, leading to website inaccessibility or inequality in accessing technology (digital divide). Therefore, it is crucial to ensure all students have access to the necessary resources.

## Discussion and Recommendations

Our umbrella review found that research on flipped classrooms covered various areas but could be grouped into three main categories (see Table 6). The main focus was to evaluate FCs' effectiveness and

understand how teaching strategy could be enhanced to better accommodate the FC model. There are several key factors contributing to the success of the FC model. One of these factors is the compatibility of the approach with the specific context; greater compatibility will result in more successful long-term implementation. Therefore, it is important to consider the overall instructional system design rather than just focusing on specific operational aspects. Additionally, achieving positive results in FCs requires balancing the workload of activities before, during, and after class instead of placing too much emphasis on only one phase (Youhasan et al., 2021).

However, the current focus is mainly on designing pre-class and post-class activities, with less attention to in-class time. Future research should consider using design-based methodologies (Inayah et al., 2023) and promoting action research to maximize the positive impact of in-class interventions on student learning in FCs. A consistent and holistic approach will help ensure that all activities support each other and actively engage the learners.

Flipped classrooms, particularly in higher education, have been widely adopted to support student-centered learning. However, FC has been criticized for lacking a clear theoretical framework (Inayah et al., 2023; Koh, 2019). In our analysis, we found that several reviews needed a solid theoretical foundation to support their research focus and practices. Other reviews, such as Wohlfart and Wagner (2023), identified a similar issue. The absence of explicit references to educational and learning theories and the wide range of student-centered pedagogical practices may present challenges for instructors interested in implementing FCs. These reviews do not provide sufficient evidence to demonstrate the effectiveness of strategies, which raises concerns about the current theoretical foundation of FC research. While constructivism predominantly prevails, other theories were identified in this umbrella review.Understanding the theoretical basis underlying different approaches to FCs is vital to establishing effective pedagogical strategies for replication. Hence, this umbrella review highlights the necessity of employing a clearly defined theoretical framework that aligns with contemporary learning theories to steer the development, implementation, and assessment of FCs. By contemporary learning theories, we mean theories that reflect on the dynamics of the digital era (e.g., connectivism), since FC is a model that pays great attention to the quality of interaction in the learning and teaching processes.

A growing body of evidence shows the impact of general and additional design considerations of the FC. In addition to a handful of design features commonly reported in most studies that may be related to the flipped approach's effectiveness, the SRs we analyzed have provided insights for improving the implementation of FCs (as detailed in the results of RQ4 and RQ5). The effectiveness of the FC could be enhanced by considering the time students spend in class and the duration of the flipped intervention (Xu et al., 2019). While flipped learning studies vary in the number and type of pedagogical details they provide (Låg & Sæle, 2019), it has been challenging to identify the specific aspects of flipped learning that most influence its success. This challenge stems from a lack of systematic reporting (Låg & Sæle, 2019). Similarly, we found only a few eligible studies that provided the necessary information to analyze these variables reliably. To address this issue, future studies on the effectiveness of flipped learning need to consistently, thoroughly, and systematically report implementation details, design, and pedagogical features as suggested in this umbrella review.

Also, one possible reason for the heterogeneity and variations in the outcomes of flipped learning is the difference in context. The academic discipline may play a role in the success of flipped models. The argument is that applied STEM fields, such as engineering or the health professions, may be particularly well-suited to flipped learning (Bredow et al., 2021). Subject areas that require practicing specific skills

or problem sets, such as languages and mathematics, have also been identified as suitable for this method (Bredow et al., 2021; Karabulut-Ilgu et al., 2018). Our umbrella review revealed unique recommendations for effectively implementing FCs in the most popular disciplines. Nevertheless, the prevalence of FCs in STEM disciplines has raised questions about whether the nature of these disciplines might have influenced their design. If FCs became equally popular in the social sciences, alternative design approaches might be explored (Hew et al., 2021). We strongly encourage researchers and practitioners to critically discuss and evaluate different strategies and apply them across various fields and subject areas to maximize their effectiveness.

Despite the widespread use of the FC model, its full potential has yet to be realized (Hoshang et al., 2021). Most reviews have focused on studies with short FC interventions (e.g., one academic semester or less) and have been based on small sample sizes (e.g., fewer than 100 participants), as illustrated in Table 9. This suggests that some findings regarding the opportunities offered by the FC model may need further investigation. Longitudinal studies are necessary to identify best practices for implementing the FC model and to determine whether contextual factors, such as subject areas, academic levels, the complexity of learning content, and duration of the flipped intervention, have any impact on its effectiveness (Lo & Hwang, 2018).

Another significant gap in the FC research is the need for more thorough reflections on the challenges of this model. Upon analyzing SRs investigating FCs, it was discovered that many of the same challenges are reported and repeated in existing literature (e.g., Li & Fang, 2015). However, the challenges remain general and lack specific details on how they might differ across disciplines, potentially leading to redundant efforts. Planning customized activities and FC designs without this information is challenging. The literature offers some recommendations to address these challenges, including providing sufficient training and support for teachers and students, ensuring all students have access to necessary resources, and fostering a supportive learning environment that promotes engagement and participation among students. In addition, MOOCs are suggested to complement teachers' lack of capability to create their own videos. We also recommend using open education resources (OER) for the same purpose besides content preparation.

Regarding gaps in flipped learning research, there is a need for reviews that diverge from the typical research on flipped learning, which focuses on micro-level (teaching and learning) aspects. Studies encompassing macro-level (systems, theory, methods, global perspectives) and meso-level (institutional management and organization) aspects will broaden the base of knowledge on FL. This would offer a more comprehensive understanding of this subject matter.

Finally, it is strongly recommended that factors that can influence the implementation process and affect students' preferences, such as students' cultural orientations and educational traditions, be considered. Understanding cultural differences is critical for the successful and sustainable implementation of the FC method (Karagöl & Esen, 2019). For instance, some cultures may prefer strict teacher guidance and authority, which could impact the efficacy of flipped pedagogies.

# Conclusion

This umbrella review examined the potential advantages of flipped learning in promoting student-centered learning. Flipped classrooms can be an effective method for enhancing learning results, as long as they are customized to address the particular requirements of different disciplines and situations.

The review presented recent trends and patterns in FC research and practice in education. It also highlighted key challenges that need to be addressed to fully benefit from this approach, such as the limited evidence of the impact of FCs beyond general outcomes. Enhanced theorization of FC research and practice would facilitate the effective implementation of this approach. Customized design and a more profound exploration of the alignment of flipped classrooms' design with subject areas and contexts were among the critical topics addressed in this review. In addressing these issues, we hope this umbrella review provides insights for educators and researchers seeking to optimize the use of FCs in teaching and learning environments.

# References

(an asterisk indicates systematic reviews included in this umbrella review)

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education*, *126*, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

Aromataris, E., Fernandez, R. S., Godfrey, C., Holly, C., Khalil, H., & Tungpunkom, P. (2014). Methodology for JBI umbrella reviews. In *Joanna Briggs Institute Reviewers' Manual: 2014 Edition, Supplement* (pp. 1–34). The Joanna Briggs Institute. https://ro.uow.edu.au/articles/chapter/Methodology_for_JBI_umbrella_reviews/27794958?file=50564202

*Arslan, A. (2020). A systematic review on flipped learning in teaching English as a foreign or second language. *Journal of Language and Linguistic Studies, 16*(2), 775–797. https://doi.org/10.17263/jlls.759300

Bagley, S. (2020). The flipped classroom, lethal mutations, and the didactical contract: A cautionary tale. *PRIMUS, 30*(3), 243–260. https://doi.org/10.1080/10511970.2018.1555196

*Baltacı, H. S. (2022). A snapshot of flipped instruction in English Language teaching in Turkiye: A systematic review. *Turkish Online Journal of Distance Education, 23*(4), 255–269. https://doi.org/10.17718/tojde.1182793

*Banks, L., & Kay, R. (2022). Exploring flipped classrooms in undergraduate nursing and health science: A systematic review. *Nurse Education in Practice*, *64*, Article 103417. https://doi.org/10.1016/j.nepr.2022.103417

Bergmann, J., & Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day* (pp. 120-190). Washington DC: International Society for Technology in Education.

Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for "flipping out": A systematic review of the flipped classroom in nursing education. *Nurse Education Today*, *38*, 15–21. https://doi.org/10.1016/j.nedt.2015.12.010

Birgili, B., Seggie, F. N., & Oğuz, E. (2021). The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis. *Journal of Computers in Education, 8*(1), 365–394. https://doi.org/10.1007/s40692-021-00183-y

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *2013 ASEE annual conference and exposition, Conference proceedings*. American Society for Engineering Education. https://www.researchgate.net/publication/285935974_The_flipped_classroom_A_survey_of_the_research

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K–12: A systematic review. *Computers & Education*, *151*, Article 103819. https://doi.org/10.1016/j.compedu.2020.103819

Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, *7*(2), Article e012545. https://doi.org/10.1136/bmjopen-2016-012545

Braun V., & Clarke V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2),77–101. https://doi.org/10.1191/1478088706qp063oa

Bredow, C. A., Roehling, P. V., Knorp, A. J., & Sweet, A. M. (2021). To flip or not to flip? A meta-analysis of the efficacy of flipped learning in higher education. *Review of Educational Research, 91*(6), 878–918. https://doi.org/10.3102/00346543211019122

Brunton, J., & Thomas, J. (2012). Information management in systematic reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 83–106). SAGE.

*Cevikbas, M., & Kaiser, G. (2022a). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. *ZDM Mathematics Education, 55*(1), 177–191. https://doi.org/10.1007/s11858-022-01388-w

*Cevikbas, M., & Kaiser, G. (2022b). Promoting personalized learning in flipped classrooms: A systematic review study. *Sustainability, 14*(18), Article 11393. https://doi.org/10.3390/su141811393

Crompton, H., Burke, D., & Gregory, K. H. (2017). The use of mobile learning in PK–12 education: A systematic review. *Computers & Education*, *110*, 51–63. https://doi.org/10.1016/j.compedu.2017.03.013

*Divjak, B., Rienties, B., Iniesto, F., Vondra, P., & Žižak, M. (2022). Flipped classrooms in higher education during the COVID-19 pandemic: Findings and future research recommendations. *International Journal of Educational Technology in Higher Education*, *19*, Article 9. https://doi.org/10.1186/s41239-021-00316-4

El Gamal, H. (2022). Is flipped approach a panacea? A systematic review of trends, conceptions, and practices of a decade of research. Asian Journal of Distance Education, 17(2),153-180. https://doi.org/10.5281/zenodo.7245601

*Ekici, M. (2021). A systematic review of the use of gamification in flipped learning. *Education and Information Technologies*, *26*, 3327–3346. https://doi.org/10.1007/s10639-020-10394-y

Eppard, J., & Rochdi, A. (2017). A framework for flipped learning. In I. Arnedillo Sánchez, P. Isaias, & L. Rodrigues (Eds.), *Proceedings of the International Conference on Mobile Learning 2017* (pp. 33–40). International Association for Development of the Information Society. https://www.proceedings.com/content/036/036166webtoc.pdf

*Evans L., Vanden Bosch, M. L., Harrington, S., Schoofs, N., & Coviak C. (2019). Flipping the classroom in health care higher education: A systematic review. *Nurse Educator*, *44*(2), 74–78. https://doi.org/10.1097/NNE.0000000000000554

*Farmus, L., Cribbie, R. A., & Rotondi, M. A. (2020). The flipped classroom in introductory statistics: Early evidence from a systematic review and meta-analysis. *Journal of Statistics Education*, *28*(3), 316–325. https://doi.org/10.1080/10691898.2020.1834475

*Fernández-Martín, F.D., Romero-Rodríguez, J.-M., Gómez-García, G., & Ramos Navas-Parejo, M. (2020). Impact of the flipped classroom method in the mathematical area: A systematic review. *Mathematics*, *8*(12), Article 2162. https://doi.org/10.3390/math8122162

*Fung, C.-H., Besser, M., & Poon, K.-K. (2021). Systematic literature review of flipped classroom in mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 17*(6), Article em1974. https://doi.org/10.29333/ejmste/10900

*Gerber, A., & Eybers, S. (2021). Converting to inclusive online flipped classrooms in response to COVID-19 lockdown. *South African Journal of Higher Education, 35*(4), 34–57. https://doi.org/10.20853/35-4-4285

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). SAGE.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews, 1*, Article 28. https://doi.org/10.1186/2046-4053-1-28

* Gianoni-Capenakas, S., Lagravere, M., Pacheco-Pereira, C., & Yacyshyn, J. (2019). Effectiveness and Perceptions of Flipped Learning Model in Dental Education: A Systematic Review. *Journal of dental education*, *83*(8), 935–945. https://doi.org/10.21815/JDE.019.109Giannakos, M. N., Krogstie, J., & Chrisochoides, N. (2014). Reviewing the flipped classroom research: Reflections for computer science education. In E. Barendsen & V. Dagiené (Eds.), *Proceedings: Computer Science Education Research Conference '14* (pp. 23–30). ACM. https://doi.org/10.1145/2691352.2691354

*Gianoni-Capenakas, S., Lagravere, M., Pacheco-Pereira, C., & Yacyshyn, J. (2019). Effectiveness and perceptions of flipped learning model in dental education: A systematic review. *Journal of Dental Education*, *83*(8), 935–945. https://doi.org/10.21815/JDE.019.109

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hammersley, M. (2020). Reflections on the methodological approach of systematic reviews. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research* (pp. 23–39). Springer VS. https://doi.org/10.1007/978-3-658-27602-7_2

*Hendrik, H., & Hamzah, A. (2021). Flipped classroom in programming course: A systematic literature review. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(02), 220–236. https://doi.org/10.3991/ijet.v16i02.15229

Hew, K. F., Bai, S., Huang, W., Dawson, P., Du, J., Huang, G., Jia, C., & Thankrit, K. (2021). On the use of flipped classroom across various disciplines: Insights from a second-order meta-

analysis. *Australasian Journal of Educational Technology*, *37*(2), 132–151. https://doi.org/10.14742/ajet.6475

Hoshang, S., Hilal, T. A., & Hilal, H. A. (2021). Investigating the Acceptance of Flipped Classroom and Suggested Recommendations. *Procedia Computer Science*, *184*, 411–418. https://doi.org/10.1016/j.procs.2021.03.052

Hwang, G.-J., Yin, C., & Chu, H.-C. (2019). The era of flipped learning: Promoting active learning and higher order thinking with innovative flipped learning strategies and supporting systems. *Interactive Learning Environments 27*(8), 991–994. https://doi.org/10.1080/10494820.2019.1667150

*Inayah, S., Juandi, D., Darhim, D., Prabawanto, S., & Jupri, A. (2023). The impact of flipped classroom implementation in mathematics learning at school: Systematic literature review. *Journal Analisa*, *9*(1), 59–73. https://doi.org/10.15575/ja.v9i1.18226

Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. *American Journal of Distance Education*, *9*(2), 7–25. https://doi.org/10.1080/08923649509526885

Karabulut-Ilgu, A., Jaramillo Cherrez, N., & Jahren, C. T. (2018). A Systematic Review of Research on the Flipped Learning Method in Engineering Education. *British Journal of Educational Technology, 49*, 398-411. https://doi.org/10.1111/bjet.12548

Kapur, M., Hattie, J., Grossman, I., & Sinha, T. (2022). Fail, flip, fix, and feed—Rethinking flipped learning: A review of meta-analyses and a subsequent meta-analysis. *Frontiers in Education, 7*, Article 956416. https://doi.org/10.3389/feduc.2022.956416

*Karagöl, I., & Esen, E. (2019). The effect of flipped learning approach on academic achievement: A meta-analysis study. *Hacettepe University Journal of Education*, *34*(3), 708–727. https://doi.org/10.16986/HUJE.2018046755

*Kazeminia, M., Salehi, L., Khosravipour, M., & Rajati, F. (2022). Investigation of flipped classroom effectiveness in teaching anatomy: A systematic review. *Journal of Professional Nursing*, *42*, 15–25. https://doi.org/10.1016/j.profnurs.2022.05.007

*Koh, J. H. L. (2019). Four pedagogical dimensions for understanding flipped classroom practices in higher education: A systematic review. *Educational Sciences: Theory and Practice*, *19*(4), 14–33. https://jestp.com/menuscript/index.php/estp/article/view/756/709

Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *AERA Open, 5*(3), 1–17. https://doi.org/10.1177/2332858419870489

Lee, K., Zawacki-Richter, O., & Cefa Sari, B. (2022). A systematic literature review on technology in online doctoral education. *Studies in Continuing Education*, *46*(1), 38–64. https://doi.org/10.1080/0158037X.2022.2135499

Lewin, K. (1952). *Field theory in social science: Selected theoretical papers* (D. Cartwright, Ed.). Harper and Brothers.

Li, S., & Fang, D. (2015). Factors of college students in the flipped classroom. In F. Khoshnoud & T. Li (Chairs), *2015 10th International Conference on Computer Science & Education* (pp. 805–809). IEEE. https://doi.org/10.1109/ICCSE.2015.7250355

Lin, H.-C., & Hwang, G.-J. (2019). Research trends of flipped classroom studies for medical courses: A review of journal publications from 2008 to 2017 based on the technology-enhanced learning model. *Interactive Learning Environments, 27*(8), 1011–1027. https://doi.org/10.1080/10494820.2018.1467462

Liu, L., Hew, K. F., & Du, J. (2024). Design principles for supporting self-regulated learning in flipped classrooms: A systematic review. *International Journal of Educational Research, 124*, Article 102319. https://doi.org/10.1016/j.ijer.2024.102319

Lo, C. K. (2020). Systematic reviews on flipped learning in various education contexts. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic reviews in educational research* (pp. 129–143). Springer VS. https://doi.org/10.1007/978-3-658-27602-7_8

Lo, C. K. (2023). Strategies for enhancing online flipped learning: A systematic review of empirical studies during the COVID-19 pandemic. *Interactive Learning Environments*, *32*(7), 3517–3545. https://doi.org/10.1080/10494820.2023.2184392

*Lo, C.K., & Hew, K.F. (2022).Design principles for fully online flipped learning in health professions education: a systematic review of research during the COVID-19 pandemic. *BMC Medical Education, 22* (720). https://doi.org/10.1186/s12909-022-03782-0

Lo, C. K., & Hew, K. F. (2017). A critical review of flipped classroom challenges in K–12 education: Possible solutions and recommendations for future research. *Research and Practice in Technology Enhanced Learning, 12*(1), Article 4. https://doi.org/10.1186/s41039-016-0044-2

Lo, C. K., & Hwang, G. J. (2018). How to advance our understanding of flipped learning: Directions and a descriptive framework for future research. *Knowledge Management & E-Learning, 10*(4), 441–454. https://doi.org/10.34105/J.KMEL.2018.10.027

Lundin, M., Bergviken Rensfeldt, A., Hillman, T. et al. (2018). Higher education dominance and siloed knowledge: a systematic review of flipped classroom research. *International Journal of Education Technology in High Education, 15*(20). https://doi.org/10.1186/s41239-018-0101-6

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R., Allard, S., & Levine, K. (2015). Peer review: Still king in the digital age. *Learned Publishing, 28*(1), 15–21. https://doi.org/10.1087/20150104

*Özbay, Ö., & Çınar, S. (2021). Effectiveness of flipped classroom teaching models in nursing education: A systematic review. *Nurse Education Today*, *102*, Article 104922. https://doi.org/10.1016/j.nedt.2021.104922

*Park, J. H., Han, W. S., Kim, J.,& Lee, H. (2021). Strategies for flipped learning in the health professions education in South Korea and their effects: A systematic review. *Education Sciences, 11*(1), Article 9. https://doi.org/10.3390/educsci11010009

Pulley, P. G. (2014). Blending face-to-face and technology: Implementing flipped K–12 classrooms. In L. Kyei-Blankson & E. Ntuli (Eds.), *Practical applications and experiences in K–20 blended learning environments* (pp. 105–119). IGI Global.

Rasheed, R. A., Kamsin, A., Abdullah, N. A., Kakudi, H. A., Ali, A. S., Musa, A. S., & Yahya, A. S. (2020). Self-regulated learning in flipped classrooms: A systematic literature review. *International Journal of Information and Education Technology*, *10*(11), 848–853. https://doi.org/10.18178/ijiet.2020.10.11.1469

Rees, R., & Oliver, S. (2012). Stakeholder perspectives and participation in reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 17–34). Sage.

Sahin, M. C. (2009). Instructional design principles for 21st century learning skills. *Procedia—Social and Behavioral Sciences, 1*(1), 1464–1468. https://doi.org/10.1016/j.sbspro.2009.01.258

Şensöz, M. Z., & Erdemir, N. (2022). A systematic review of graduate studies on flipped classrooms in English language teaching in Turkey. *ELT Research Journal, 1*(1), 1–15.

Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine, 99*(4), 178–182. https://doi.org/10.1258/jrsm.99.4.178

The Joanna Briggs Institute. (2014). *Methodology for JBI umbrella reviews*. The Joanna Briggs Institute.

Wohlfart, O., & Wagner, I. (2023). Teachers' role in digitalizing education: An umbrella review. *Educational Technology Research and Development*, *71*(2), 339–365. https://doi.org/10.1007/s11423-022-10166-0

Xiao, J. (2023). Introduction to history, theory, and research in ODDE. In O. Zawacki-Richter, & I. Jung (Eds.), *Handbook of open, distance and digital education: Towards an informed approach to ODDE* (pp. 15–25). Springer. https://doi.org/10.1007/978-981-19-2080-6_1

*Xu, P., Chen, Y., Nie, W., Wang, Y., Song, T., Li, H., Li, J., Yi, J., & Zhao, L. (2019). The effectiveness of a flipped classroom on the development of Chinese nursing students' skill competence: A systematic review and meta-analysis. *Nurse Education Today*, *80*, 67–77. https://doi.org/10.1016/j.nedt.2019.06.005

*Youhasan, P., Chen, Y., Lyndon, M., & Henning, M. A. (2021). Exploring the pedagogical design features of the flipped classroom in undergraduate nursing education: A systematic review. *BMC Nursing*, *20*(1), Article 50. https://doi.org/10.1186/s12912-021-00555-w

Zhou, M., & Li, Z. (2018). Blended mobile learning in theatre arts classrooms in higher education. *Innovations in Education and Teaching International*, *56*(3), 307–317. https://doi.org/10.1080/14703297.2018.1447389

Zawacki-Richter, O. (2009). Research Areas in Distance Education: A Delphi Study. *The International Review of Research in Open and Distributed Learning*, *10*(3). https://doi.org/10.19173/irrodl.v10i3.674

Zawacki-Richter, O. & Naidu, S. (2016). Mapping Research Trends from 35 Years of Publications in "Distance Education". *Distance Education, 37*(3), 245-269. https://doi.org/10.1080/01587919.2016.1185079

Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (2020). *Systematic reviews in educational research: Methodology, perspectives and application*. Springer. http://link.springer.com/10.1007/978-3-658-27602-7

Zawacki-Richter, O., & Bozkurt, A. (2023). Research Trends in Open, Distance, and Digital Education. In O. Zawacki-Richter & I. Jung (Eds.), *Handbook of Open, Distance and Digital Education* (pp. 199– 220). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-2080-6_12

Zawacki-Richter, O., Cefa, B., & Bai, J. Y. H. (2025). Towards reproducible systematic reviews in Open, Distance, and Digital Education (ODDE) – an umbrella mapping review. *Review of Education*. https://doi.org/10.1002/rev3.70031

*Zou, D., Luo, S., Xie, H., & Hwang, G.-J. (2020). A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning*, *35*(8), 1811–1837. https://doi.org/10.1080/09588221.2020.1839502

Zuber, W. J. (2016). The flipped classroom, a review of the literature. *Industrial and Commercial Training, 48*(2), 97–103. https://doi.org/10.1108/ICT-05-2015-0039


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Book Review: *Resisting the Dehumanization of Refugees*

**Editors:** Yasmeen Abu-Laban, Michael Frishkopf, Reza Hasmath, and Anna Kirova (AU Press, 2024, 420 pages). ISBNs: 978-1-771-99410-1 (paperback), 978-1-771-99411-8 (PDF), 978-1-771-99412-5 (epub).
https://doi.org/10.15215/aupress/9781771994101.01

**Reviewed by:** Staci B. Martin, EdD, *Portland State University, Portland, Oregon, United States*

On a bus to the Geneva Airport, four Swiss police officers boarded, asking only my seatmate and me for tickets. We stood out as the only two individuals of color among the rest of the passengers on board. The police officer challenged my seatmate's pass, noting the lack of a name, to which he, speaking with an American accent, explained he received it from his hotel. After they exited, I remarked, "I hope I'm on the right bus." He smiled and replied, "Me too." I discovered he worked for the U.S. State Department. He asks me where I am from. I say, "Oregon." I ask him where he is from. He says, "Virginia." We both carry on a longer conversation, never repeating the words. No, really, where are you from? We continued our conversation, implicitly understanding we share an American identity without needing to question each other's origins.

I share this story as it illustrates the many layers of navigation when it comes to traveling to a foreign country that one day you may call home. The word "home" is a term that is used in general terms; however, it often encompasses the emotional and personal meaning that includes feelings of belonging and familiarity.

*Resisting the Dehumanization of Refugees*, edited by Abu-Laban, Frishkopf, Hasmath, and Kirova, has a familiar quality when I read it. The book explores the profound struggle to cocreate a sense of home and belonging amidst the relentless dehumanization refugees face from restrictive policies, biased media portrayals, and societal prejudice. It adopts a transdisciplinary approach, featuring contributions from scholars, practitioners, and refugees themselves, who focus on the Canadian context and draw on broader international comparisons that support both the mechanisms of dehumanization and strategies for resistance.

The book is structured into four parts, each comprising three to five chapters that explore a particular theme. In the first two parts, the authors explain how immigration policies and the media systemically dehumanize refugees, while educational institutions can either exacerbate or mitigate this. The last two sections explore rehumanization through state apologies. The chapters emphasize the vital role of refugee agency and artistic expression in building a sense of home.

Part 1 focuses on how immigration policies and media narratives systemically dehumanize refugees. In turn, these policies create vulnerability and precarity for refugees. The policies and media reinforce the negative stereotype narrative.

Abu-Laban's chapter compares the immigration policies of and public discourse in Canada, the United States, and Australia. It also examines how these factors contribute to either the dehumanization or humanization of refugees in the three countries. This chapter lays the foundation for the book by defining dehumanization as the act of viewing some individuals as fully human while perceiving others as limited or inferior.

Ayres's chapter explores the "double movement," which the author contends is the interdependency of state policies, media, and social action that might create migrant and refugee precarity. There is also a countermovement that emerges from collective action and mobilization on the Canada–United States borderlands. This chapter is intriguing as it offers a glimpse of the collective protests between bordering countries that humanize the migrant and refugee experience.

Birjandian's chapter is the first of several chapters told from a first-person narrative perspective. This chapter offers a rare glimpse into a refugee's journey, detailing the arduous steps of resettlement (e.g., asylum process, resettlement, and integration) into a new country.

Khasanova's chapter examines how Canadian print media depicts Syrian refugees, particularly after the 2015 Paris attacks, resulting in their dehumanization. The chapter critiques the media's misrepresentation of Islam through three dominant narratives: Islam as Other, Islam as anti-Semitic, and the portrayal of Islamic/Muslim terrorism.

Part 2 examines the dual role of educational institutions, which can either exacerbate or counter dehumanization depending on their approach to refugee students' needs.

Bhattacharyya, Songose, and Wilkinson's chapter explores the experiences of Yazidi refugee families with children. They focus on how families engage with the Canadian education system and how that system either fails or facilitates humanization and integration. They advocate for more multifaceted approaches to create more accessible and supportive environments, thereby lessening refugees' transition and strengthening their integration.

Jalal's chapter is another first-person narrative that explores the challenges faced by refugee youth as they navigate between two cultures—the host and refugee contexts. This chapter poses the seemingly innocent question: "Where are you from?" As seen in this chapter, it is not so innocent and raises questions about where one belongs and the location of their identity, particularly for young people. This chapter is significant, as it served as the impetus for why I shared my own story at the beginning of this review.

Gutiérrez Rodríguez's chapter examines the effectiveness of transcultural learning experiences for asylum seekers in an academic setting in Germany. It highlights strengths, such as course accessibility, cooperative learning environments, and place-based instruction, as well as addressing the negative factors that are

brought on by power imbalances in day-to-day living issues (e.g., inaccessible transportation, restricted movement due to status) and German language proficiency, as well as limited access to university resources.

Kirova's chapter poses the question of how refugee children experience school in Canada, which often fails to meet their needs adequately. The chapter concludes by considering how educational theorists can develop a framework for humanizing education, drawing on the principles of intercultural education. Kirova promotes this approach as one that cultivates understanding, respect, and equitable relations across diverse groups through meaningful dialogue, critical thinking, and a sense of belonging. This approach means moving beyond mere cultural awareness to challenge inequities actively.

Part 3 explores strategies for rehumanization and supports the role of state apologies in acknowledging past wrongs so that reconciliation can happen.

Bakan's chapter examines the complex "politics of apology" through the lens of Canada's apology to Jewish refugees. The chapter explains that state apologies often serve to redeem or favor the apologizer rather than create reparations that support rehumanization. Bakan, as a Jewish person whose parents were the children of refugees, moves her critique to contemporary issues, emphasizing that efforts to hold Israel accountable for its actions are often unfairly dismissed as anti-Semitic, thereby silencing critical Palestinian voices.

Hasmath, Ho, and Kay-Reid's chapter explains how state apologies can hinder or contribute to the rehumanization of refugee, Indigenous, and ethnic minority groups. They argue that for an apology to be reparative, it must include recognition of the harm done, a commitment(s), and accountability with tangible redress.

Gurnett's chapter offers a person-centered perspective on displacement. The chapter emphasizes the importance for refugees of fostering hope, inclusion, and a sense of "home" as a means of countering dehumanization. The chapter critically examines how Canada often highlights only the success stories of refugees in its refugee policies. In sharing just the success stories, systemic shortcomings and the lived realities of displaced individuals are often overshadowed.

Part 4 provides spaces for cocreating meaningful collaborative action in the resistance to dehumanization through artistic expression. This action focuses on the role of refugees' agency, self-determination, and belonging. Individuals can reclaim their narratives, assert their humanity, and build a sense of home.

Barzanji's chapter presents a first-person narrative of a writer in exile. The author highlights how artistic expression and writing can be used to resist dehumanization while asserting one's humanity. This chapter is significant in that it shows how poetry can transform personal suffering and imprisonment into universal themes of humanity (e.g., persistence, courage, resistance, and hope).

Harrington and Waissi's chapter explores how Kurdish refugees in Canada negotiate and express their cultural identity. The chapter conveys a message of hope by demonstrating the resilience of Kurdish culture and identity within the diaspora through cultural preservation and expression.

Frishkopf's chapters highlight the role of music in rehumanizing refugees, exploring its therapeutic impact on the psychological trauma of displacement. Mapfumo and Chikawa showcase music as a tool for political

change. Frishkopf's concluding insights then offer practical applications for enacting rehumanization through musical initiatives.

*Resisting the Dehumanization of Refugees* brought me to familiar places in my teaching (e.g., supporting third spaces, higher education in protracted contexts), research (e.g., community-based action research and art-based approaches), and critical hope and despair peace-building scholarship. As a reader, I appreciated the first-person narratives, the different formats of the chapters (e.g., poetry and interviews), and the profound hope of humanizing refugees in the face of restrictive policies, biased media portrayals, and societal prejudice.

This is a timely and yet timeless book that showcases how refugees reclaim their narratives, preserve their cultural heritage, and utilize creative outlets to convey their experiences, resilience, and unique perspectives. This book is written in a way that is accessible and that aims to foster empathy and understanding among host and refugee communities. I believe the book effectively challenges stereotypes and offers practical solutions to combating Othering that contributes to dehumanization.


**I**nternational **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 26, Number 4

**November – 2025**

# Book Review: *Brave New Words—How AI Will Revolutionize Education (and Why It's a Good Thing)*

**Author:** Salman Khan (Viking, 2024, 272 pages), ISBN: 9780593656952 (hardcover), 9780593656969 (e-book)
https://doi.org/10.1080/00131911.2025.2533064

**Reviewed by:** Taoufik Boulhrir, *Fordham University*

## Introduction

Salman Khan's *Brave New Words* emerges at a pivotal moment in educational history, when artificial intelligence (AI) is alternately celebrated as a transformative force and denounced as a threat to the human dimensions of teaching. Drawing on his rich experience as the founder of Khan Academy, a globally renowned non-profit organization producing freely available educational videos and exercises, Khan weaves a conversational narrative that eschews dry technical jargon in favor of vivid case studies and practitioner anecdotes. This book is geared toward readers who are new to AI, whether as a parent, teacher, or education policy maker. Rather than serving as a step-by-step guide or an exhaustive chronicle of AI's evolution, the book offers a practitioner's reflection on how emerging technologies can be adapted to align with institutional goals and real-world classrooms. Khan brings readers into his conversations with education innovators. The author's enthusiasm for AI technology in education sometimes outpaces a deeper engagement with its long-term social and pedagogical implications.

## A Casual Scholarly Conversation

Salman Khan writes with a casual conversational style that bridges the gap between AI scholar and layperson. Rather than cloaking his insights in dense academic prose, he opts for clear language and anecdotes modeling the pedagogical openness and accessibility that he advocates. Khan removes technical jargon so readers can focus on the book's proposition that the emerging uses of AI can positively transform education.

Khan incorporates interviews and concise storytelling vignettes from the perspectives of parents, classroom teachers, school administrators, and education researchers to document instances of AI experiments in use. These case-based accounts include adaptive math exercises deployed in under-resourced schools and AI-facilitated historical simulations in humanities courses. The author is transparent about his reliance on pilot studies, pilot program feedback, and his own institution's data, acknowledging their strengths as well as their limitations. Khan invites the reader to consider an open-ended exploration of what AI might become in the hands of thoughtful educators.

## Conversations With AI as a Pedagogical Choice

Throughout the book, Khan explores how AI can be designed as a "Socratic tutor," a pseudo-collaborator that prompts students with questions rather than supplying ready-made answers. Khan recounts failed attempts to get GPT-3.5 to perform as a Socratic tutor. It was not until GPT-4 was available that this could be achieved, and the book shares how by using the simple prompt: "You will be a Socratic tutor. I will be your student. Don't give me answers" (p. 31). Khan's approach to applying a dialogic tradition to the use of AI in education is recognizably aligned with constructivist scaffolding, to incrementally push the learner slightly beyond their current abilities.

The reliance on Socratic dialogue for educational use of AI begins to fray when we consider what human dialogue offers that an algorithm cannot. True Socratic dialogue is laden with emotional nuance, cultural reference points, and nonverbal cues, all of which an AI tutor can only approximate. While Khan shows how AI can simulate empathy by offering reassurance before a daunting problem, it lacks genuine affective understanding. Cultural idioms or community-specific values may be flattened into generic responses, and the absence of shared lived experience may leave machine-mediated exchanges feeling hollow. In this sense, AI conversations risk becoming transactional exchanges of prompts and replies, rather than rich, mutual explorations of ideas.

Might the overreliance on AI-driven questioning threaten to displace the relational heart of teaching? Human tutors attune not only to what students answer, but also to why they hesitate, what emotions surface, and how individual identities shape inquiry. Khan's own examples of memory-driven personalization, where the system recalls a student's love of soccer to frame future lessons, demonstrates the promise of adaptive learning. A strength of AI is in looking backwards at a massive amount of data. In this soccer-loving-student scenario, AI may correctly remember a past interest of a learner but have no access to data to show that the learner is no longer interested in soccer. As a pedagogical tool, AI conversations hold real potential, if educators critically calibrate its use to preserve empathy, context-sensitivity, and cultural richness that define human dialogue.

## Ethical and Equitable Use of AI in Education

The book's techno-optimism risks glossing over the reality that real-time AI interactions require substantial broadband connectivity and that global disparities in telecommunications infrastructures can render even the most advanced tools moot in under-resourced settings. Educational technologies are ever evolving. Using AI is not a panacea for resolving funding, policy, and systemic inequality in education.

AI reflects and reproduces existing power structures, decisions about which voices get digitized, whose histories are digitized, and whose cultural frameworks inform the defaults of machine learning models. Khan's *Brave New Words* portrays the assumption that bias can be engineered away through fair algorithms and the collection, reporting, and analysis of data. This reliance on AI technology may be an underestimation of the complex social and political concerns that exist in education.

## AI's Place in Educational Practice

Khan stresses that artificial intelligence will not replace teachers but rather empower them to engage in more meaningful educational interactions. He posits that AI would shoulder routine tasks, freeing educators to cultivate deeper mentoring relationships, personalized instruction, and creative curriculum design. Unwittingly, he echoes the techno-optimism advocated by proponents of new educational technologies throughout history.

Khan's commitment to individualized learning permeates *Brave New Words*, reflecting the origins of the Khan Academy in offering tailored video lessons to learners at their own pace. This book pays little attention to the educational roles of the relational dynamics of student-to-student or student-to-teacher interactions. He argues repeatedly that personalized pathways, driven by adaptive algorithms, unlock each student's unique potential more effectively than one-size-fits-all instruction. A reader might see Khan's pursuit as exactly that which he is rallying against, as he advocates building a better, more complex one-technology-for-all educational tool. This insistence on individualization privileges the solitary learner's journey and risks underplaying the value of human interactions.

Similarly, Khan's treatment of standardized testing and learner autonomy deserves academic perusal. The book presents testing data as a neutral metric of progress, often citing improvements in quiz scores following AI interventions, while dismissing broader critiques that standardized exams can narrow curricula and entrench inequities. Testing, at its best, provides actionable feedback; at its worst, it incentivizes teaching to the test and overlooks the holistic development of critical thinking and creativity. His support for data-driven assessments aligns with his emphasis on measurable gains but ignores how overreliance on test scores can distort educational priorities. Although autonomy is well established as an effective drive for learning, students require not only adaptive prompts but also opportunities to collaborate, debate, and co-construct knowledge, practices that standardized testing alone cannot capture.

## Closing Comments

First and foremost, Khan frames AI adoption as a challenge of an attitudinal leap. The book succeeds as a reflective, example-driven narrative that encourages educators to engage with AI not as passive consumers of technology but as active co-designers of new pedagogical pathways. His vivid case studies, from adaptive tutoring experiments to AI-assisted humanities simulations, offer concrete illustrations of how institutions' use of emerging tools can align with their core educational vision, mission, values, and purpose. Khan demonstrates an approach for using AI not only as a tool, but also as a digital collaborator in instructional design. In practice, deploying AI responsibly requires more than a willingness to try new tools; it demands structural literacy about policy contexts, critical awareness of ethical trade-offs, and ongoing reflection on how technology reshapes power and relationships in the classroom.

Looking ahead, several critical questions emerge from what is written in the book. First, which pedagogical values, formative assessment, autonomy, or socio-emotional learning, should guide AI integration into education? Second, how can educators safeguard their professional agency and maintain trust with students when AI assumes more automated roles in feedback, grading, or personalized support? Finally, what level of data collection and reporting is required for AI to effectively track progress and tailor responses without

infringing on student privacy? By addressing these questions, educators, technologists, and policymakers will inform how AI in education evolves in ways that may affirm both human dignity and pedagogical rigor. One question that arises from what is not written in the book is: "How might AI be used to initiate, build, and foster relationships between actors within the digital classroom setting, such as learner-to-learner?"

## Acknowledgements

I would like to express my sincere gratitude to Dr. Peggy Lynn MacIsaac for her insightful editorial support and thoughtful guidance; her careful scrutiny and encouragement were instrumental in shaping the clarity and depth of this review.